TIME DEPENDENT DENSITY FUNCTIONAL THEORY MODELING OF CHIROPTICAL PROPERTIES OF AMINO ACIDS IN SOLUTION

by

Matthew David Kundrat
August 26, 2008

A dissertation submitted to the
Faculty of the Graduate School of
the State University of New York at Buffalo
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Department of Chemistry

Acknowledgments

First of all I really must acknowledge those responsible for getting my research career started. I thank Professor Berny Schlegel for introducing me to computational chemistry, and for always being available to talk science with me, even during the times when no one else would. I'd also like to thank the friends I still have at Wayne State University for continuing to believe in me and thank all the friends I have made in Buffalo who wanted me to succeed. In addition I thank the members of the General Motors R&D staff; were it not for the work I did with them my publications list would not be quite as impressive as it is.

To the past and present members of the Autschbach research group at UB, including Boris Le Guennic, Mykhaylo Krykunov, Brendan Mort, Shaohui Zheng, Eva Zurek, Mariusz Sterzel, Aijun Ye, Mark Rudolph, Ajitha Devarajan, and Alexander Gaenko, I am grateful for your camaraderie, friendship and insight. Kudos go out to the UB CCR, for spoiling me with a first rate computing cluster, and to Marek Freindorf for helpful discussions regarding molecular dynamics. On the subject of MD, I cannot be more grateful to Marcel Swart of the Institut de Química Computational for his ongoing collaboration and endless feedback with his DRF-90 program. I would like to also thank our mutual friend Stan van Gisbergen of SCM for funding two of my trips to ACS conferences and introducing me to the inside world of computing software. Thanks also to Ben Wagner, for finding me those hard to get journal articles.

Of course, no thesis would be complete without acknowledging my committee members, Profs. Jim Garvey, Harry King, and Javid Rzayev for their time and constructive feedback. Finally, I thank my adviser, Professor Jochen Autschbach, for treating me not as a peon but as a colleague, even from day one. Were it not for his help and endless understanding I would not be where I am today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Time Dependent Density Functional Theory (TDDFT) and the COnductor-like Screening MOdel (COSMO) of solvation were used to model the specific rotation and Optical Rotatory Dispersion (ORD) of various amino acids in solution. Zwitterionic, cationic, dicationic, anionic and dianionic forms of amino acids were investigated and the results compared with experimental literature data obtained in neutral, acidic and basic conditions, as appropriate. It was found that TDDFT modeled the specific rotation of all the forms of the amino acids with an approximately equal level of accuracy. One source of error encountered is that the model overestimated the extent of intramolecular hydrogen bonding for the zwitterions, affecting the calculated mole fractions of the different conformers thus having an impact on the specific rotation. The physical origin of the Clough-Lutz-Jorgensen's rule was investigated, a general property of alpha amino acids where the specific rotation of an acidified solution is nearly always more positive in value than the specific rotation of the amino acid by itself. Results were consistent with the functional side chains of the amino acids behaving similarly to negatively charged groups when the molecules were in their zwitterionic forms, while those same groups behaved in a fashion akin to positively charged perturbers in the cationic, protonated forms of the molecules. Finally, the specific rotation of amino acids has be calculated using explicit solvation molecules, as opposed to the continuum model that has been used previously. It was found that simple three - point charge models used in place of quantum mechanical waters were able to reproduce, to an extent, the inductive effects of those solvent molecules on the chiroptical responses of the solute amino acids.

CHAPTER 1: INTRODUCTION, BACKGROUND AND GENERAL OUTLINE

Any molecule whose structure lacks a plane of symmetry, center of inversion or any other improper axis of rotation is chiral. Chiral compounds can form as two stereoisomers that are mirror images of one another, called "enantiomers". Such enantiomers are indistinguishable from one another by most standard analytical techniques; they have the same melting point, boiling point, IR spectra, NMR spectra, UV-Vis spectra and the same reactivity toward achiral reagents. However, different enantiomers do have differing reactivity toward other chiral substances. This is particularly important in biological chemistry, where one enantiomer of a substance may have a desired biological effect where its mirror image may be ineffective or even have an adverse size effect.

As such, it is important for a chemist to be able to reliably distinguish between opposing enantiomers. Perhaps the most robust method with which to make such assignment of "absolute configuration" is through X-ray diffraction (XRD). When it works the method reveals the arrangement of all the heavy (non-hydrogen) atoms in a crystal, which allows for an unambiguous assignment of absolute configuration. This method is not without its drawbacks. XRD equipment is rather expensive and requires specialized training to operate. The test is time consuming. Growing crystals that are suitable for analysis can be an arduous task. In fact some substances will not crystallize at all, and for those another method must be found to assign absolute configuration.

Fortunately, other methods exist for distinguishing between enantiomers. One such method is based on the fact that opposing enantiomers have the opposite interaction with a plane of polarized light. A solution of one enantiomer when placed in the path of a beam of polarized light will rotate that plane a given number of degrees, depending on the structure of the

molecule, the color of the light, the path length, the concentration and the temperature. A solution of that enantiomer's mirror image, measured under identical conditions, will rotate that same plane of polarized light by the same number of degrees, but in the opposite direction. This technique, known as polarimetry, has grown to become one of the methods that practically all chemists learn at the undergraduate level. The equipment needed is relatively inexpensive, widely available and easy to use. The actual measurement of the extent to which a solution of an enantiomer rotates a place of polarized light, the measurement of that molecule's "specific rotation" requires only a little more time than that spent preparing the solution to be measured.

The sign of the specific rotation that is measured is characteristic of the identity of the enantiomer. However there is no simple correlation between the observable response that is specific rotation and the internal molecular structure that defines its absolute configuration. Over the years various empirical and semi-empirical methods have been developed towards the goal of bridging the gap between specific rotation and structure. These have been reviewed recently by Polavarapu.[1, 2]

Computational chemistry can provide this information.[3-7] Early computational benchmarking on modeling optical rotation has been performed on molecules in the absence of solvent. From a computing perspective gas phase measurements are the easiest to model. However, experimentally most measurements are carried out in solution. This complicates the work of the computational chemist, as solvent effects may have a significant impact on the observed optical rotatory dispersion of a solution. Recent advances in this field have been made toward modeling chiroptical properties under the influence of solvent.[8, 9] Improvements in the accuracy of these modeling techniques have prompted an increasing number of chemists to rely on computational methods to assign absolute configurations.[10-14]

Most of the molecules studied thus far can be found in approximately the same geometries in solution as in the gas phase. With such molecules one may save computational costs by optimizing the molecule in gas phase, then treating solvent effects on the energy and response properties afterwards. But not all molecules are so well behaved. Some molecules adopt significantly differing geometries in vapor and in solution. In the most difficult cases the solution phase geometry is not even stable in the gas phase. The amino acids fall into this category of molecules, forming zwitterions in aqueous solution but reverting to their neutral form upon precipitation or vaporization. Part of the purpose of this work is to demonstrate that hybrid-DFT calculations employing large diffuse basis sets and an appropriate solvent model can provide reasonable optical rotation values for neutral, acidic and basic amino acids solutions.

Historical development of chiroptical measurement and modeling

As the technique of polarimetry developed, a standard method of measurement took hold. A solution of an analyte was prepared, usually using water as the solvent. The solution was placed in a transparent glass tube of a known length, and that tube was placed in the polarimeter instrument. A plane-polarized light source was then passed sample. The D-line of yellow light generated by a sodium vapor lamp, a doublet centered at 589.3nm, became a consensus choice for the light source. This was likely due to its widespread availability and the sensitivity of the detector to this color of light, the earliest detector being the human eye. "Room temperature" was a natural choice for the temperature of such measurements, which were typically made between 20 and 25 degrees Celsius. The length of the polarimeter tube could vary, but 10cm was a convenient size to use. Concentration of the solution could also vary, although analytical chemists were motivated to keep the concentrations low enough so that the relationship between

concentration and optical rotation remained linear. As such, the standard units of specific rotation became deg·cm^3/(g·dm), which referred to the degrees of rotation of the plane of polarized light, the concentration of the solution, in grams per cubic centimeter, and the length of the polatimeter tube, in decimeters.

As specific rotation data from numerous compounds measured under these standard conditions became available, early researchers strived to relate these chiroptical observations to those absolute configurations of the molecules that were known. In the era before modern computers, scientists relied heavily on empirical rules to try to make this relationship. Such rules derived from observed patterns observed in the measured specific rotation of compounds with absolute configurations that had been established by other means. For a review of these methods see the work of Polavarapu.[1]

Many of these classical methods were limited by the fact that each was designed for a specific class of molecules that were related in structure. In particular, some methods focused on sets of molecules that had a particular functional group. The historical rules that relate most closely with my work are the "sector rules", which try to relate specific rotation to the spatial relationship of "chromophores" to the other atoms in a molecule.[15] Used in this context, a chromophore is a functional group with π electrons that are subject to excitation by lower energy photons than are the core electrons or those in a molecules σ bonding network; they need not actually make the compound appear colored.

The chromophore thought to be primarily responsible for the specific rotation of the amino acids, which my project centers on, is the carbonyl chromophore. Such sector rules that have evolved for this chromophore include a "quadrant rule" and various forms of an "octant rule". These rules vary in the complexity with which they section off space about a

chromophore, but all have one thing in common: A perturbing atom that orients itself in a positive sector has a positive effect on the molecule's measured specific rotation, and an atom oriented in a negative sector has a negative effect on that rotation.

Amino acids found in their zwitterionic form or anionic forms possess chromophores that are slightly more complicated. These molecules possess a carboxylate chromophore, which contain a carbon with not one but two bonds to oxygen atoms that, due to resonance both have significant π character. But Jorgensen reasoned that a sector rule for this chromophore could be rationalized if it is though of as two overlapping carbonyl chromophores.[16] The way that two sets of octants overlap to form a sector model for the carboxylate chromophore is illustrated below:

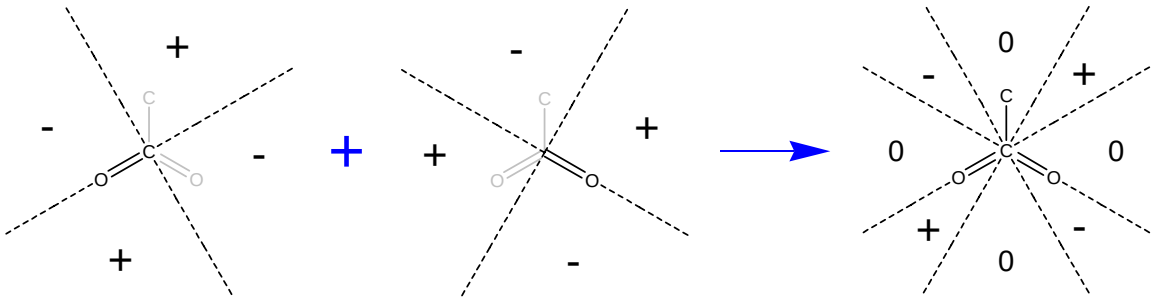


Figure 1-1: Formation of Jorgensen's sector rule for the carboxylate chromophore. Plus and minus signs shown refer to the sign of the sector above the plane of the paper, signs of corresponding sectors below the page are the reverse. Note that this diagram illustrates a superposition of two carbonyl groups; it is *not* meant to imply a pentavalent carbon atom.

Before the age of modern computer, such sector rules represented the state of the art in relating specific rotation to structure. But even before these empirical rules were developed, the fundamental mathematics behind chiroptical techniques had been laid out. Consider the difference in refractive index of a compound with respect to left handed and right handed circularly polarized light. This is related to the electric-magnetic polarizability tensor, β in the following equation:

$$d(t) = d^0 + \alpha\cdot\varepsilon(t) - \beta/c \cdot dB/dt + \text{higher order terms} \qquad (1\text{-}1)$$

Here d(t) is the net dipole of the molecule, d^0 is its permanent dipole moment, $\alpha\cdot\varepsilon(t)$ is the electric polarizability, B is the magnetic polarizability and c is the speed of light. This method tends to be referred to by the moniker, "linear response method", which is unfortunate since the term "linear response" is rather generic and can refer to other things in a different context. Those interested in the derivations may for example consult the reviews of Polavarapu[1, 18],Crawford[6], Pecul and Ruud,[5] and Autschbach[19].for overviews of the mathematics.. The essence of the linear response method is that optical rotation may be calculated at a desired wavelength without explicit computation of all the excitations using time dependant density functional theory. This method is not subject to the truncation error of the sum over states method which will be mentioned later. A draw back of this method is that, in some of the implementations of it available in current software, none damping effects of electronic excitations are taken into account, therefore the optical rotation that one models in the vicinity of an excitation will be abnormally large in magnitude, and a calculation of optical rotation at an excitation energy will fail completely, since the undamped linear response method yields a singularity at all excitation wavelengths.

For a reference for the equations in this section, see for example the recent review of Crawford and co-workers.[17] From this fundamental equation can be derived the Sum Over States (SOS) equation, which can be used to compute specific rotation based on the sum of these calculated rotational strengths:

$$G'(\omega) = \frac{-2\omega}{\hbar} \cdot \sum_{n\neq 0} \frac{R_{0n}}{\omega_{0n}^2 - \omega^2} \qquad (1\text{-}2)$$

In this equation, ω is the angular frequency ($2\pi \cdot \nu$) of light for which specific rotation is calculated in atomic units (usually at 589.3nm), ω_j is the angular frequency of an excitation and R_{0n} the corresponding rotatory strength from equation (1-4). The sum formally runs over the complete set of excitations of the molecule. This proves to be a weakness of the sum over states equation, since in order to rigorously compute specific rotation using it; one must take into account all possible excitations as are calculated. For all but the smallest molecules computed with the smallest basis sets, the number of excitations is impractically large. Therefore the sum must be truncated at some point to make the equation (1-2) soluble. This truncation becomes a source of error in the calculation.

From G' specific rotation, [α] may be calculated:

$$[\alpha]_{\omega} = \frac{(72.0\times10^{6})\hbar N_A \omega}{c^2 m_e{}^2 M} \cdot \frac{1}{3} Tr(G') \qquad (1\text{-}3)$$

Here h-bar is Planck's constant divided by 2π, N_A is Avogadro's number, c is the speed of light in m/s, m_e is the mass of an electron in kg, M is the molecular weight of the compound in atomic mass units and Tr(G') represents the sum of the diagonal elements of G'.
This equation, developed by Rosenfeld in the 1920's, and those derived from it would later form the basis for the modern quantum chemical algorithms that we now use for modeling chiroptical properties:

$$R_{0n} = \mathrm{Im}(\langle\Psi_0|\mu|\Psi_n\rangle \cdot \langle\Psi_n|m|\Psi_0\rangle) \qquad (1\text{-}4)$$

In this equation, named for its developer, R_{0n} represents the rotatory strength of a transition from state 0 to state n, Ψ_0 and Ψ_n represent the wave functions of those respective states, and μ and m represent the sums of the electronic and magnetic dipole operators,

respectively. The "Im" in the notation means that the rotational strength is equal to the *imaginary* part of the dot product of the factors in parentheses.

One final method by which optical rotatory dispersion (including specific rotation) may be computed is from known circular dichroism by using the Kramers-Kronig (KK) transformation. The CD spectrum used for the transformation may be either measured or calculated. The equation is depicted below. Note the similarity to the sum over states form in equation (1-2):

$$[\phi]_\lambda = \frac{2}{\pi}\int_{\lambda'=0}^{\infty}[\theta(\lambda')]\frac{\lambda'}{\lambda^2 - \lambda'^2}d\lambda' \qquad (1\text{-}5)$$

Here $[\phi]_\lambda$ is the molar rotation at wavelength λ, and $\theta(\lambda')$ is the circular dichroism at wavelength λ'. Specific rotation, $[\alpha]_\lambda$, can of course be computed from this molar rotation by multiplying the molar rotation by 100 and dividing by the molecular weight of the compound. A major advantage of the KK transformation is the speed at which it can be carried out. No wave functions need be calculated, and so the time required for such a transformation depends primarily on the level of precision used for the integration grid. The disadvantage of this method is that the quality of the molar rotation that is output is dependant on the quality of the CD spectrum used for input. If the CD spectrum were generated by the sum over states method, then the results of a KK transformation of that spectrum would have the same truncation error as a molar rotation generated directly by the SOS method.

The relationship between the mathematical methods and the chiroptical response properties is summarized by Figure 1-2. Of these three methods that can be used to compute ORD (and specific rotation), the linear response method is preferred for those wavelengths of light that are far from an electronic excitation. At wavelengths closer to an excitation, the SOS and KK methods become more practical.

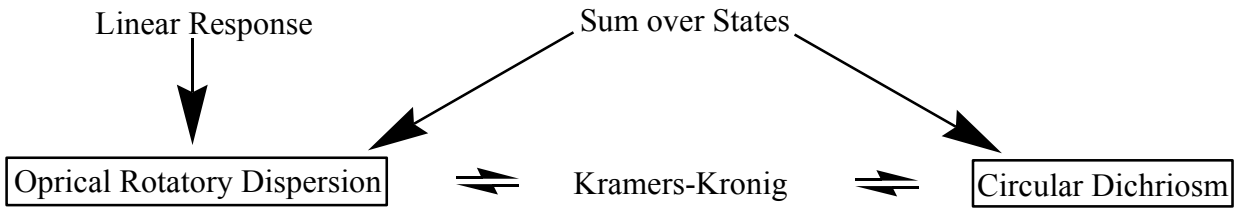


Figure 1-2: Relationship of the chiroptical responses and the mathematics that can be used to compute them. Specific rotation, also of interest, is a subset of ORD.

Outline of thesis

This thesis has begun with a brief review of the old empirical models which were used to try to model specific rotation. The contemporary computational methods that may be employed to model chiroptical responses and the mathematics on which they are based have been introduced. The remaining chapters, constituting the bulk of this work, deal not with the derivation of the fundamental mathematics, but the application of established computational models to solve the problem of modeling the chiroptical responses of amino acids solutions.

The second chapter of this dissertation constitutes a benchmarking chapter. The research contained therein did not make it into print, owing to the interest in efficiency and the fact that much of it reinforces principles that have already been established by others in the research field, albeit in different contexts. However, such information should prove quite useful to anyone wishing to continue the work that I have started. Subjects in this chapter include choice of basis sets / density functionals, solvent models, and issues having to do with the timing and geometries generated in molecular dynamics simulations.

The third chapter is the fourth paper that I published, and the first that I submitted as part of my research at UB. It involved modeling the specific rotation and the transparent region optical rotatory dispersion (ORD) of some of the smallest amino acids, alanine, proline and serine. Among the amino acids, these were natural targets for computational study due to their

relatively low molecular weights and small number of possible conformations. Modeling of the specific rotation of alanine zwitterion and cation, as well as that of the proline cation, zwitterion and anion were successful. Resulting specific rotations tended to be of the correct sign, and slightly larger in magnitude than those of experiment.

My second publication here was a subset of a larger paper primarily authored by Mykhaylo Krykunov, a postdoctoral research fellow formerly at this university. While not a separate publication, my work contributed significantly to the paper, helping to validate the computer code that he had developed with a "real world" example. My contribution, which I have adapted to a brief chapter, takes a more detailed look at the ORD of proline, using the Kramers-Kronig (KK) transformation to extend the modeling of this dispersion curve to include the wavelengths of light which the molecule absorbs, the so called "anomolous region". This dispersion curve was found to be most correct when the sum of the rotary strengths in the partial sum-over-states of the CD calculation was closest to zero, in keeping with the sum rule.

Chapter five, my third publication at UB, modeled the chiroptical properties of the aromatic amino acids: histidine, phenylalanine, tyrosine and tryptophan. These properties depend on two distinct chromophores, the carboxylic acid/carboxylate group, present in all protonated/deprotonated amino acids, in addition to an aromatic side chain which differs amongst the amino acids. The magnitude of the influence of this aromatic group on specific rotation varied significantly depending on the nature of aromatic groups. This influence can be explained through a combination of the effects of the selection rules for electronic transitions, as well as the energies at which those transitions take place.

The foregoing publications set the stage for my magnum opus, which was recently published in the Journal of the American Chemical Society. This publication, which constitutes

chapter six of my thesis, investigated Clough-Lutz-Jirgensons rule. This rule is an old semi-empirical rationale for the consistent perturbation of the molar rotation of solutions of L-amino acids by lowering the pH of those solutions. Here I used TDDFT to marry the sector rule based rationale with our modern understanding of molar rotation based on ab-initio electronic calculations. From a practical aspect I found that TDDFT was more consistent at modeling the change in molar rotation accompanied by a zwitterionic to cationic transition of an amino acid than it is modeling the absolute value of the molar rotation of either of the corresponding cationic or zwitterionic molecules.

All my work published thus far at this university has been modeling of properties using static molecular structures optimized by density functional theory. Chapter seven, which was recently submitted for publication, will be my first paper that employs empirically derived molecular dynamics simulations generating a variety of structures for investigation. This opens the door to investigating vibrational effects on chiroptical properties, as well as allowing the use of explicit solvation, as opposed to the continuum solvation modeling I had used in my forgoing work. Here I found that explicit solvation using simple point charge type water molecules could effectively simulate some of the solvation effects on chiroptical responses.

Chapter eight begins with a summary of my published and presented work, and its impact on the scientific community. Every researcher should strive to have his work accepted and validated by others, and the reception that I have gotten serves as evidence that this is happening. This final chapter ends with a brief outlook for the future of the research in my field, and the challenges that must be overcome for it to be successful.

2. BENCHMARKING AND CALIBRATION OF METHODS USED

Introduction

In the course of one's term as a PhD student, there are always results that go unpublished. In my particular case most of my work has been published, and chapters 3-7 are dedicated to those papers. However there is work that for whatever reason has not made it into those papers. This includes the results of many benchmarking calculations, timing computations, method validation and calibration…all of the work that must be done in order to successfully complete a research project, but which does not fit into the limited space given to journal articles. This chapter is set aside for such work.

Chapter 2 begins with a discussion of basis sets used. While this choice is typically just a compromise between the speed of computations with small basis sets and the robustness of those with large sets, here the situation was more complex. I found that diffuse sets were the best for TDDFT response calculations, while a basis not augmented with diffuse functions proved better for geometry optimizations and energy computations. The role of intermolecular hydrogen bonding, as well as the effects solvation has on the molecules contributed to this conclusion.

Next the choice of functional is discussed. While the B3LYP hybrid functional was found sufficient for geometry optimizations and energy calculations, it was not always the best choice for response calculations; for these other hybrids gave better results. There is not one "best" hybrid for all molecules, and the choice of optimal functional appeared to depend on the charge of the molecule being modeled.

Then this chapter takes a look at the issue of timing of sampling in molecular dynamics simulations. It was found that at 300K it takes a glycine – water system between 1 and 2 picoseconds to randomize. Surprisingly, this was about the same whether the solute was fixed or

flexible, indicating that the glycine molecule rearranges itself on a similar time scale as the water molecules around do.

Next differing explicit water solvation models and their effects on energy and optical rotation were considered. I looked at several static point charge based models, and one where point charges were assigned dynamically based on the direct reaction field model. Despite differences in the models, all had comparable effects on solvation energy as did full quantum mechanical water molecules.

Finally, the effects on molecular geometry on energy and molar rotation were investigated. In particular, the consequences of rotation about the carbon-carbon bond of glycine were considered in some detail. The dihedral angle about this bond correlated very strongly with the molar rotation of the molecule, in keeping with the principles of the sector rules.

Choice of basis sets

I entered this university with a "bigger is better" philosophy for basis set choice. In principle, there is some probability of finding an atom's electron at *any* distance from the nucleus. A basis set approximates this infinite space with a finite number of functions. As such, I endeavored to use a basis that was so large that either: 1. the property which I was measuring had converged, i.e. increasing the basis set size would not significantly alter the results or 2. the basis set was so big that the calculations became impractical due to memory or computing time requirements. What I found during the course of my work is that the choice of basis set is not as simple as I first thought.

For my first publication[20] at the University at Buffalo, I followed the recommendations of Grimme et al[21], using Dunning's d-aug-cc-pVDZ basis for computations on the small amino acids. For geometry optimization, I had to abandon this doubly diffuse basis for the singly

diffuse aug-cc-pVDZ, for reasons which I will discuss later, having to do with the solvent model and hydrogen bonding. For the response calculation I retained the doubly diffuse set, at least for this first publication, as the second diffuse function, according to Woon and Dunning, was important for the computation of dipole polarizability of neon, and as such it was reasonable to expect it to have an effect on such response properties of larger molecules as well. I confirmed this is the case, as can be seen in figure 2-1.

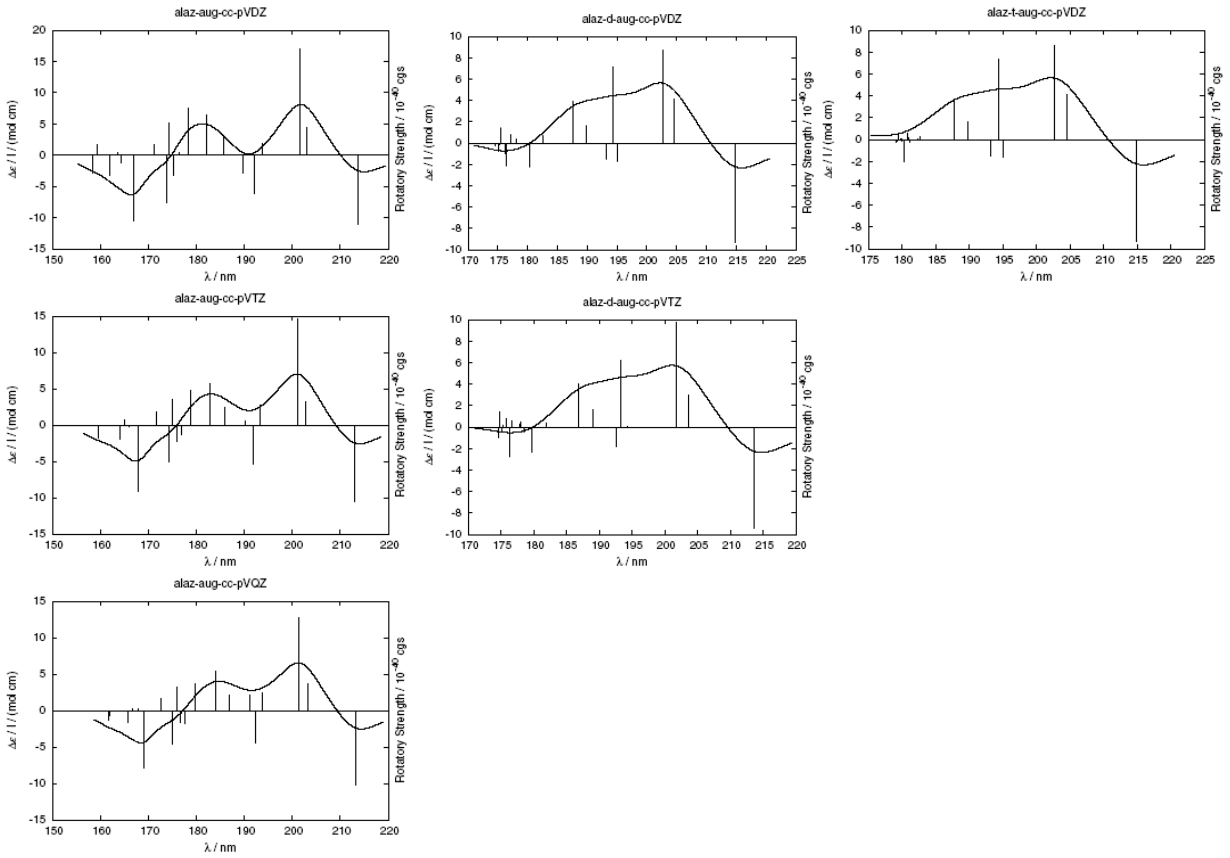


Figure 2-1: The computed CD spectrum of alanine with various basis sets. The number of diffuse functions (one, two or three) increases from left to right. The cardinal number of the basis (double, triple, or quadruple zeta) increases from top to bottom. The B3LYP functional was used for all these computations. Geometries were optimized at the B3LYP/TZVPP level of theory, and COMSO solvation was used throughout.

Adding a second set of diffuse functions to the aug-cc-pVDZ basis significantly affects the lowest electronic transitions of the alanine zwitterion (though some of these were later found to be spurious charge transfer transitions). Addition of a third set of diffuse functions has little effect, but to increase the time of the computation. The change from a double to triple zeta basis also has little effect on the lowest energy transitions, which have the greatest impact on the optical rotation at 589.3nm. Going from a triple to quadruple zeta basis has no discernable effect on the CD, however it tremendously increases computational time, from less than one day to over eleven days (Table 2-1). When all of these factors are considered, the d-aug-cc-pVDZ basis

appears to be the optimal choice for modeling the sodium D-line optical rotation and low energy CD spectra of these compounds.

Table 2-1: Time (hours:minutes:seconds) required to complete a single point (dscf) and a CD spectrum through the first 20 excitations (escf) for the solvated alanine zwitterion. A single 2.8Ghz Intel Pentium 4 CPU was used for each calculation.

Basis	dscf time	escf time	total time	primitive functions	contracted functions
aug-cc-pVDZ	0:08:47	1:03:30	1:12:17	335	201
aug-cc-pVTZ	1:28:40	17:36:25	19:05:05	571	437
aug-cc-pVQZ	18:50:49	249:35:32	268:26:21	948	802
d-aug-cc-pVDZ	0:26:52	3:31:41	3:58:33	417	283
d-aug-cc-pVTZ	6:24:07	81:00:50	87:24:57	730	596
t-aug-cc-pVDZ	1:21:37	11:12:30	12:34:07	499	365

That the doubly augmented basis sets appear optimal for response calculations does not necessarily mean that they are best for geometry optimizations or single point calculations. The d-aug-cc-pVDZ basis set was useful for optimizing the smaller amino acids glycine and alanine, but failed for the larger one, serine. The reason is directly related to the implementation of the COMSO solvation model. While I could not isolate the exact reason why these calculations failed to converge, one possible source of error with continuum models is charge lying outside of the solvation cavity.[22] As this problem is known to be the most severe with electronically diffuse systems, I briefly investigated how much this escaped charge problem might be affecting the modeling of the amino acids.

The implementation of COMSO in the Turbomole code contains an algorithm for compensating for escaped charge, the details for which appear in the work of Klamt and Jonas.[22] The program computes how much of the molecular charge is in the primary COSMO radius, defined here as 1.17 times the Bohr radii of the atoms. It then calculates the charge that is contained within a second shell around the molecule, 1 angstrom outside the primary shell, and empirically corrects the energy of the system for this escaped charge. Charge that is found outside this second shell has escaped the solvent model completely, and is a more serious source

of error. The quantity of charge that escapes the primary solvation shell and that which escapes the secondary shell as well is plotted for various basis sets for differing ionization states of glycine is figure 2-2.

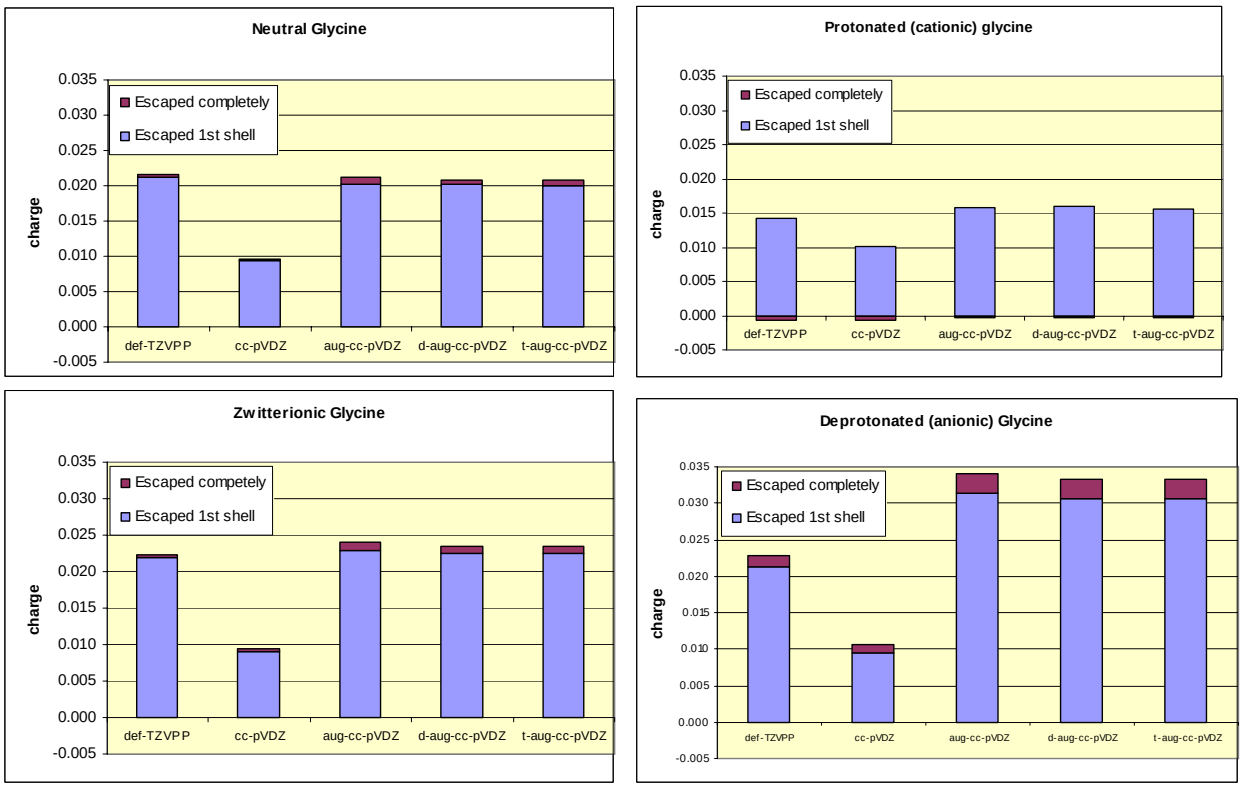


Figure 2-2: The "escaped charge" (in atomic units) of the glycine molecule as a function of basis set used. The primary solvent shell is 1.17 times the Bohr radius of the atoms from the nuclei; the secondary shell is an additional angstrom away. The B3LYP functional was used for all calculations.

Two patterns are apparent in the charts, both of which are expected. First of all, the molecules with more diffuse electronics have more of an issue with charge escaping the COSMO shells around the solute. The most charge escaped from the deprotonated glycine, which carries an overall negative molecular charge and as such has its outermost electrons most weakly held and at the greatest distance from the nucleus. Zwitterionic and neutral glycine have the next highest quantity of escaped charge. Though both of these molecules have an overall zero charge, the zwitterion should and does have a somewhat larger amount of charge lying outside the

respective COSMO cavities due to its negatively charged carboxylate group. The protonated glycine with its positive charge has the least amount of charge escaping the first solvation shell, with the quantity that escapes the second shell being numerically insignificant.

With respect to the basis set, the larger more diffuse sets allowed for more escaped charge than did the smaller sets and those without diffuse functions. The reason is trivial: if a basis set does not contain functions diffuse enough to penetrate beyond the COSMO solvent radius, then electronic density cannot be found there. The increased quantity of escaped charge in computations with the augmented basis sets may have contributed to the fact the geometry optimizations with some of these molecules did not sufficiently converge when a diffuse basis was used. This is part of the reason I abandoned the use of the aug-cc-pVDZ for geometries after my first publication with it, using instead the default TZVPP basis for this purpose in subsequent publications.

Another reason for switching to the def-TZVPP basis for geometry optimizations and energy calculations is it appeared to give slightly better energies for my molecules than Dunnig's basis sets did. As shown in figure 2-3, the energy gap between that of a neutral glycine molecule and a solvated zwitterion is marginally closer to that of experiment with the def-TZVPP basis than with any of the other basis sets used. This may be partially due to the fact that less charge escapes the COSMO cavity with this basis than with the augmented basis sets; it may also be due in part to the fact that the TZVPP basis was one of the basis sets for which the solvent model was originally calibrated.

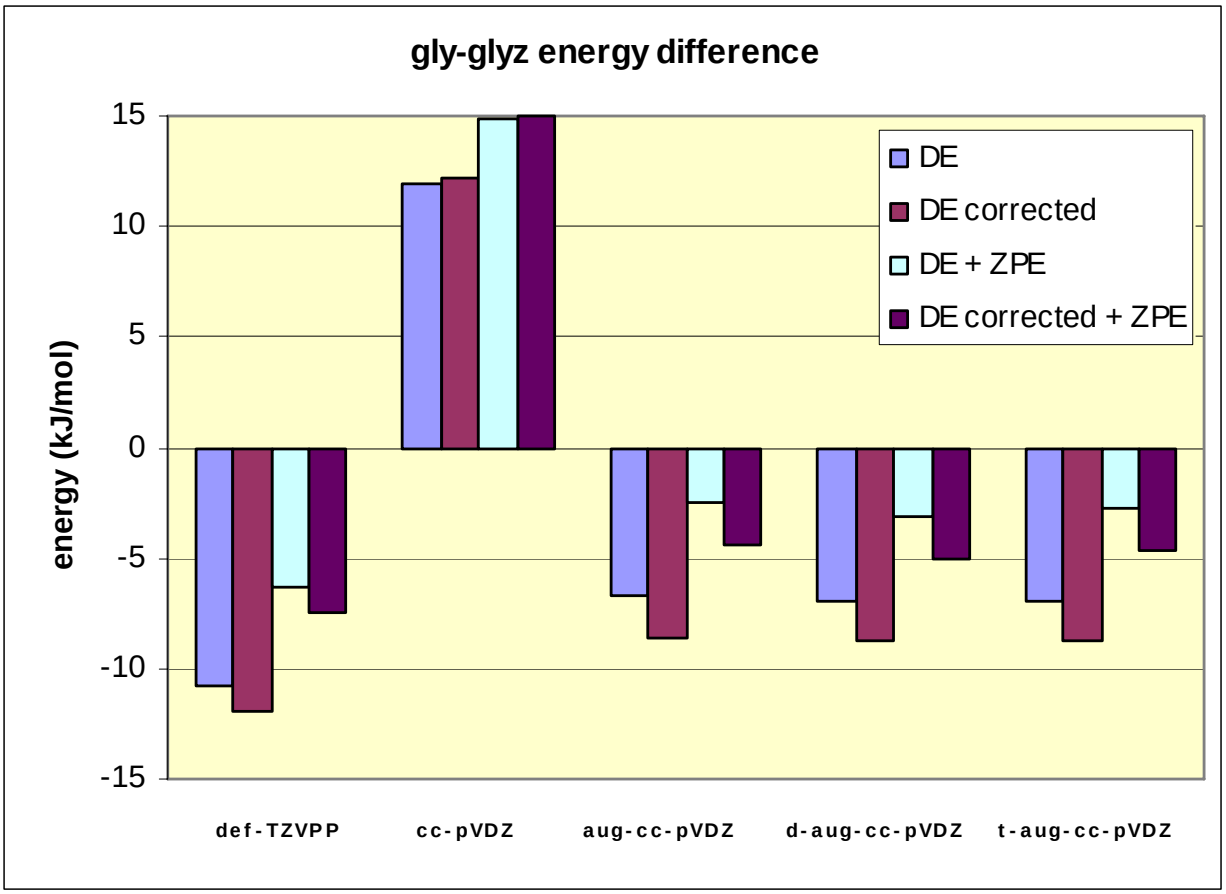


Figure 2-3: The energy difference between neutral solvated glycine and the solvated glycine zwitterion as calculated with the B3LYP functional and selected basis sets. The energy is calculated with and without the COSMO correction for outlying charge as well as with and with out zero-point vibrational corrections. For reference, the experimentally derived value for the neutral-zwitterionic energy difference is 30.4kJ/mol.[23]

Ultimately, using the def-TZVPP basis for geometry optimizations and energy calculations resulted in relative energies that were marginally closer to experimental values. Part of this may be the better compatibility of the TZVPP basis with the solvent model. Some improvement may have been the omission of diffuse functions, since diffuse functions promote long range interactions, like hydrogen bonding, and amino acid conformers with intramolecular hydrogen bonds tended to be too low in energy, especially with diffuse basis sets. However even with the COSMO/TZVPP method, accurate calculation of molecular energies remained a significant source of error when computing accurate Boltzmann populations is critical.

Choice of functional

In addition to the basis set, the choice of the density functional used is critical to the development of a theoretical model. For geometry calculations, the difference between pure and hybrid DFT functionals did not appear to be significant, so the choice of functional was not investigated to any great depth here. I used B3LYP for all geometry and energetic calculations merely for consistency; it was chosen due its well established status as being reliable for such purposes.

For response calculations, the choice of functional proved critical. My earliest computations of amino acid specific rotation with the Amsterdam Density Functional (ADF) program, using "pure" DFT, lead to rotations that were far greater in magnitude than those of experiment. This ultimately prompted the abandonment of ADF in favor of Turbomole, for which hybrid DFT response was implemented.

My choice of hybrid evolved as my research progressed. For my first paper published here, the B3LYP functional was used. In subsequent works, BHLYP was used as well. Most recently, I have gravitated toward PBE0, due to its producing a first excitation energy for glycine that is closest to that modeled at the CC2 level of theory. The critical component appears to be the amount of exact exchange in the hybrid: for the zwitterionic amino acids, about 25% appears optimal. But there is no "correct" hybrid for all situations, as for the cationic amino acids, less exact exchange could be used and good results still obtained. For the anionic forms, more exact exchange appears to be appropriate.

DRF-90 timing

In order to model a dynamic system via a series of static snapshots the first thing that needs to be determined is what frame rate to conduct the sampling at. For the motion picture

industry, standard practice is to record a frame once every $1/24^{th}$ of a second. To record at a much slower frame rate would result in missing valuable information about the motion of the actors. To record at a much faster frame rate would be a waste of film when the change between frames is negligible.

The TDDFT used to model the properties of a molecular system can perceive changes on a much smaller time scale than human beings watching a movie in a theater, so a much faster frame rate must be in molecular modeling. The quantum mechanical calculations that are performed on each frame consume time and computing resources, so I do not want our frame rate to be too fast, lest I be repeatedly running expensive computations on a system that is not changing significantly between each frame. Nor do I want our frame rate to be too slow, since running a molecular mechanics dynamics simulation requires resources too, and I do not want information from a dynamics run to go to waste if it is of value.

To this end, I sought a sampling rate that would yield geometries that were independent of one another. In other words, the energy and chiroptical response properties that we are interested in studying should randomize completely between frames, such that any similarity in these properties between adjacent frames would be a coincidence. Doing this involves looking at the quantities of interest, energy and molar rotation, over reasonably large data sets (2048 frames) which were collected at differing time intervals ranging from 1 to 10000 femtoseconds. Next each adjacent pair of data points were averaged, to give new data sets consisting of half as many (1024) points. Then the standard deviation of the values for each original data set is divided by the standard deviation of the corresponding second set. If the two frames are completely identical, then this resulting ratio should be equal to one, since the standard deviation of a set does not change if identical numbers are averaged. If however the adjacent frames

contain independent data, then the ratio of the standard deviations should, by the principle of signal averaging, approach the square root of the number of frames included in this average, in this case two. The results of this analysis are depicted in figure 2-4.

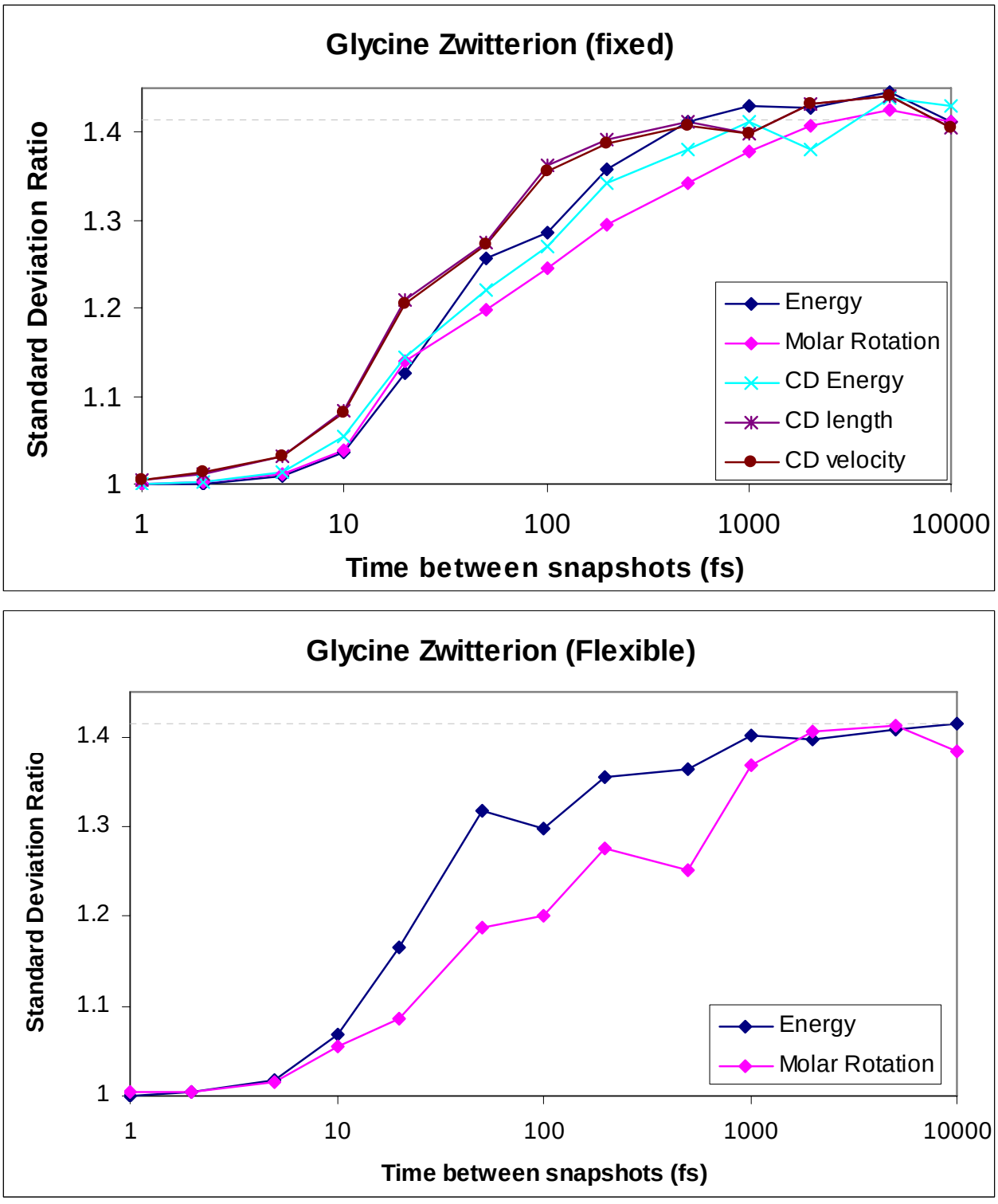


Figure 2-4: Relative iterframe independence for select properties of glycine as a function of frame rate. The data on the top are from a system where the solute geometry was fixed, for the data on the bottom the solute was flexible. The dashed line represents the square root of two, the point at which the frames can be deemed independent. Calculations were performed with the DRF-90 program at 300K with a time step of 1fs.

At the far left of each graph where frames are samples at a rate of one every femtosecond, there is very little difference between the frames. As the time between samples increases through 10, 100 and 1000 femtoseconds the adjacent frames become more independent, as is evident in the standard deviation ratio increasing from a low of one and slowly converging to the square root of two, at which independence is achieved. Note that it does not exactly approach the square root of two, nor does it approach that value in a perfectly smooth fashion; this is a consequence of the finite sampling set. But even with this limitation it is apparent that interframe independence is achieved in around 1000 to 2000 femtoseconds.

Ultimately, this information did not make it into publication, as I switched molecular dynamics programs from DRF90 to GROMACS. This switch, prompted by the relative maturities of the programs, resulted in a speed up of nearly two orders of magnitude. Thus, choosing a time between frames that was as short as possible became less of an issue, as molecular dynamics computation time became nearly 100 fold cheaper I could simply choose an arbitrarily large number of steps between frames. In fact, in the work I recently submitted for publication 10000fs were allowed to elapse between the geometries that were collected for subsequent ab-initio calculations. However, the analysis of the timing of the molecular dynamics remains valid, and may be useful to those wishing to pursue such studies.

Choice of solvent model

In order to successfully model an amino acid with explicit solvation, some type of non-quantum mechanical solvent model is essential. This is true because the large number of water molecules required to solvate an amino acid, combined with the large number of geometries for which quantum mechanical response calculations would simply overwhelm our computing resources. The original choice for this was the Direct Reaction Field (DRF) model, which can

take into account the multipole moments of water (if given enough degrees of freedom) and the polarization induced on that water by neighboring solute or solvent molecules, and to feed that effect back so that polarization is felt by the solute. Unfortunately, as of this writing, this method is only fully implemented in the Amsterdam Density Functional code. Response functions using hybrid DFT have yet to be perfected with this code, and because of the serious problems with spurious charge transfer when pure DFT functionals are used to model the zwitterionic molecules under investigation, hybrid DFT (or a wave function based method that does not have the charge – transfer problem) is essential.

As such, a method to effectively simulate the polarization induced by explicit solvent molecules had to be developed that was compatible with the Turbomole program, which runs quite efficiently and has hybrid DFT response fully implemented. Turbomole typically does not give out source code to its end users, and it has not thus far given this code to this group. Therefore any method of explicit solvation must be compatible with being scripted into standard Turbomole input. For this purpose point charges were selected. The magnitudes and positions of these point charges relative to their respective nuclei were adopted from the TIP waters of Jorgensen and coworkers.[24] The three such waters I considered are illustrated in figure 2-5.

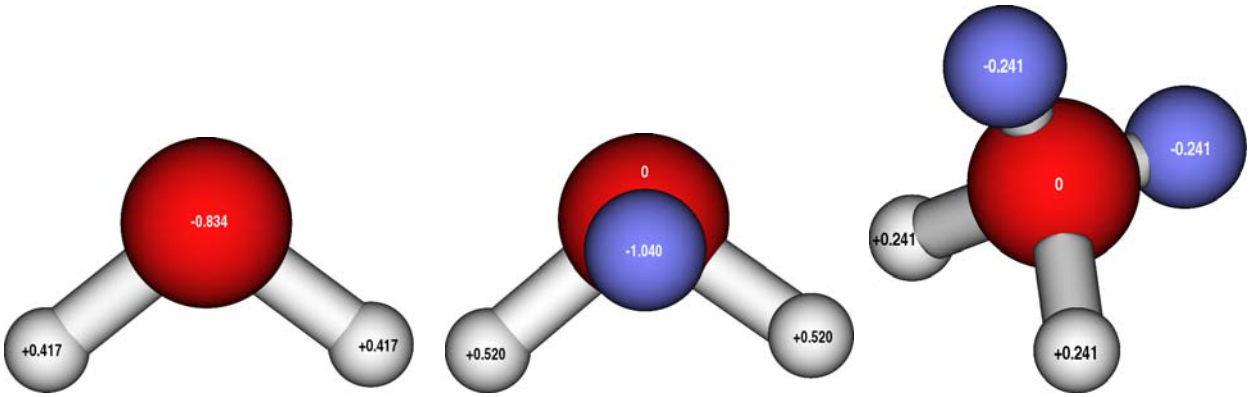


Figure 2-5: The point charge based water models considered in this work. Numbers represent the point charges used, in atomic units

The numerals in the TIP waters refer to the number of points at which a partial charge and / or a nucleus is located, be this three, four or five points. All of these molecules have the same nuclear geometries, with fixed bond lengths of 0.9572 angstroms and bond angles of 104.52 degrees. They differ only in the placement and magnitude of the partial charges. For the TIP3P model, the point charges are centered on the nuclei, and partial charges of -0.834 and +0.417 were used for the oxygen and hydrogen atoms, respectively. For the TIP4P water, partial charges increase to -1.040 and +0.520, respectively and the charge for the oxygen is not centered on the nucleus, but instead moved 0.15 angstrom towards the interior of the molecule on a line bisecting the two hydrogen atoms. For the TIP5P model charges of +0.241 are used on the hydrogen atoms, with the balancing counter charge being split into two parts located roughly in the directions that the lone pairs for the oxygen atoms would be, at 0.7 angstrom from the oxygen atom in a plane orthogonal to that of the water nuclei, with a charge – oxygen – charge angle of 109.47 degrees.

These three represent just a few of hundreds of water models developed by molecular modelers over the years. Because the chiroptical response properties that I am measuring depend on polarization of the solute that these solvent molecules should affect, it makes sense to consider the multipole moments of these differing solvent models. These data are available on Matrin Chaplin's website at http://www.lsbu.ac.uk/water/models.html, which is an invaluable resource for those interested in the molecular modeling of water. Some of the pertinent data from here are summarized in table 2-2.

Table 2-2: Comparison of the dipole and quadrupole moments of the selected point charge water models with those computed at the MP2 level of theory. Ab-Initio data are from Tu and Laaksonen[25]

	Dipole (Debye)	Quadrupole (Debye * Angstrom)		
Model	z	xx	yy	zz
TIP3P	2.35	-2.30	0.00	-1.38
TIP4P	2.32	-2.86	0.00	-1.60
TIP5P	2.29	-1.33	0.76	-0.42
Gas Phase	1.855			
Tu and Laaksonen	2.65	-4.27	-7.99	-5.94

Overall, the dipole and quadrupole moments for the established water models appear to be somewhat smaller than those computed ab initio. This is consistent with my conclusions in chapter 7, where I found the changes in chiroptical response induced by TIP3P waters is somewhat weaker than that obtained from quantum mechanical waters. It should be possible in the future to tune the dipole moments of these molecules in order to achieve better agreement with QM diploes. The quadrupoles could be tuned as well (as could the octapoles, the hexadecapoles, et cetera), though for those more charge locations would be needed on each water molecule, as a simple C_{2V} symmetrical 3 point charge system does not have the degrees of freedom needed to solve for all of these multipoles simultaneously.

To examine the effects of these differing water models I ran a 64ps molecular dynamics simulation of a glycine zwitterion in a sphere of 128 water molecules with the DRF-90 program[26], collecting geometry snapshots every ps. I then selected the water molecule nearest the glycine solute in each geometry and ran a series of quantum mechanical calculations on those glycine water systems with Turbomole, averaging the results from the 64 configurations. I then did the same with glycine solvated by two, three, four…up to eight water molecules. Some of the results of theses calculations, are depicted in figure 2-6.

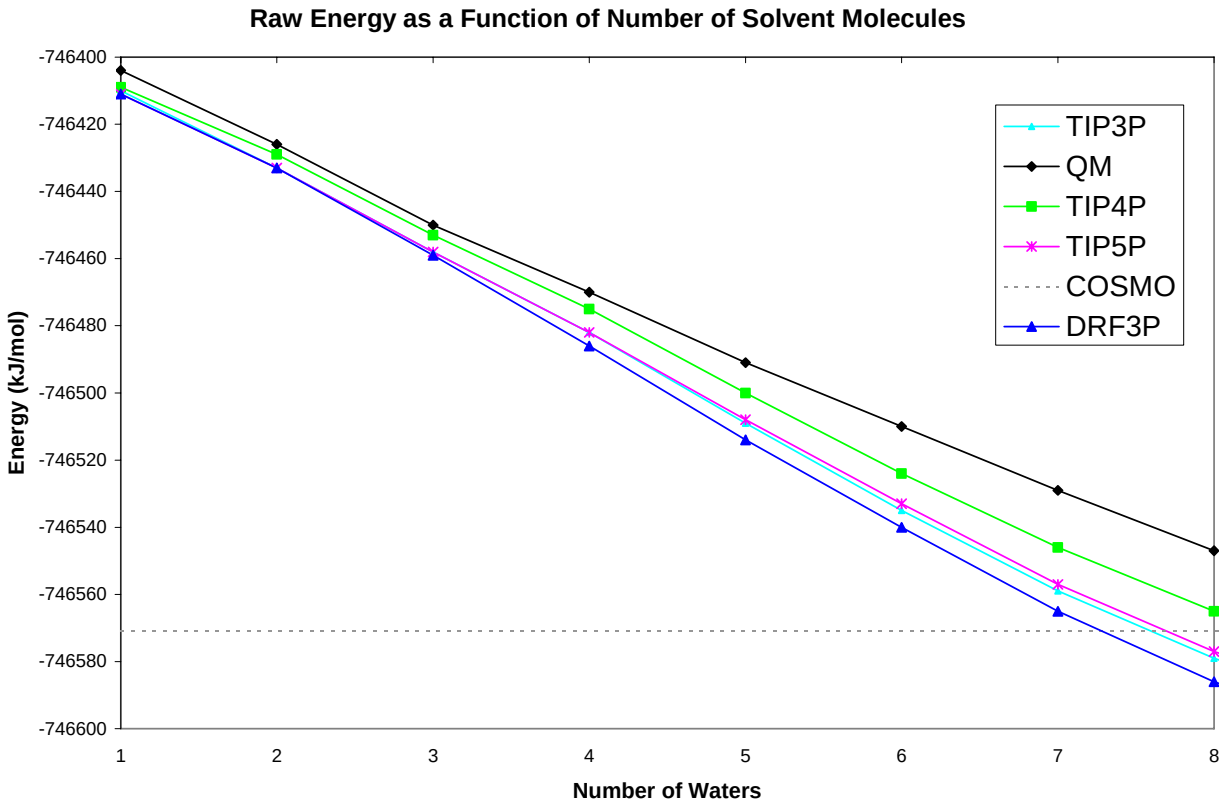


Figure 2-6: The energy of a glycine – water system as a function of number of waters at the B3LYP/aug-cc-pVDZ level of theory

Each point charge based solvent model produces a similar trend to that seen by the addition of a corresponding number of full fledged QM water molecules. (The partial solvation energies shown for the QM systems were calculated by subtracting the energy of the water molecules from that of the solute.) All of the point charge waters used to solvate glycine lowered the energy of the system more than comparable QM waters did. Of course point charge models do not simulate all of the effects that QM waters do, steric repulsion being one of these, which may account for some of the energy difference between the two. Amongst the TIP models, the 3 point and 5 point had surprisingly similar effects on the energy of the system, despite the difference in the nature of the charge distributions between the two models. The

TIP4P model, which has charges that are greater in magnitude than the TIP3P but placed closer together, shows less of an effect on energy than the other TIP models do.

The only charge model where the charges are not fixed quantities, the DRF3P model, shows the greatest solvation effect. In this model charges occupy the sites of the water nuclei, as with TIP3P. However the TIP charges are static, having constant charges of -0.834 and +0.417 on the oxygen and hydrogen atoms, respectively. In the DRF3P model, these charges are recomputed for each water for each configuration, based upon the polarizability of the water molecules using the DRF model. On average, these charges were -0.764 and 0.382 for oxygen and hydrogen. Thus, the DRF3P charge waters have a greater impact on energy than do the static TIP3P waters, despite having partial charges that on average are lower in magnitude.

The effects of the dihedral angles on energy and molar rotation

This final section of chapter 2 deals with research which has not yet been submitted for publication, but will be in the future. Because it is helpful in the discussion of molecular structure and chirality, I include it here. It involves the O-C-C-N dihedral angle of glycine zwitterion, a structural portion which is present in all amino acids. As this angle varies during the course of a molecular dynamics simulation, the energy of the system and its chiroptical response vary as well. This section shows how the energy and molar rotation of glycine correspond strongly to this dihedral angle.

The energy of a solvated glycine molecule as a function of this angle is plotted in figure 2-7. The first thing of note from this graph is that the energy of the system varies much more when explicit point charge water molecules are used for solvation than when the COSMO continuum is used. This is to be expected since the large variations in the geometries formed by

the explicit waters can lead to greater changes, both positive and negative, of the glycine – water system.

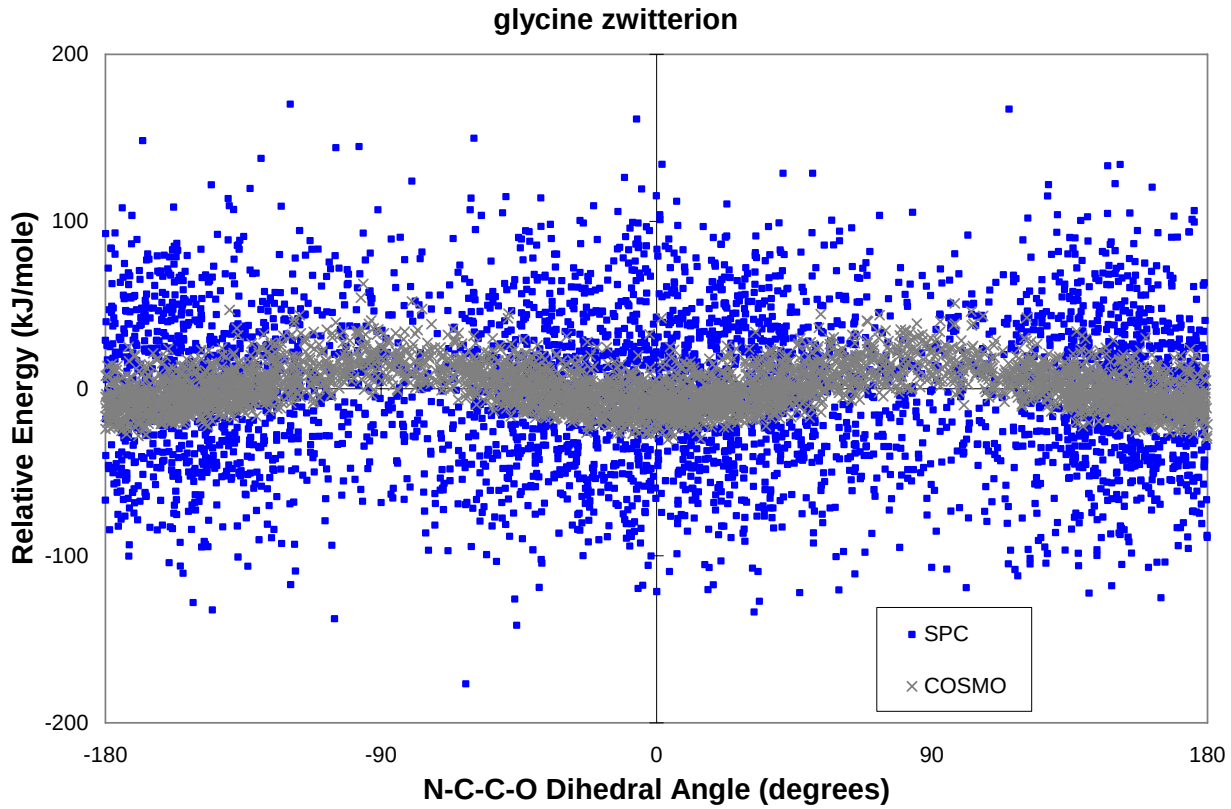


Figure 2-7: The energy of a glycine system as a function of the N-C-C-O dihedral angle. The energies reported are for the same dynamics run, and include either 256 point charge based water molecules or the COSMO solvation model. All energies are reported relative to the average energy for the model. For the geometries the GROMACS program was used, and snapshots were recorded at 10ps intervals.

The second point I want to draw attention to in figure 2-7 is the correlation between energy and population. The energy of the system is calculated by Turbomole at the PBE0/aug-cc-pVDZ level of theory. The population of the geometries with their particular dihedral angles is determined by the all-atoms force-field in GROMACS. The geometries of glycine that constitute local minima *could* differ between the QM and MM methods, but they do not. The trend of the QM energies indicated that the molecule has a local minimum structure with its dihedral angle at 0 degrees (or 180 degrees, due to the degeneracy in the C_{2v} symmetrical

carboxylate group). This is in agreement with the QM optimization calculations on glycine in chapter 3. In that work I found the glycine structure with the ±90 degree dihedral to be a saddle point, which is consistent with the date in figure 2-7. These data also indicate that the energetics of the MM simulation agree with the QM results. The geometries are governed by the physics of molecular mechanics, and these have resulted in far more structures with dihedral angles close to 0 and 180 degrees than to ±90 degrees, indicating that 0 and 90 degrees are also minima and saddle points at the molecular mechanics level of theory as well.

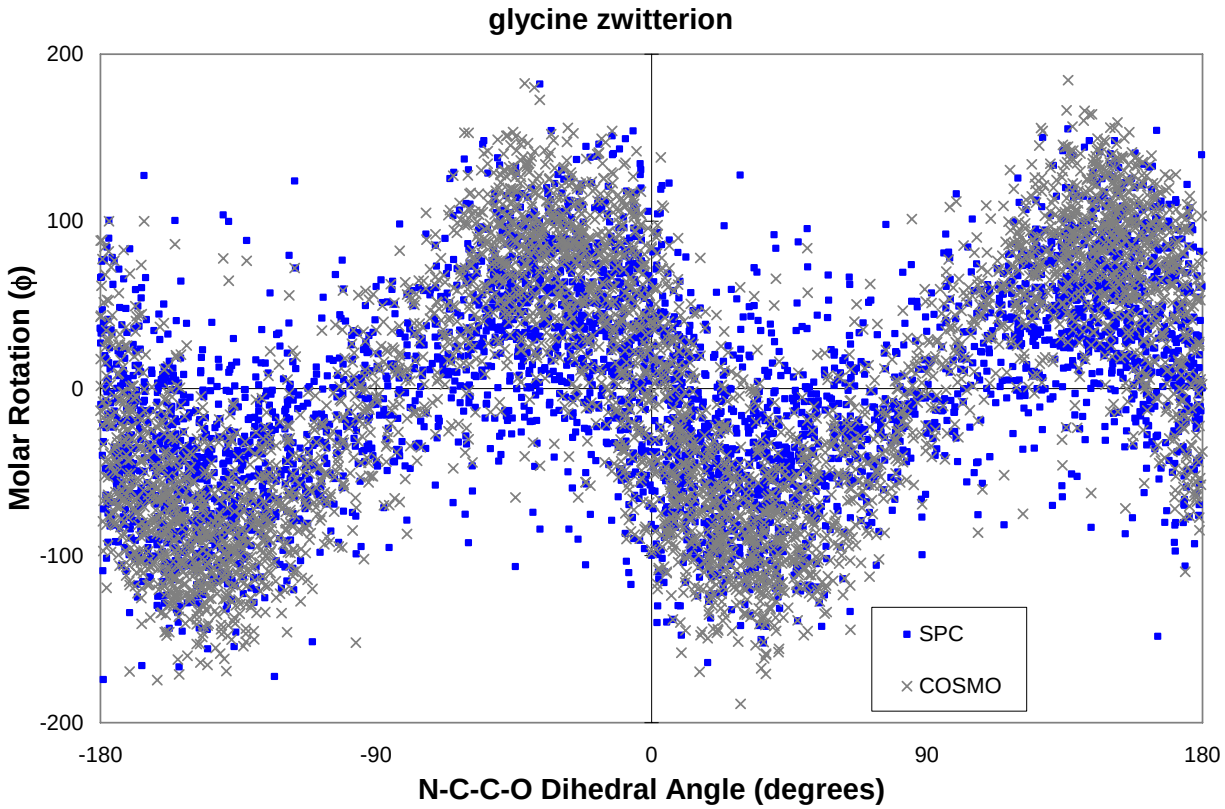


Figure 2-8: The molar rotation of a glycine system as a function of the N-C-C-O dihedral angle. The energies reported are for the same dynamics run, and include either 256 point charge based water molecules or the COSMO solvation model. All energies are reported relative to the average energy for the model. For the geometries the GROMACS program was used, and snapshots were recorded at 10ps intervals.

While the energy of a solvated glycine molecule is an *even* function of its N-C-C-O dihedral angle, the molar rotation is an *odd* function. The energy of the glycine molecule goes

up as the dihedral angle deviates from its ideal at zero degrees, regardless of the direction of the perturbation. The molar rotation, however deviates in equal and opposite directions depending on which way the molecule twists. This is in keeping with the sector rule for optical rotation. As the N-C-C-O dihedral changes, the functional groups that perturb the symmetry of the otherwise C_{2v} symmetrical carboxylate group move either above or below horizontal symmetry plane of that group. As the sign of the optical activity "sectors" are opposites on opposite sides of this plane, it makes sense that the molar rotation would change in opposing ways depending on to which side of the plane those perturbing groups deviate.

Also note that every time that the N-C-C-O dihedral angle is a multiple of 90 degrees, the molecule assumes approximate C_{2v} symmetry, but for other perturbations caused by the vibrations elsewhere in the molecule, as well as movement in the explicit solvation shell, where one is present. As such these angles represent "nodes" in oscillating molar rotation pattern. These are the geometries in which the glycine molecule is most symmetrical, and as such its molar rotations are the smallest in magnitude. Conversely, when the N-C-C-O dihedral angle deviates most from multiples of 90 degrees, the molar rotations tend to be the greatest.

CHAPTER 3: MODELING THE CHIROPTICAL PROPERTIES OF SMALL AMINO ACIDS IN SOLUTION

Kundrat, Matthew D.; Autschbach, Jochen. Time dependent density functional theory modeling of chiroptical properties of small amino acids in solution. Journal of Physical Chemistry A; **2006**; *110*(11) pp 4115-4123.

Kundrat, Matthew D.; Autschbach, Jochen. Comparison of methods of density functional theory modeling of the specific rotation of amino acid solutions. Poster presented at 230th ACS National Meeting, Washington, DC, Aug. 28-Sept. 1, 2005, COMP 215.

ABSTRACT

Time Dependent Density Functional Theory (TDDFT) and the COnductor-like Screening MOdel (COSMO) of solvation were used to model the specific rotation and Optical Rotatory Dispersion (ORD) of alanine, proline and serine solutions. Zwitterionic, cationic and anionic forms of amino acids were investigated and the results compared with experimental literature data obtained in neutral, acidic and basic conditions, respectively. It was found that TDDFT consistently underestimated the electronic excitations of the molecules, leading to calculated optical rotations that are of the correct sign but somewhat larger in magnitude than those of experiment. An additional challenge was encountered in the modeling of serine, an amino acid with a strong tendency to form intramolecular hydrogen bonds. The model used overestimated the extent of such hydrogen bonding for the zwitterions while possibly underestimating such bonding for the cationic form. This effect on the calculated mole fractions of the different conformers had an impact on the specific rotation.

INTRODUCTION

Any molecule whose structure lacks a plane of symmetry, center of inversion or any other improper axis of rotation is chiral and can be made as two stereoisomers, enantiomers that are mirror images of one another. Such enantiomers and identically prepared solutions thereof will rotate a plane of polarized light in equal an opposite directions. However without additional information the sign of the rotation alone is not enough to assign the absolute configuration of an enantiomer. Computational chemistry can provide this information.[1, 4, 6, 27] Early computational benchmarking on modeling optical rotation has been performed on molecules in the absence of solvent. From a computing perspective gas phase measurements are the easiest to model. However, experimentally most measurements are carried out in solution. This complicates the work of the computational chemist, as solvent effects may have a significant impact on the observed optical rotatory dispersion of a solution. Recent advances in this field have been made toward modeling chiroptical properties under the influence of solvent.[8, 9] Improvements in the accuracy of these modeling techniques have prompted an increasing number of chemists to rely on computational methods to assign absolute configurations.[10-13] Confidence in the computational methods has grown to the point that it has been used to "correct" older experimentally derived absolute configurations.[14]

Most of the molecules studied thus far can be found in approximately the same geometries in solution as in the gas phase. With such molecules one may save computational costs by optimizing the molecule in gas phase, then treating solvent effects on the energy and response properties afterwards. But not all molecules are so well behaved. Some molecules adopt significantly differing geometries in vapor and in solution. In the most difficult cases the

solution phase geometry is not even stable in the gas phase. The amino acids fall into this category of molecules, forming zwitterions in aqueous solution but reverting to their neutral form upon evaporation.

In 2004 Pecul et al. conducted a study of the conformational effects on the optical activity of two amino acids in the gas phase.[28] When comparing these calculations to experimental data gathered from aqueous solutions, they found that the specific rotations calculated at the Hartree-Fock (HF) level of theory were closer to experiment than those obtained at the Density Functional Theory (DFT) level despite the fact that DFT should and generally does yield optical rotations that are closer to experiment.[29] In particular, the authors found DFT to be "clearly unreliable" for the proline anion. The zwitterions were not studied. The authors did note that amino acids tend to be found in zwitterionic form in neutral aqueous solution and that the bulk of the optical activity caused by amino acid solutions is attributable to these forms. However, this issue was not investigated further.

Part of the purpose of our paper is to demonstrate that hybrid-DFT calculations employing large diffuse basis sets and an appropriate solvent model can provide reasonable optical rotation values for neutral, acidic and basic amino acids solutions. First we will test our computational methods on glycine, an achiral amino acid which we know should have an optical rotation of exactly zero. Next we will model the optical rotation of solutions of the smallest chiral amino acid alanine, which we will use to show some of the merits and shortcomings of DFT modeling of optical activity and show how through a cancellation of errors sometimes HF results can be closer to experiment than those of DFT when one looks at an optical rotation measured at a single frequency instead of a range of frequencies. Then we will model the solution phase optical activity of the slightly larger and more conformationally complex proline,

where we will begin to discuss the importance of Boltzmann-averaging optical rotations from different conformers in order to obtain results that best agree with those of experiment. Finally we will attempt to extend our model to serine, an amino acid whose –OH functional group makes it especially prone to intra-molecular hydrogen bonding, and we will show the challenges such interactions pose to our current method of modeling chiroptical response properties.

COMPUTATIONAL METHODS

All data were computed with the Turbomole[30] quantum chemical software, version 5.7. Except where otherwise noted all calculations were performed with B3-LYP[31] hybrid functional. All molecular geometries were optimized with the aug-cc-pVDZ basis set from the Turbomole library. All response calculations were with the d-aug-cc-pVDZ set[32], which previously has been shown to work quite well for TDDFT calculations of optical rotations.[21]

All optimizations and response calculations were performed with the COnductor-like Screening MOdel (COSMO)[33] of solvation. Solvent model parameters were configured using the cosmoprep program of the Turbomole package. The dielectric constant of the solvent was set to 78; all other solvent parameters were left at program default values. Default atomic radii (Bondi radii x 1.17)[34] were used.

Initial geometrical parameters were set using the Molden[35] graphical interface program and its default parameters. First the alanine zwitterion structure was drawn and optimized, afterwards Molden was used to modify this template into various conformations of the other amino acids that were themselves optimized. Cationic, anionic and neutral amino acid structures were derived from their corresponding optimized zwitterionic structures by adding, removing or changing the location of a hydrogen atom as appropriate followed by re-optimization. No

symmetry restrictions could be imposed on the structure of glycine during optimization since the COSMO code did not support symmetry.

Accuracy limitations[36] of COSMO apparently kept some of our geometries from meeting our convergence criteria during optimization attempts with the doubly augmented basis set, thus all reported geometries and zero-point energies were calculated with the singly augmented set. All structures were confirmed local minima having no imaginary vibrational frequencies as calculated with the NumForce program. This numerical method of frequency computation was used since the analytical frequency module of the software was incompatible with the COSMO solvation method. In this paper the energy, "D" of a particular conformer is defined as the sum of the electronic energy, the solvation energy from the COSMO model and the zero point energy calculated by NumForce. "ΔD" is defined as the energy of a particular conformer relative to the lowest energy conformation. Boltzmann factors were calculated based on this relative energy at the temperature of 293K. Different conformations of the same molecule generally had zero point energies and solvation energies that were within 1-2kJ/mol of one another.

Since the Turbomole code presently does not support the use of Gauge Including Atomic Orbitals (GIAOs, also referred to as London Atomic Orbitals) strictly speaking all calculated optical rotations are gauge origin dependent. Our gauge origin is defined as the center of mass in each molecule. Such origin dependency is known to diminish as the basis set size increases, and from a practical standpoint reasonably reliable results for small molecules are obtained by using the large augmented basis sets that are always needed to calculate reliable optical rotations.[3, 12] Data in the literature support this conclusion. Using the aug-cc-pVDZ basis set Pecul et al. reported that for proline in its neutral and cationic form, optical rotations varied by less than 5

deg cm^3 g^{-1} dm^{-1} when computed with and without GIAOs.[28] Earlier Ruud and Helgaker used the larger d-aug-cc-pVTZ basis set to ensure "near gauge-origin independence" for their non-GIAO calculations.[37] As we use the d-aug-cc-pVDZ basis here, we can infer that the origin dependence effect on our computed optical rotations should be small compared to other errors inherit in the calculations.

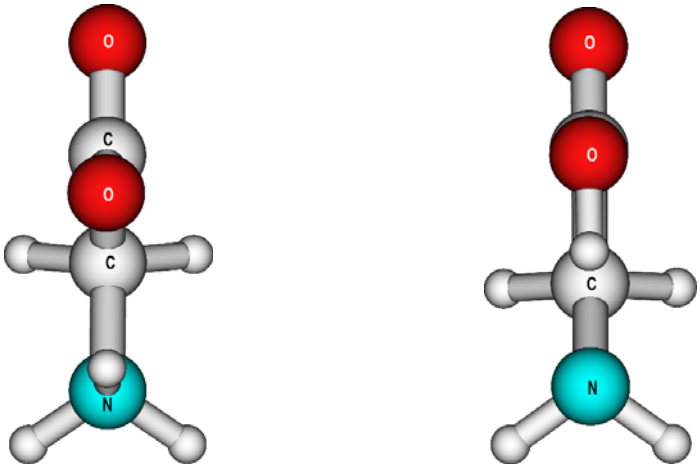
Except where otherwise noted specific rotations were calculated at the wavelength of the sodium D line (589.3nm). All specific rotations are reported in units of $deg \cdot cm^3/(g \cdot dm)$. Computed optical rotatory dispersion (ORD) curves were calculated at 10nm intervals from 600 to 220 nanometers. Experimental ORD plots were scanned from their respective graphics in the literature, digitized using the WinDIG program,[38] then converted from molecular rotation to specific rotation and plotted alongside the calculated curves.

RESULTS AND DISCUSSION

Glycine

Glycine is the smallest genetically encoded amino acid, and as the only achiral one it is the only one with a vanishing optical rotation. Much computational work has been published on glycine, and a good part of this work was directed toward correctly predicting the stability of the glycine zwitterion in solution. For example Jensen and Gordon have published the results of computational studies on the stabilization of the glycine zwitterion with explicit water molecules, what is sometimes referred to as a "discrete solvent model".[39] However without a thorough isotropic sampling of many solvent – solute configurations an arbitrary addition of solvent molecules around even an achiral solute can lead to an asymmetrical system which exhibits a significant optical rotation – that is the solvent molecules themselves cause optical activity. Continuum solvent models[40] such as COSMO do not have this drawback.

We performed some calculations with the COSMO solvation method to determine if it could stabilize the glycine zwitterion without the need for the addition of explicit water molecules. These attempts were successful, and we obtained the glycine zwitterion structure shown on the left side of figure 3-1. By this method the zwitterionic form of glycine is calculated to be more stable than the neutral form on the right by a ΔD of 4.6kJ/mole. This differs by nearly an order of magnitude from the experimental value of 30.4kJ/mole.[23] Several years ago Tortonda et al. obtained similar results in modeling neutral and zwitterionic glycine with an ellipsoidal cavity continuum solvent model, but nonetheless they succeeded in calculating an aqueous glycine infrared absorbance spectrum that was comparable to that of experiment.[41] Likewise for the purposes of the present paper any error in the neutral-zwitterion energy difference does not necessarily have a profound effect on the validity of the computed response properties. All that is needed is a method to calculate a reasonable zwitterionic amino acid structure that is energetically stable. The COSMO model appears to fulfill this requirement.



GlyZ Gly

Figure 3-1: Glycine neutral and zwitterionic forms optimized at the B3LYP + COSMO/aug-cc-pVDZ level of theory. Mirror plane is perpendicular to page.

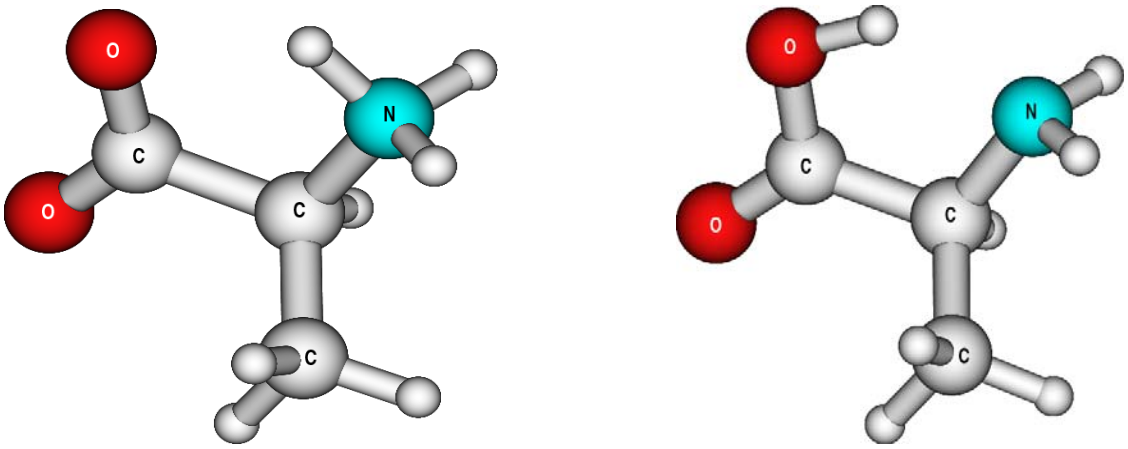
Solutions of glycine do not exhibit optical rotation. From it may be concluded that either: 1. the glycine molecule has a structure with a plane of symmetry or 2. there is a fast exchange between degenerate chiral conformers. The former case was found to be true as the optimized structure of the glycine zwitterion molecule appears to have C_s symmetry.

As the convergence criteria are tightened it is apparent that this molecule would converge to a structure that has a mirror plane. Unfortunately the imprecision inherit in the way that the COSMO model is implemented prevented the use of the tight geometry convergence criteria that one would like to use for the optimization of molecules. With the DFT grid tightened to "m5", the SCF convergence tightened to 10^{-7}, and the COSMO Number of geometrical Segments Per Atom (NSPA) number increased to 162, the molecular energy fluctuated by about 10^{-5} hartree by the end of the optimization cycle. Since these geometries inevitably did not converge as well as we intended, the optical rotatory data calculated based on those geometries also deviates from ideal. This is most apparent with glycine which we know should have a specific rotation of zero. The optimized glycine structure we obtained, which to the human eye does appear to be perfectly

symmetric, in fact has a O-C-C-N dihedral of 0.147 degrees instead of zero. This slight asymmetry results in a calculated specific rotation of -2.3 deg·cm^3/(g·dm). This number is indicative of approximately how much variance should be expected by computing specific rotations in case small numerical errors in the structure due to a solvation model and the DFT integration grid may occur.

Alanine

Alanine serves as the prototype chiral amino acid. Its relatively small number of electrons lends to rapid calculations of structures and response properties. In addition the fact that alanine has fewer atoms than the other chiral amino acids means that it will have fewer local minimum structures that need to be investigated. In fact for the alanine zwitterion we found only one minimum.



AlaZ Ala

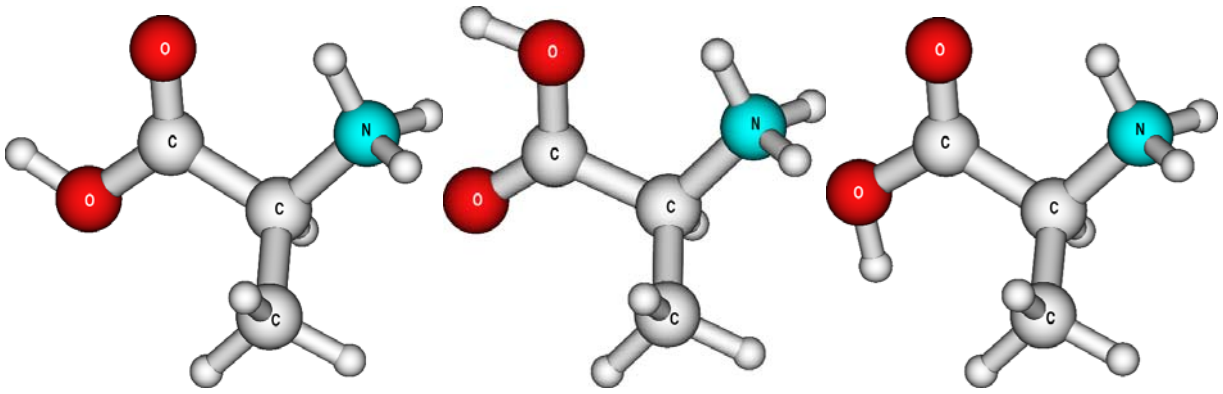
Figure 3-2: Optimized structures of alanine zwitterion (AlaZ) and neutral form (Ala).

For this zwitterionic structure we calculated a specific rotation of +4.0 deg·cm^3/(g·dm) which compares reasonably well with the experimental value of +2.42.[42] Among the common amino acids that are optically active, alanine has the smallest specific rotation in its zwitterionic form. Djerassi noted that this small rotation corresponds to the fact that two of the groups

attached to the alanine's chiral carbon, $-NH3^+$ and -CH3, are isoelectronic with one another and that the optical activity of the molecule must result from the charge difference of the N and C nuclei and the perturbation this causes to the electronic structure (otherwise the molecule would have a plane of symmetry).[43]

Thus at this initial stage we appear to have correctly calculated the sign of a very small specific rotation of a single molecule at a single wavelength caused by a small electronic perturbation of an otherwise symmetrical molecule. Without further information it would be fair to argue that such an agreement between calculation and experiment for a specific rotation so small in magnitude may merely have been the result of good fortune.[4] To gain more confidence in the method it would be desirable to compare the optical rotation of a molecule at various wavelengths closer to the excitation energy where the rotation is greater. Fortunately some experimental ORD plots for various amino acid solutions can be found in the literature, although those experiments were performed at a pH of one where alanine is found predominately in its cationic (protonated) form.

To facilitate comparison with the experimental the optimized geometries of the alanine cation were calculated. Protonation of the alanine zwitterion occurs at one of the oxygen atoms of the COO^- group. This reduces the site symmetry of the group from C_{2v} to C_s and makes possible a greater number of local minimum conformations for the cation than for the zwitterion. Three such minima were found here. They are depicted in figure 3-3.



I II III

Figure 3-3: Optimized local minimum structures of the alanine cation

Structure I is predicted to dominate at room temperature, with a Boltzmann weight of nearly 97%. As can be seen in Table 3-1 the contribution to the Boltzmann-averaged specific rotation from the two higher energy conformers is negligible. As such for the rest of this work the only cationic amino acid conformations that will be considered will be those analogous to structure I, the structures with the additional proton attached to the oxygen most distant from the amino group and orientated between the two carboxylate oxygen atoms.

Table 3-1: Relative energies and specific rotations of alanine cation conformers
B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann factors calculated from ΔD at 293.15K

Conformer	ΔD(kJ/mol)	Boltzmann factor	specific rotation ($deg \cdot cm^3/(g \cdot dm)$)	Boltzmann factor x specific rotation
I	0.0	0.97	34.6	33.4
II	8.3	0.03	-17.3	-0.5
III	12.5	0.01	22.3	0.1
Boltzmann Average Specific Rotation				33.0
Experimental Specific Rotation				13.7

Experimental data are from ref. 29, pp 203.

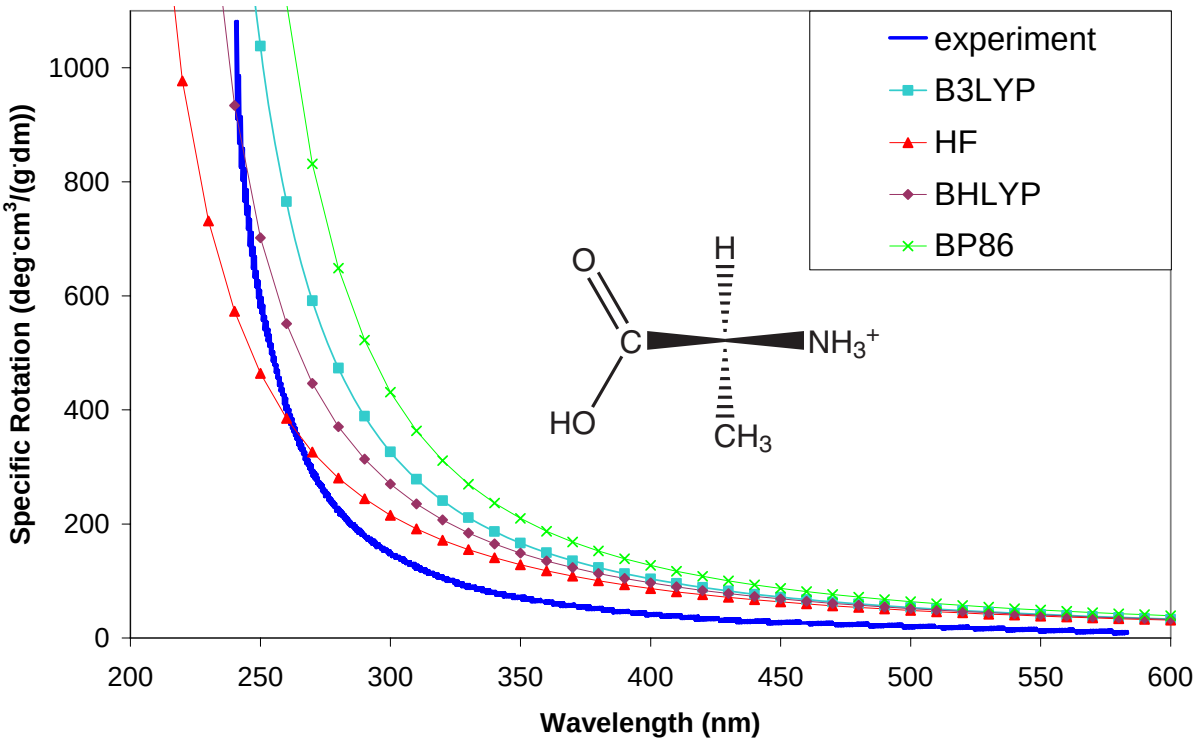


Figure 3-4: Specific rotation of cationic alanine as a function of wavelength. Methods used include Hartree-Fock (HF), Becke "half-and-half" LYP (BHLYP) and Becke-Perdew 86. Experimental data from ref. 30, pp 217.

Figure 3-4 shows that the shape of the ORD for alanine in acidic solution has been faithfully reproduced. A comparison of experimental measured optical rotation to our calculated value at 589.3nm would show that the calculated value is greater in magnitude than that of experiment; TDDFT is known to overestimate the magnitude of ORD with respect to both experiment and CCSD calculated results.[44] Such a deviation could arise from an overestimation of the magnitude of the rotatory strengths, an underestimation of the electronic excitation energies, or a combination of both.

What may also be concluded from the spectrum that may not be determined from single wavelength calculations is that the calculated optical rotations are not merely too high in

magnitude, but in fact the entire calculated ORD curve appears to be red-shifted compared to the experiment. This error also seems to be characteristic of B3LYP in general, as it is know to yield electronic excitation energies that are somewhat lower than experiment.[45] Conversely, the Hartree-Fock method is known to overestimate the electronic excitation energies, which results in a calculated ORD curve that is blue-shifted.[11] If the only error occurring were the underestimation of the electronic transition energies, changing the functional from DFT to HF should make the magnitude of the calculated specific rotation smaller than that of experiment. Since this is not the case for the alanine cation, it is apparent that some other error is present in the calculations that is causing the magnitude of the specific rotation to be overestimated, and changing to an uncorrelated method just partially cancels this error. It is not our aim to "tune" the fraction of Hartree-Fock exchange in the hybrid functional for best agreement with experiment as the B3LYP functional is already known to generally yield quite accurate linear response properties. We are simply pointing out how a cancellation of errors can sometimes make HF results appear closer to experiment if one only looks at the rotation at 589.3nm instead of modeling Since the B3LYP functional will be used throughout the rest of this work one must keep in mind that these predictable errors are likely to be present in the calculations of the more complex amino acids as well.

Table 3-2: Calculated lowest excitation wavelength and specific rotation for alanine
Geometries at B3LYP/aug-cc-pVDZ, response calculations with d-aug-cc-pVDZ basis

		HF	BHLYP	B3LYP	BP86
Cation	Excitation Wavelength (nm)	177.2	195.0	206.4	213.7
	Specific rotation at 589.3nm	32.6	34.1	34.6	40.9
Zwitterion	Excitation Wavelength (nm)	168.5*	192.3	215.6	239.2
	Specific rotation at 589.3nm	-1.2	-2.5	3.9	18.0

*- Sign of rotatory strength for lowest energy transition does not match sign calculated with the density functional methods

Proline

Next to alanine and the achiral glycine, proline is arguably the next easiest amino acid to model. Proline is ubiquitously referred to as an "amino acid" due to its biological role, though it is sometimes termed an "imino acid" as the nitrogen atom, bound in a five-membered ring is bonded to one less hydrogen atom than is the case for the other amino acids. The unique ring moiety formed by one nitrogen and four carbon atoms limits the conformational space of proline thus limiting the number of calculations that must be performed to model its optical rotation.

For the proline zwitterion only two conformations were found. These conformers differ by the direction of puckering of the five membered pyrrolidine ring, and can be identified by whether the γ carbon of the ring is bent towards the carboxylate group (endo) or away from the carboxylate group (exo). It is not completely clear from the literature whether the endo or the exo configuration is energetically favored at neutral pH.

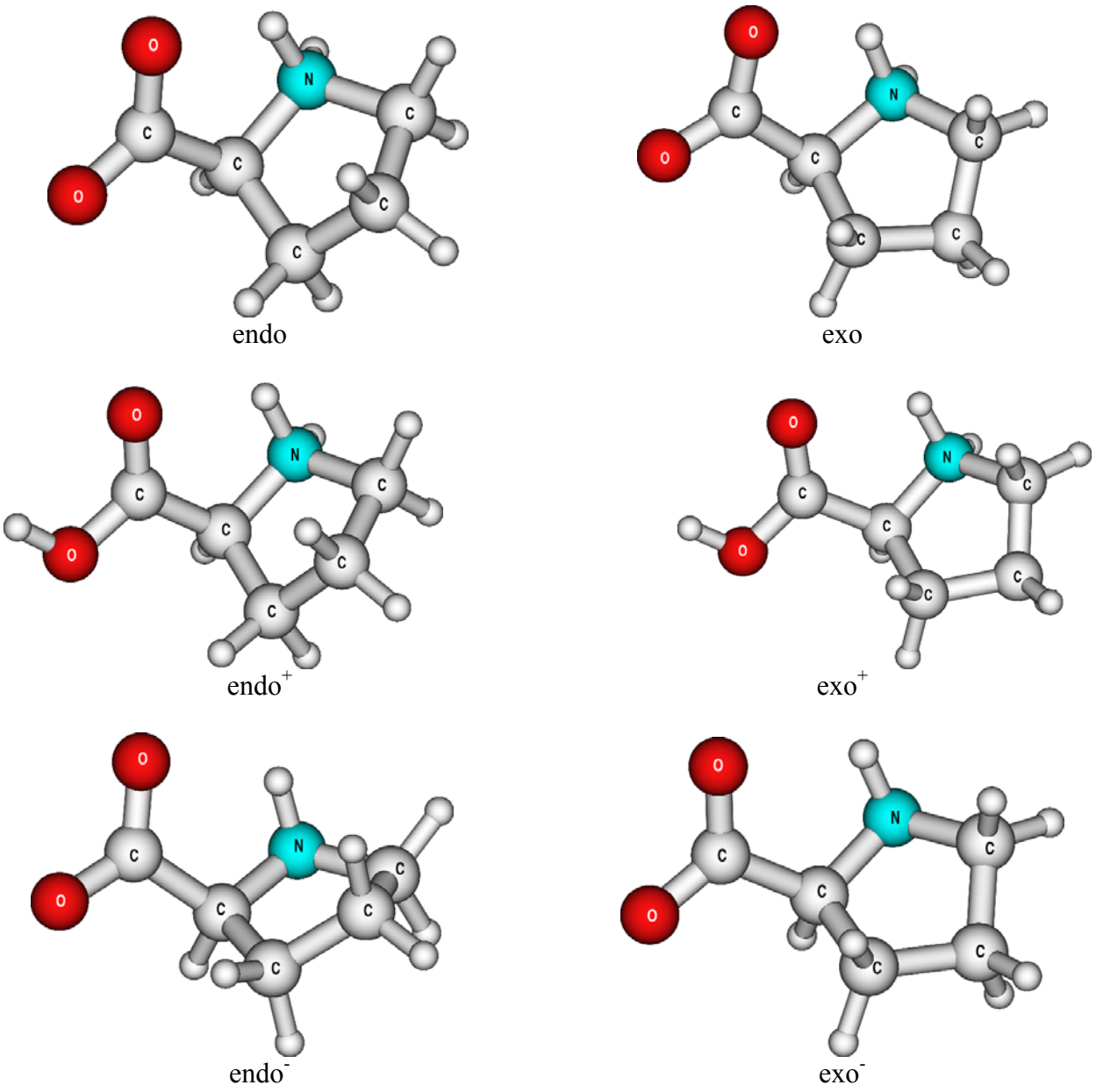


Figure 3-5: Low lying optimized structures of the proline zwitterion, cation and anion. A pair of higher energy anions with the imino hydrogen on the opposite side of the ring, as well as two pair of higher lying cation conformers with the COOH groups in configurations analogous to alanine structures II and III were also found, but since they did not have significant Boltzmann populations at room temperature they are not shown here.

In 1978 Jankowski et al. performed semi-empirical calculations on proline based on the Karplus equation and NMR data and found a ratio of endo to exo conformations of 63:37 at a pH of 7.2.[46] A few years later Haasnoot and coworkers used a similar method to deduce that these proline zwitterions have a 50:50 mole fraction in solution, and thus the ΔG between the two was precisely zero.[47] Stepanian et al. calculated that the in the vapor phase the neutral endo conformer is favored by 2.0 kJ/mole over the exo at the CCSD(T)/6-31++G** level of theory.[48] Pecul and coworkers performed gas-phase B3LYP/aug-cc-pVDZ calculations on proline and found that the lowest energy endo conformer was energetically favored over the exo by 1.63kJ/mole in the neutral form and 0.84kJ/mole in the protonated form.[28] Most recently Cappelli et al. calculated the two conformers to be less than 0.1kJ/mole apart using the B3LYP/6-31+G(d) method and the IEF-PCM solvent model, with the *exo* form actually becoming favored over the *endo* form by around 3.95kJ/mole when three explicit water molecules are added to the model.[49]

Our calculations indicate that the endo conformer may be slightly favored in the zwitterionic form, by a ΔD of 1.6kJ/mol. With only two conformers that are close in energy, even a small change in energy can alter the ratio of Boltzmann populations considerably. An average computed based on our calculated 66:34 ratio, which closely agrees with Jankowski's semi-empirical ratio, results in a specific rotation of -115.4 $deg \cdot cm^3/(g \cdot dm)$ for the zwitterionic form of proline. A 50:50 average suggested by Haasnoot's experimental data gives an average specific rotation of -128.8 $deg \cdot cm^3/(g \cdot dm)$, which is actually further from the experimental value of -85.0.[42] Regardless of the Boltzmann populations used here the product is always a calculated specific rotation that is of the correct sign and somewhat higher in magnitude than the

experimentally measured value. This could be expected based on previous experience with optical rotation calculations using the B3LYP functional, including our results for alanine.

Our calculations on the proline cation showed two conformers that are so close in energy that a conclusive assignment of the ground state is not possible, although it appears that the exo form is slightly favored over the endo by 0.4kJ/mole, which yields a 54:46 exo to endo ratio. Experimentally derived data agree that the endo and exo forms are equally populated at low pH.[47, 50] Cappelli et al. calculated a slightly higher energy difference that results in a 72:28 exo to endo ratio with the IEF-PCM model.[49] Pecul et al. calculated that the lowest energy *endo* conformation is favored over the *exo* by 0.84kJ/mole in gas phase.[28] As was the case with the zwitterionic form there is disagreement in the literature regarding the exo to endo ratio. Again our average specific rotation calculated for the proline cation is of the same sign as the experimental value regardless of the Boltzmann factors used. More insight can be gained by comparing the calculated ORD to experiment.

Table 3-3: Relative energies and specific rotations of proline conformers

B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann factors calculated from ΔD at 293.15K

Conformer	ΔD(kJ/mol)	Boltzmann factor	specific rotation ($deg \cdot cm^3/g \cdot dm$)	Boltzmann factor x specific rotation
Zwitterions				
endo	0.0	0.66	-86.7	-57.1
exo	1.6	0.34	-171.0	-58.3
Boltzmann Average Specific Rotation				-115.4
Experimental Specific Rotation				-85.0
Cations				
exo^+	0.0	0.54	-108.8	-59.1
$endo^+$	0.4	0.46	-46.4	-21.2
Boltzmann Average Specific Rotation				-80.3
Experimental Specific Rotation				-52.6
Anions				
$endo^-$	0.0	0.56	-74.3	-41.8
exo^-	0.6	0.44	-243.3	-106.5
Boltzmann Average Specific Rotation				-148.3
Experimental Specific Rotation				-93

Experimental data are from ref. 29, pp 7967.

As can be seen in figure 3-6 the ORD of the endo conformer is consistently lower in magnitude than that of experiment, and the ORD from the exo conformer is higher. Our Boltzmann averaged ORD provides a better fit to experiment than the exo or endo conformers alone and better fits the shape of the ORD, including the trough around 270nm. Generally we see the same consistent deviation from experiment we saw with alanine: the calculated values are more intense and the calculated ORD is red shifted.

At wavelengths lower than about 250nm the dependence of the optical rotation on the wavelength becomes very strong, and the numerical agreement with experiment worsens. This is a predictable result of the way in which the ORD is calculated. The computational code used in this work does not take into account the finite lifetimes of the electronic states. As a result the

ORD curves for the endo and exo conformers each will exhibit a singularity as they approach their excitation energies, calculated to be at 207.2 and 208.3 nanometers, respectively, with positive rotatory strengths.

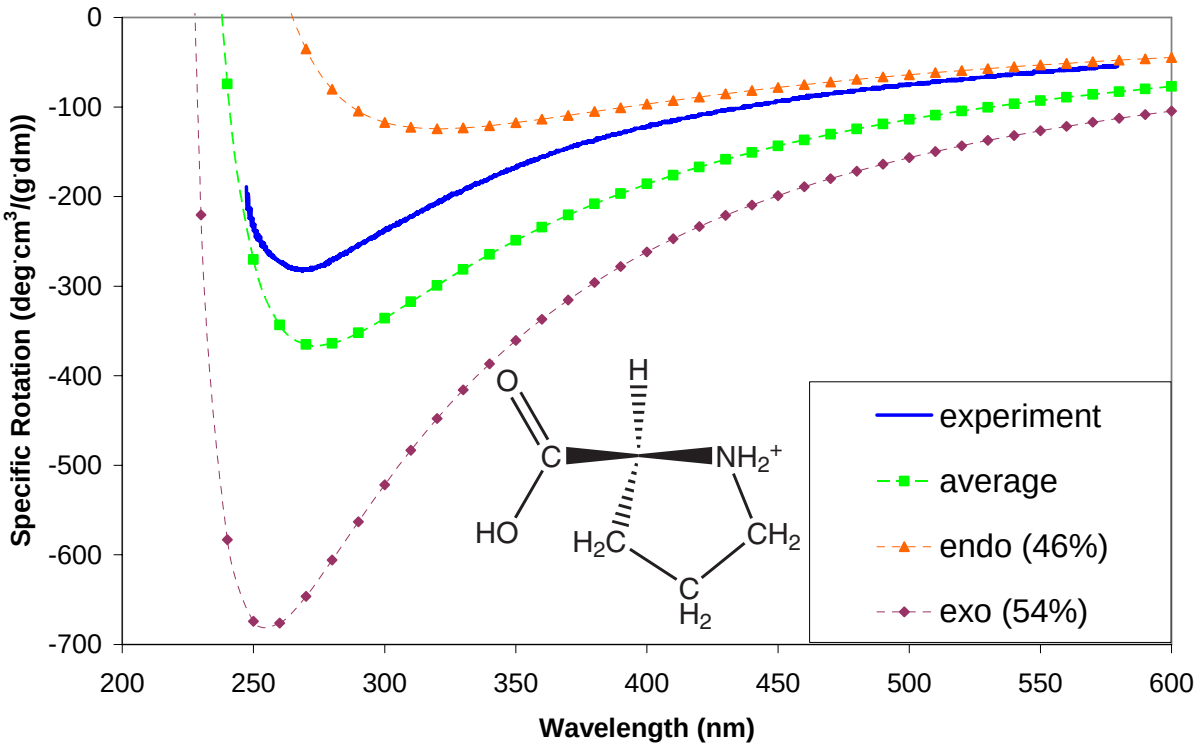


Figure 3-6: Specific rotation of cationic proline as a function of wavelength Experimental data is from ref. 30, pp 222.

Experimentally one would expect the typical bisignate ORD close to 200 nm corresponding to a positive Cotton Effect. On the other hand the trough at 270 nm as well as the long wavelength limit of the optical rotation must be dominated by high lying excitations with positive rotatory strength and are thus expected to be reproduced well by computations without finite-lifetime damping such as performed here.

Specific rotation data at high pH is also available in the literature for proline. Consequently, calculations on anionic proline in solution were performed to compare those

results with experimental values. Four low lying minimum structures were found differing from one another by exo/endo ring puckering and by whether the imino proton is on the same or opposing side of the ring as the carboxylate group. The two structures with the imino proton *cis* to the carboxylate group were both calculated to be about 12kJ/mole lower in energy than the lowest lying *trans* configuration; therefore only data from these two lowest lying conformers are reported hereafter.

Of these two structures the endo configuration is calculated to be favored over the exo by 0.6kJ/mol, yielding a 56:44 endo to exo ratio at room temperature. Jankowski et al. derived from NMR measurements at a pH of 12.7 that this endo to exo ratio should be 86:14.[46] Haasnoot disputed the validity of the trend cited by Jankowski that the endo configuration becomes more favored at higher pH and instead asserts that a 50:50 ratio is equally valid at all pH values.[47] Our energy calculations by themselves do not have the high level of precision that is needed to conclude which experimentally derived ratio is more correct. However some insight may be gained by comparing the effect these differing proposed exo/endo ratios have on the average specific rotation.

When the optical rotations of the exo and endo forms of proline in various ionization states are averaged based on the populations calculated in this work, a trend of increasingly negative specific rotation with increasing pH can be seen (Fig. 7). This is in keeping with experiment. If the experimentally derived 50:50 ratio[47] is used to average our data from the respective conformers, this trend still holds true, although a somewhat larger deviation from experiment is seen for the more alkaline solutions. However if the experimental ratios suggesting that the endo form becomes much more populated at high pH[46] are used to average

our optical rotation data, then the specific rotation of proline is calculated to become less negative with increasing pH, and this is inconsistent with experiment.

In the absolute sense the average specific rotations calculated in this work deviate more for the anionic form than for other forms of proline. Recently Pecul et al. modeled the specific rotation of anionic proline with TDDFT and found the results unacceptable. Their calculated optical rotation was based almost entirely upon what in this work is termed the exo$^-$ conformer, which has an optical rotation that is much higher in magnitude than that of the lowest energy endo$^-$ conformer. Unfortunately Ref. 12 did not consider data from the endo$^-$ conformer. As a result their calculated specific rotation was several times higher in magnitude than that of experiment. Based on this calculated optical rotation the authors concluded that, "for the anionic form of proline, the DFT values are clearly unreliable because the excitation energies are strongly underestimated, probably because of the DFT self-interaction problem."[28]

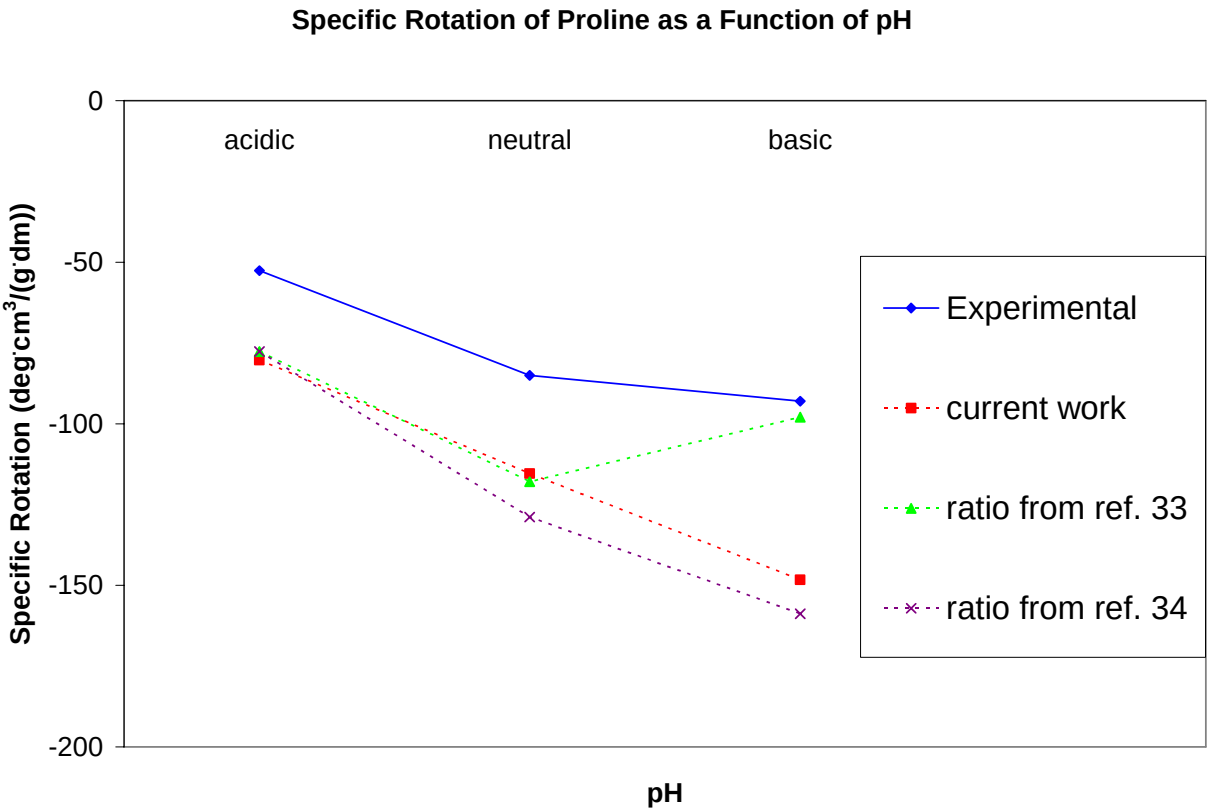


Figure 3-7: The pH effect on optical rotation data for proline. Experimental values from ref. 29, pp 7967. Experimentally derived exo/endo ratios used to compute averages are from refs 33 and 34.

The results in this instant work do not lead to the same conclusion. It is true that as one proceeds from modeling cations to modeling anions the lowest excitation energies decrease. Thus in a first approximation one might expect the relative errors in the optical rotation calculations to increase along the cation/zwitterion/anion series. However with a large diffuse basis set, modeling of solvent effects, and based on a conformational search that includes the endo$^-$ conformer our data indicate that the specific rotation of the proline anion can be modeled about as well as that of the cation or zwitterion. The data in table 3-3 show that the calculated specific rotations for the proline cation and anion both are of the correct sign and exceed their experimentally measured counterparts by just over 50% in magnitude. The underestimation of

the excitation energy seems to affect the modeling of cationic, zwitterionic and anionic proline's specific rotation relatively equally.

Serine

Serine can be considered a derivative of alanine. In terms of number of atoms and molecular weight it is the smallest of the chiral amino acids after alanine and is thus attractive from a computational point of view. However, in contrast to proline with its atoms bound in a ring, all of the bonds in serine are relatively free to rotate. This gives rise to a number of possible conformations. The primary rotamers, formed by rotation about the C_α-C_β bond are depicted in figure 3-8.

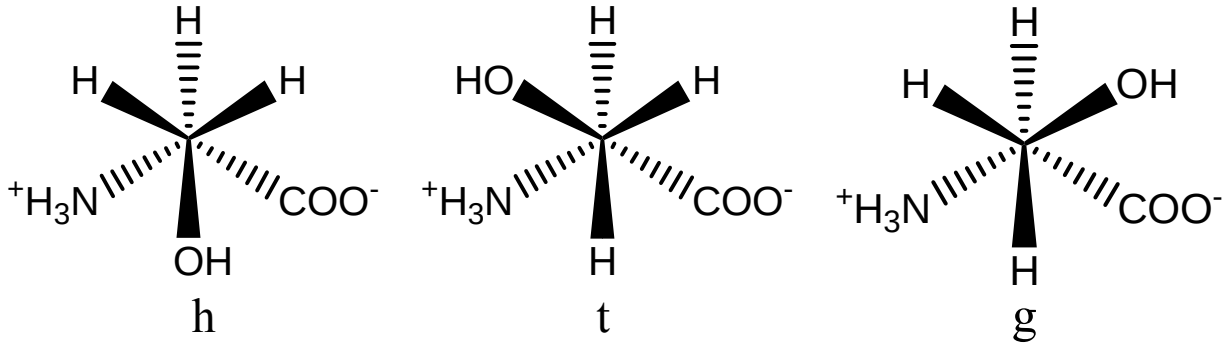


Figure 3-8: Primary rotamers of serine; page is perpendicular to the C_β-C_α bond.

The naming system is adopted from the works of Marten et al.[51, 52] The "t" and "g" refer to situations where the *largest groups* are trans and gauche to one another; this differs from the definitions used in other naming conventions.[53] The "h" designation apparently stems from the word "hindered" which was assigned at a time when it was presumed that the configuration with all of the largest functional groups adjacent to one another would be the least favored conformation. Ironically, experiments have since shown that the h configuration is in fact the most energetically favored conformation for serine, followed by the t and the g.[54, 55]

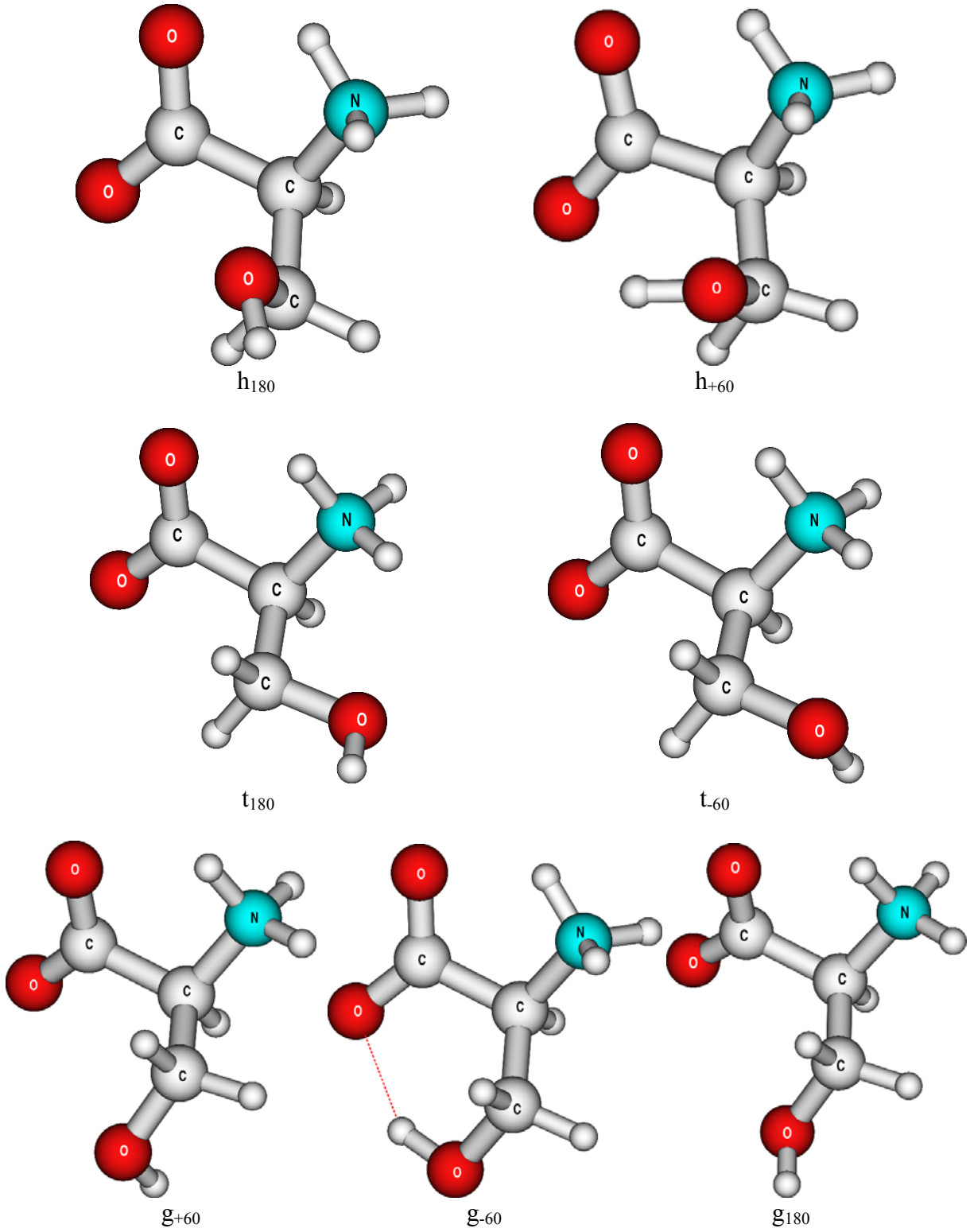


Figure 3-9: Optimized structures of the serine zwitterion. Structure g_{-60} depicts the strongly hydrogen bonded conformation with an O-H distance = 1.87A

For each of the primary rotamers depicted in Figure 3-7 there can be three possible subrotamers due a rotation of the –OH group. This could give rise to up to nine local minima. In actuality only seven were found, since one each of the h and t rotamer structures did not converge to a minimum due to a steric interaction between the $-NH_3^+$ and the –OH groups. The optimized structures are depicted in figure 3-9. The subrotamers are designated as +60, -60 and 180, which refers to the approximate value of the H-O-C_β-C_α dihedral angle.

Our calculated Boltzmann factors generally follow the trend that the total mole fraction of the h rotamers exceeds that of the t's which in turn exceeds that of the g's. However this agreement is not as good as one would like it to be. By far the biggest source of error has been the intramolecular hydrogen bonding within the g rotamers.

In modeling the average specific rotation of serine, the unusually low energy of one conformer alone was sufficient to disturb our Boltzmann averaging. This structure is depicted as "g_{-60}" in figure 3-9. It contains an intramolecular hydrogen bond between the –OH and the -COO^- groups. This so stabilized this configuration that this structure is predicted to be the overall ground state for serine. It is the energy of this structure that caused the computed Boltzmann population of the g rotamers to be more than twice the value derived from Noszal's experiments, a value in line with other experimentally derived data.[54] Given this consistency and the assertion that such NMR coupling constant derived rotamer populations should error by no more than $\pm$8%[56] we may conclude that our calculated g rotamer population is indeed significantly higher than it should be, and that it is an overestimation of the extent of intramolecular hydrogen bonding that caused this deviation.

Table 3-4: Boltzmann Populations of Serine Rotamers
B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, calculated from ΔD at 293.15K

		h	T	g
Serine Zwitterion	Computed	0.47	0.27	.26*
	Experimental	0.61	0.27	0.11
Serine Cation	Computed	0.79	0.21	<.01
	Experimental	0.76	0.16	0.08

*- almost exclusively from the "hydrogen bonded" configuration, g_{-60}
Experimentally derived populations are from Nozsal et al.[55]

With proline such a deviation between computed and mole fractions of conformers would not have affected the validity of the computed sign of the specific rotation, since all of the relevant conformers had the same specific rotation sign. However with molecules that can adopt conformations with differing signs of large-magnitude optical rotation changes in calculated conformational populations can alter the sign of the calculated average specific rotation if the resultant value is small. Consequently errors in the calculated energies can alter the sign of the calculated specific rotation, which is what occurred with our serine zwitterion calculations.

The average specific rotation calculated for zwitterionic serine is +6.3 $deg \cdot cm^3/(g \cdot dm)$, while the experimental value from the literature is -6.83. The deviation from experiment in the calculated mole fractions (Boltzmann populations) explains this error. Table 3-4 shows that for the serine zwitterion we calculate a population of the g rotamer that is more than twice that calculated from experiment. This inflated g population comes at the expense of the population of the t rotamers, which is therefore calculated to be significantly lower than that of experiment. The g rotamers are calculated to give a large positive optical rotation, +56.3 $deg \cdot cm^3/(g \cdot dm)$. The h rotamers contribute a much smaller positive rotation, +7.2 $deg \cdot cm^3/(g \cdot dm)$. If our computed Boltzmann factors of .26 for g and .47 for t are used, the resulting contributions to the average, +14.8 $deg \cdot cm^3/(g \cdot dm)$ for the g and +3.4 for the h, added together are greater than the -11.9 $deg \cdot cm^3/(g \cdot dm)$ contribution from the levorotatory t rotamers and cause the average calculated

specific rotation to be positive. However when Noszal's experimentally derived Boltzmann populations were imposed on our results, the g rotamers then contribute only 6.2 deg·cm^3/(g·dm), the contribution from the h rotamers would increase to 4.4 deg·cm^3/(g·dm), and with these added to the -11.9 deg·cm^3/(g·dm) from the t rotamers yields an average specific rotation of -1.3 deg·cm^3/(g·dm), which agrees in sign with the experimentally measured value. It should be noted that the optical rotation of several of the conformers were an order of magnitude larger than the experimental average. Thus an accurate average is difficult to obtain.

Table 3-5: Relative energies and specific rotations of serine conformers
B3LYP/d-aug-cc-pVDZ//B3LYP/aug-cc-pVDZ, Boltzmann factors calculated from ΔD at 293.15K

Conformer	ΔD(kJ/mol	Boltzmann factor	specific rotation (deg·cm^3/g·dm)	Boltzmann factor x specific rotation
Zwitterions				
g_{-60}	0.0	0.26	56.3	14.8
h_{180}	0.2	0.24	10.8	2.6
h_{+60}	0.4	0.23	3.5	0.8
t_{180}	1.5	0.14	-28.3	-3.9
t_{-60}	1.8	0.13	-64.0	-8.0
g_{+60}	18.4	0.00	144.6	0.0
g_{180}	19.9	0.00	87.7	0.0
Boltzmann Average Specific Rotation				6.3
Experimental Specific Rotation				-6.83
Cations				
h_{180}^+	0.0	0.70	33.5	23.5
t_{-60}^+	4.1	0.13	-21.9	-2.8
h_{+60}^+	5.0	0.09	35.0	3.0
t_{180}^+	5.0	0.08	-17.3	-1.5
g_{-60}^+	15.7	0.00	83.9	0.1
g_{+60}^+	18.3	0.00	198.2	0.1
g_{180}^+	19.1	0.00	109.6	0.0
Boltzmann Average Specific Rotation				22.4
Experimental Specific Rotation				14.95

Experimental data are from the ref. 29, pp 8605.

The opposite problem arises when the cations of serine are studied - the method *underestimates* the stability of the g cationic rotamers. However the correct sign for the optical

rotation of the serine cation is still calculated at all frequencies, due to a cancellation of errors. As can be seen in figure 3-10, if the g population had been as high as it should have been, the Boltzmann average plot would be more in keeping with that of alanine and proline, that is the calculated ORD would be too large in magnitude and red shifted.

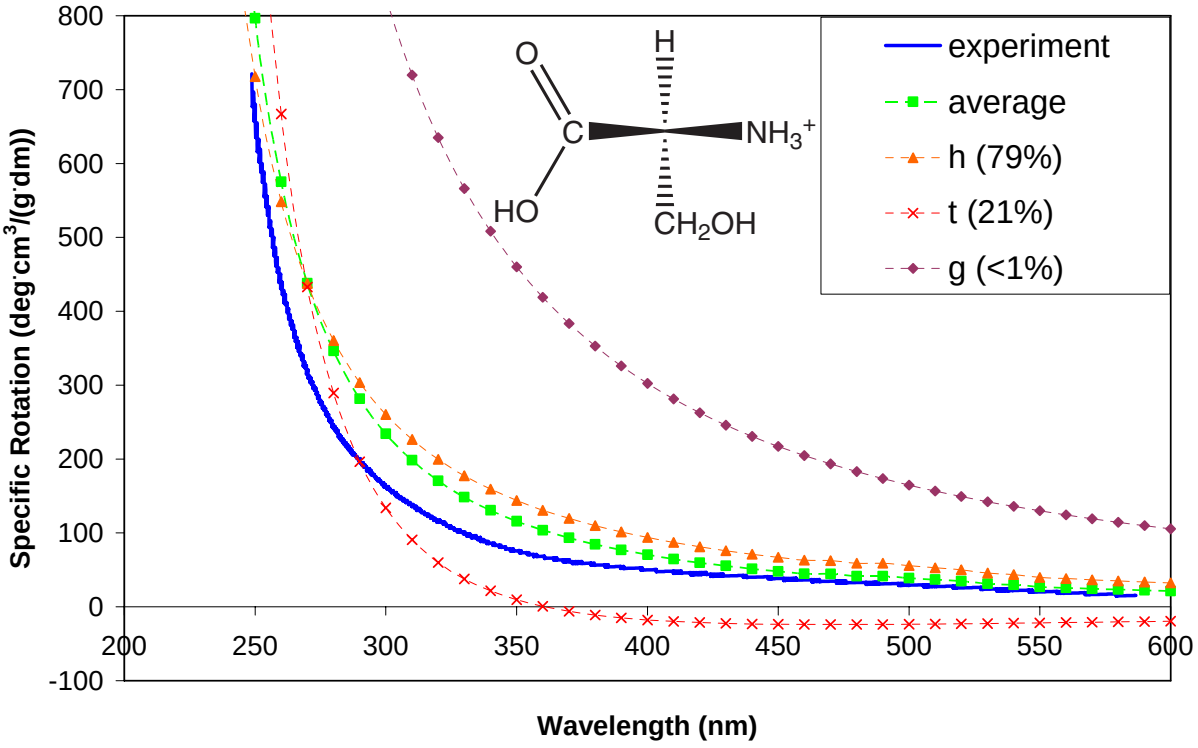


Figure 3-10: Specific rotation of cationic serine as a function of wavelength Experimental data from ref. 30, pp 217. Plots labeled "h", "t" and "g" are each Boltzmann averages of the optical rotations calculated for the two or three respective subrotamers.

Further insight may be gained by comparing the relative energies of our serine cation conformers with the calculations in the literature. Noguera and coworkers obtained the same relative energies for the primary rotamers as were obtained here in agreement with experiment: h is more stable than t that is more stable than g.[57] Their calculations did not consider solvent effects, and they found the lowest lying g conformer to be nearly 36.8kJ/mol higher in energy than the h ground state. Our calculations indicate this difference to be 15.7kJ/mol. It must be

noted that the authors of ref. 44 did not include some of the low energy local minimum geometries caused by rotation of the –OH group. However they reported the structure that we calculate to be the most stable g conformer, which is far too high in energy to be consistent with the population of the g rotamer observed by experiment. From this it may be safely concluded that it is not the solvent model that causes the deviations seen in our relative energies. Instead those deviations come from approximations that would already affect gas-phase DFT calculations. The COSMO model appears to be partially compensating for such error, though not enough to bring our calculations in line with experiment.

CONCLUSIONS

Time-dependent density functional theory response calculations have been shown capable to reproduce the optical rotatory dispersion of cationic, zwitterionic and anionic amino acids in solution with similar margins of error. The method has been used to correctly reproduce the sign of the optical rotation of small amino acids, and to faithfully reproduce the dispersion of the optical rotation of such molecules. A consistent source of error for this method appears to be the underestimation of the electronic excitation energies, which leads to an overestimation of the specific rotation of the molecules at 589 nm and an overall red-shift of their ORD curves.

The usefulness of this method is limited by the ability to predict correct relative energies and the resulting Boltzmann factors for the accessible geometric conformations. The COSMO model seems work reasonably well for this purpose, but only in cases where intramolecular hydrogen bonding is not a major factor. When intramolecular hydrogen bonding comes into play, DFT coupled with COSMO tends to overestimate the magnitude of the bonding in the zwitterions and underestimates it in the cations. If the overall optical rotation is determined by a cancellation of large contributions from individual conformers the relative errors in the energies

can lead to comparatively large deviations from the experimental optical rotation. It is likely that some explicit solvation is necessary to obtain the correct balance between intramolecular and solvent-solute hydrogen bonding.

ACKNOWLEDGMENT

The authors would like to acknowledge support from the Center for Computational Research (CCR) at the University at Buffalo for computational resources. J.A. is grateful for financial support from the ACS Petroleum Research Fund and from the CAREER program of the National Science Foundation (CHE-0447321).

CHAPTER 4: SUM OVER STATES AND KRAMERS-KRONIG MODELING

Krykunov, Mykhaylo; Kundrat, Matthew D.; Autschbach, Jochen. Calculation of CD spectra from optical rotatory dispersion, and vice versa, as complementary tools for theoretical studies of optical activity using time-dependent density functional theory. Journal of Chemical Physics; **2006**; *125* pp 194110-194122.

Comparison of ORD from KK transformation with experimental ORD from literature

This chapter is the only published chapter of my thesis which constitutes not a full paper but a subset of a paper I co-authored with a former colleague. The idea here was to use the code the Mykhaylo developed to calculate an Optical Rotatory Dispersion (ORD) curve via a Kramer-Kronig (KK) transformation of a Circular Dichroism (CD) spectrum calculated by the Sum Over States (SOS) method. This constitutes my contribution to that work.

For this it seemed appropriate to compare ORD from a KK transform of a computed CD spectrum with an ORD curve obtained experimentally. For this endeavor proline was chosen as the model compound, as it has a limited number of conformations and an optical activity of a relatively large magnitude of which much experimental data can be found in the scientific literature. Since the most extensive ORD data was obtained at a pH of 1 or less where proline is cationic, the optical response of protonated proline has been calculated.

The 500 lowest energy CD excitations of each of the two conformers were computed at the B3LYP/d-aug-cc-pVDZ level following the procedures of previous work.[58] The CD spectrum was plotted using a line width[59] of Γ=7.5 cm^{-1} and the broadened CD intensity was then transformed. The resulting dispersion curves are depicted in figure 4-1. The two conformers are designated “endo” and “exo”, which refer to the configuration of the γ-carbon. These two conformers were earlier calculated to have relative populations of 46% and 54% respectively at room temperature.[58] These factors were used for computing the “average” plotted ORD.

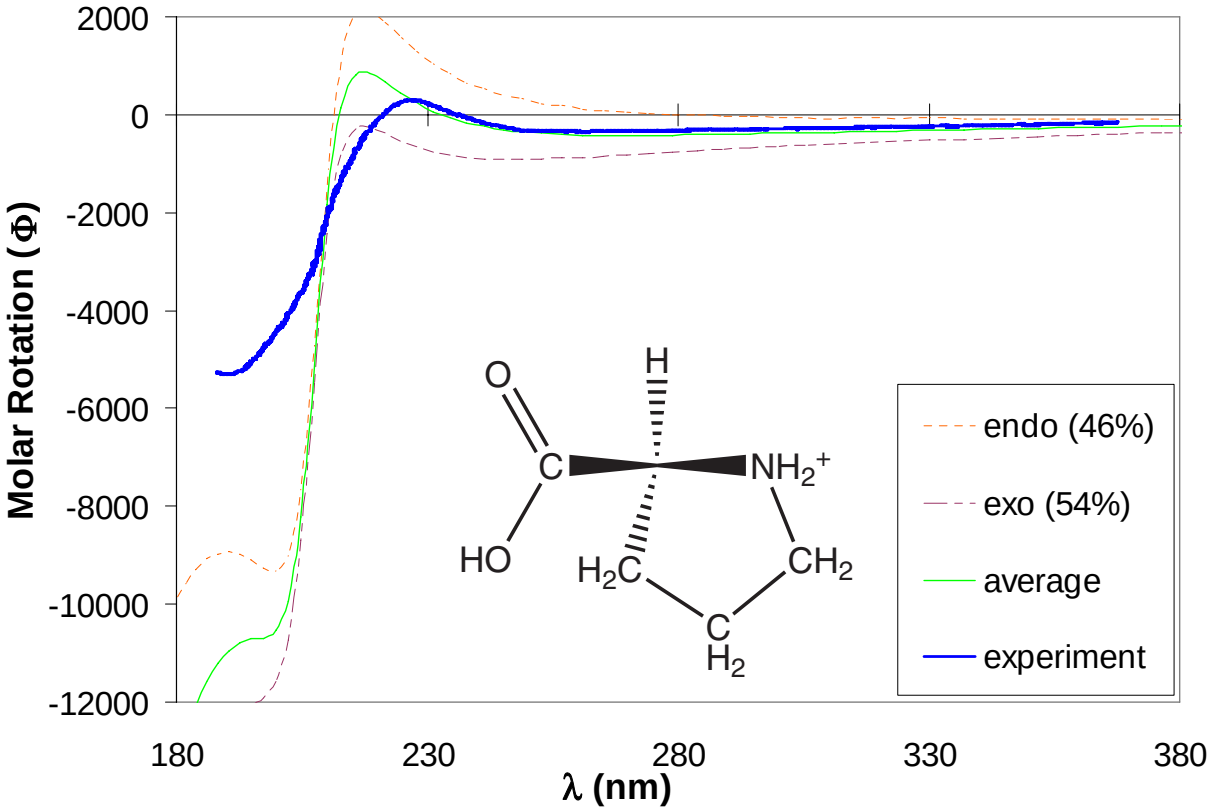


Figure 4-1: Computed ORD of the two conformers of cationic proline and the Boltzmann averaged ORD depicted alongside the experimental ORD from Iizuka and Yang.[60] The 500 lowest energy CD excitations were used for the calculation.

The KK transform performs adequately in modeling the ORD of proline in the near ultraviolet region. The trough around 260nm is accurately reproduced. The sign of the first two cotton effects also match those from experiment. The intensity of those effects are not in great agreement, however this intensity is largely dependent on the empirical broadening factor that has been applied to the excitations. This allows the ORD intensities to be varied somewhat arbitrarily by manipulation of the CD broadening factor.

The effect of truncation error was also explored. This results from a finite number of CD excitations being used to calculate a rotatory dispersion. In theory this error should be minimized when the sum of the rotatory strengths approaches zero. A graphical depiction of how molar rotation varies with the sum of rotatory strengths is shown in figure 4-2.

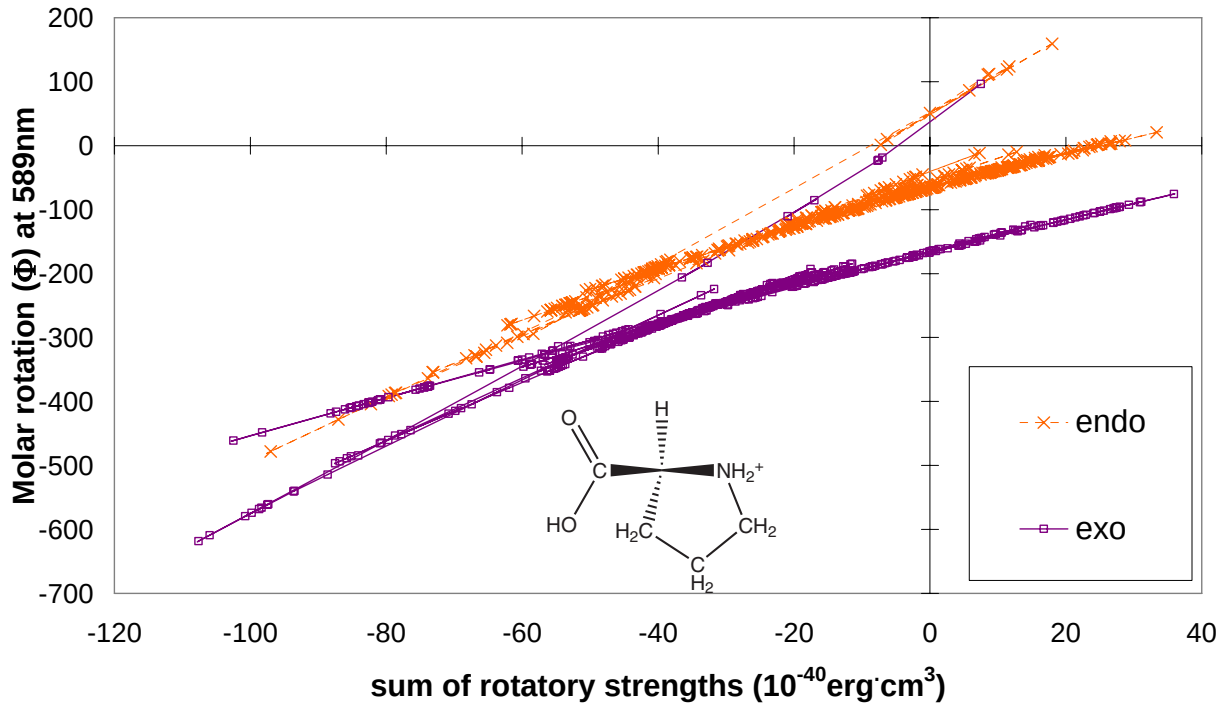


Figure 4-2: The molar rotation of each conformer of cationic proline plotted as a function of the sum of rotatory strengths. No damping factor was used.

One factor that is immediately apparent is the strong correlation of the molar rotation with the sum of the rotatory strengths. This dependence becomes stronger as the number of excitations increases, and has become approximately linear after approximately the first 100 excitations have been considered. But the data that should be focused on here is the molar rotation when the sum of rotatory strengths approaches zero, as this is the point where truncation error should be minimized. Here the molar rotation computed by KK transform of the CD should approach the molar rotation computed by direct linear response. At this point we find a molar rotation of about -66 and -165 for the endo and exo forms of proline, respectively. To facilitate comparison with the molar rotation calculated by linear response with infinite lifetimes, no broadening factor was used in this plot. For comparison the ORD of these two forms was calculated to be -53.4 for the endo and -125.3 for the endo. This deviation is consistent with the

fact that the sum rule does not apply perfectly to a finite number of excitations calculated with a finite basis.

Here I have succeeded in modeling the anomolous region ORD of the proline cation through the KK transformation of a truncated CD spectrum. The Cotton effects associated with the electronic excitations were reproduced with reasonable accuracy. The sodium D line molar rotation approaches that calculated via the linear response method when the sum of the states used approaches zero, in keeping with the sum rule.

CHAPTER 5: MODELING THE CHIROPTICAL PROPERTIES OF THE AROMATIC AMINO ACIDS

Kundrat, Matthew D.; Autschbach, Jochen. Time dependent density functional theory modeling of specific rotation and optical rotatory dispersion of the aromatic amino acids in solution. Journal of Physical Chemistry A; **2006**; *110*(47) pp 12908-12917.

Kundrat, Matthew D.; Autschbach, Jochen. Time-dependent density functional theory modeling of the chiroptical properties of aromatic amino acids in solution. Poster presented at 38th Midwest Theoretical Chemistry Conference, Columbus, OH, June 15-17, 2006.

ABSTRACT

Time Dependent Density Functional Theory (TDDFT) along with the COnductor-like Screening MOdel (COSMO) has been applied to model the specific rotation at 589.3nm and the optical rotatory dispersion (ORD) of the aromatic amino acids phenylalanine, tyrosine, histidine and tryptophan. Solution structures at low, neutral and high pH were determined. Both the anomalous dispersion absorbing (resonance) region and the lower energy (transparent) region of the ORD of the compounds were modeled. Linear response calculation of the specific rotation and ORD as well as Kramers-Kronig transformations of calculated circular dichroism spectra to model resonant ORD were compared with experimental data from the literature.

INTRODUCTION

All naturally occurring amino acids save the smallest, glycine, are chiral compounds. Each enantiomer of such a chiral compound can be distinguished from its mirror image by its specific rotation. Traditionally, specific rotation has been measured using the yellow sodium D-line of light, to which these colorless amino acids are transparent. At this wavelength the specific rotation is expected to be strongly dependent on the lowest allowed electronic excitation. Historically, this transition has been identified as the $n \rightarrow \pi^*$ transition of the carboxylate functional group that is present in all of the free amino acids.[61]

The aromatic amino acids: phenylalanine, tyrosine, histidine and tryptophan differ from their aliphatic analogs in that in addition to this carboxylate chromophore they also possess an aromatic ring chromophore. The $\pi \rightarrow \pi^*$ transitions of the aromatic system have an additional effect on the chiroptical properties of these building blocks of proteins. Such transitions are of interest since they are primarily responsible for the circular dichroism (CD) features of proteins in the near ultraviolet (UV).[62] There is a growing interest in using time-dependent density functional theory (TDDFT) to model this phenomenon. In fact, while this manuscript was in preparation Tanaka and coworkers published an article on using TDDFT to model the electronic and vibrational CD of these very amino acids.[63]

That contemporary work by Tanaka et al. focused on modeling the configurations of the aromatic amino acids that are found in the random coil configuration of polypeptides, structures that are relevant to protein modeling. This current work focuses instead on the zwitterionic form of the molecules and their various protonated and deprotonated forms in which these amino acids can be found in dilute aqueous solutions of the molecules. Also while Tanaka et al focused on modeling the absorptive chiroptical properties of these compounds, the ECD and VCD, this work

instead models the dispersive properties, the specific rotation and the optical rotatory dispersion (ORD), and how various chromophores affect these properties. This work constitutes a continuation of our previous work modeling specific rotation and ORD of the small aliphatic amino acids: alanine, proline and serine where the carboxylate chromophore was primarily responsible for the chiroptical response.[58] Here our calculations are extended to the larger, aromatic amino acids, in which two distinct chromophores contribute to the optical rotation.

To faithfully model chiroptical response properties one must first correctly determine the structures of the molecules being studied, and so this paper begins with a discussion of the optimized geometries of the amino acids. Some attention will be paid to the basis set effects on the relative energies of these geometries, and the relationship between basis set and the difficulty in correctly modeling the extent of intramolecular hydrogen bonding will be noted briefly. The computed mole fractions (Boltzmann populations) of the various conformers of some of these amino acids will be compared with experimentally derived Boltzmann populations from the literature. Next the specific rotations of select ionic states of these amino acids will be computed and the results compared with experimental rotations. Particular attention will be paid to how the two different chromophores affect specific rotation, and how this varies depending on the conformation of the molecule. For some cases where the sign of the computed and measured specific rotation do not agree at 589nm, it will be demonstrated how comparison of computed and measured optical rotatory dispersion curves would be a better method for assigning absolute configuration than comparison at 589nm alone. Finally the anomalous optical rotatory dispersion of tyrosine in the near UV will be modeled in various protonation states via the Kramers-Kronig transformation of computed CD spectra and the results compared with experimental ORD.

COMPUTATIONAL METHODS

All data were computed with the Turbomole[30] quantum chemical software, version 5.7.1. The calculations were performed with B3-LYP[31] hybrid functional as implemented in the Turbomole code (note that this uses the VWN5 local correlation functional). Molecular geometries were optimized with the default doubly polarized valence triple zeta (TZVPP) basis set from the Turbomole library; all energies reported herein were computed with this basis. Response calculations were performed with the aug-cc-pVDZ set[64], commonly used for the calculation of chiroptical properties. Where noted, corresponding response calculations were carried out using the TZVPP basis for comparison.

All optimizations and response calculations were performed with the COnductor-like Screening MOdel (COSMO)[33] of solvation applied to the ground state. Solvent model parameters were configured using the cosmoprep program of the Turbomole package. The dielectric constant of the solvent was set to 78; the Number of geometrical Segments Per Atom (NSPA) was set to 162; all other solvent parameters were left at program default values. Default optimized COSMO atomic radii were used.

Initial geometrical parameters were set using the Molden[35] graphical interface program and its default bond distances and angles. First the alanine zwitterion structure was optimized and afterwards used as a template to generate conformations of the other amino acids that were themselves optimized. Cationic, anionic and neutral amino acid structures were derived from their corresponding optimized zwitterionic structures by adding, removing or changing the location of hydrogen atom(s) as appropriate followed by re-optimization.

All structures were confirmed local minima having no imaginary vibrational frequencies as calculated with the NumForce program. This numerical method of frequency computation

was used since the analytical frequency module of the software was incompatible with the COSMO solvation method. In this paper the energy, "D" of a particular conformer is defined as the sum of the electronic energy, the solvation energy from the COSMO model and the zero point energy calculated by NumForce. "ΔD" is defined as the energy of a particular conformer relative to the lowest energy conformation. Boltzmann factors were calculated based on this relative energy at the temperature of 293K.

The Turbomole code presently does not support the use of Gauge Including Atomic Orbitals (GIAOs, also referred to as London Atomic Orbitals), so strictly speaking all calculated optical rotations are somewhat dependent on the gauge origin. Our origin is defined as the center of mass in each molecule. Such gauge origin dependence is known to diminish as the basis set size increases, and from a practical standpoint reasonably reliable results for small molecules are obtained by using the large augmented basis sets that are always needed to calculate reliable optical rotations.[3, 12] Data in the literature support this conclusion[28],[37]

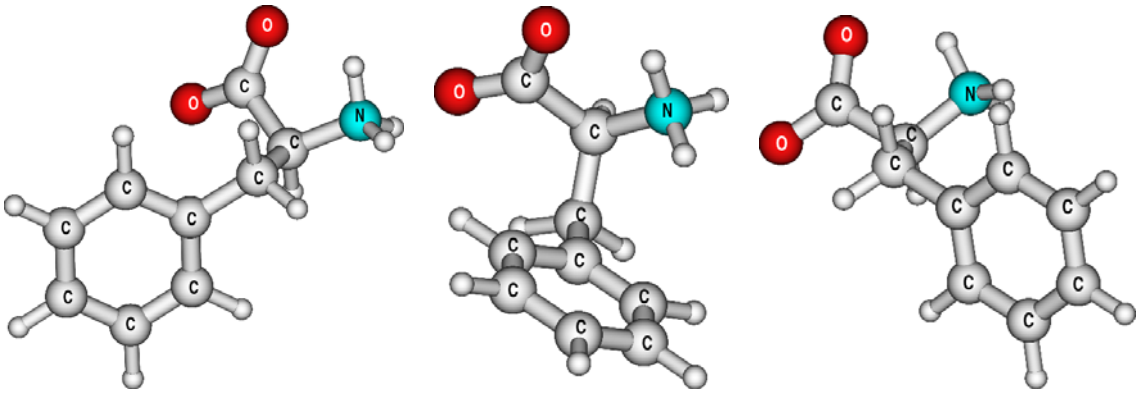
Except where otherwise noted, specific rotations were calculated at the wavelength of the sodium D line (589.3nm). All specific rotations are reported in units of deg·cm^3/(g·dm). Optical Rotatory Dispersion (ORD) curves in the transparent region were computed similarly using a direct linear response method and plotted with values calculated at 10nm intervals. Computed optical rotatory dispersion ORD curves in the absorbing region were obtained via a Kramers-Kronig transformation of computed CD spectra using a numerical integration scheme recently recommended by Polavarapu.[65] Experimental ORD plots were scanned from their respective graphics in the literature, digitized using the WinDIG program,[38] and plotted alongside the calculated curves.

RESULTS AND DISCUSSION

Structures of the aromatic amino acids

Before the chiroptical properties of the aromatic amino acids can be modeled, it must first be acknowledged that these molecules possess a conformational freedom in solution at room temperature. The mole fractions of each conformer may be calculated based on the energy of each minimum using the Boltzmann equation. As such, the specific rotation measured from a solution of molecules is actually caused by a number of conformations, and the response observed is a weighted average of the effects of each conformer.

Free amino acids in solution can be found as three primary rotamers. These will be referred to herein as "g", "t" and "h", per the naming convention used by Martin et al.[51] The $C_{carboxyl}C\ C\ C$ dihedrals of the rotamers are approximately +60, 180 and -60 degrees, respectively. These three conformations are depicted in Figure 5-1, using the phenylalanine zwitterion as an example. In the g rotamer the aromatic ring is *gauche* to the carboxylate chromophore. In the t rotamer the phenyl group is *trans* to the acid functional group. In the h rotamer the aromatic ring is found adjacent to both the amino and acidic functional group, in a configuration that can be considered *hindered* from a steric viewpoint.

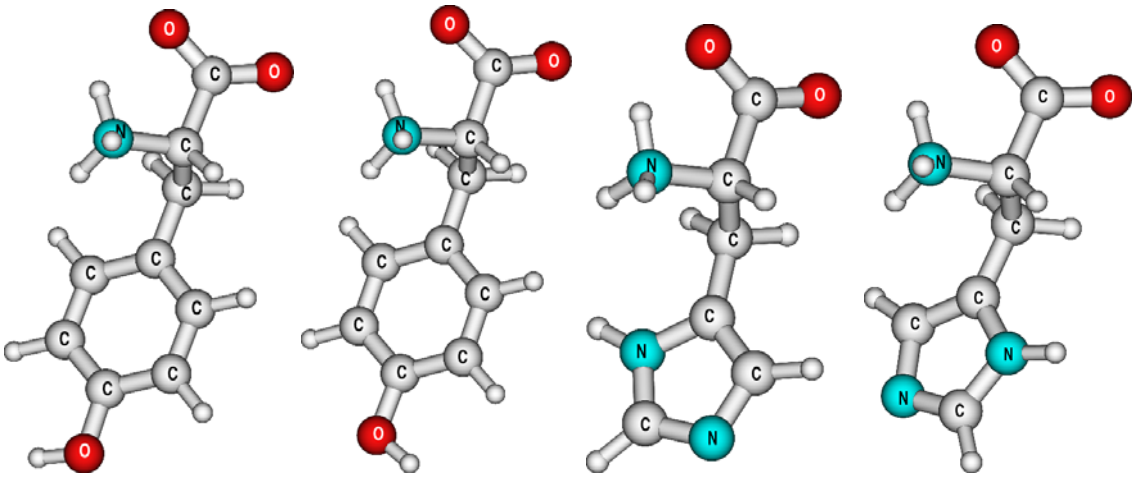


g h t

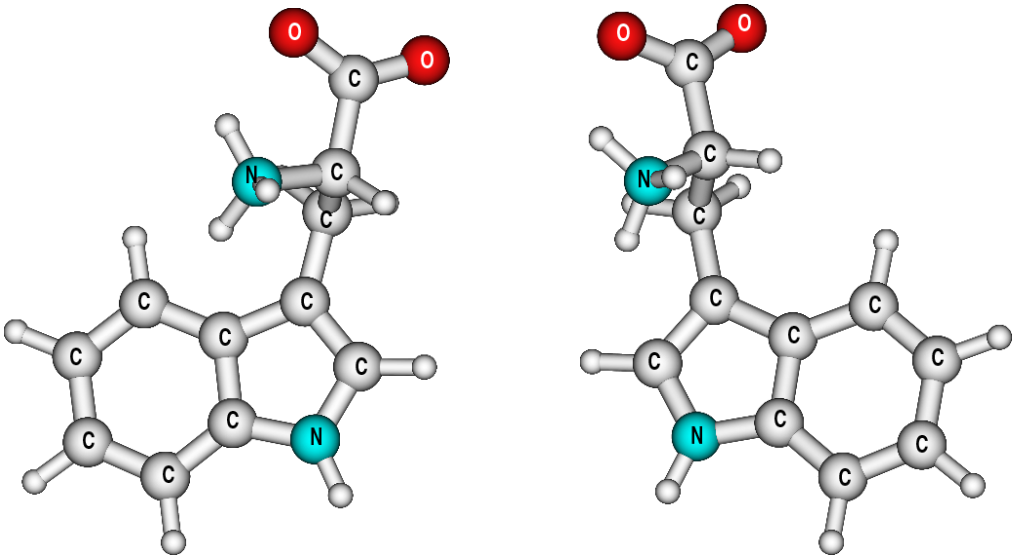
Figure 5-1: Optimized rotamers of the phenylalanine zwitterion.

Each of the phenylalanine rotamers converged to an optimized geometry with the aromatic ring approximately perpendicular to the C_α-C_β bond. No corresponding local minima with the aromatic ring parallel to this bond could be found. Therefore for phenylalanine only three conformations were considered.

For tyrosine, histidine and tryptophan more conformers are possible. These "sub-rotamers" are depicted in Figure 5-2. The tyrosine ring contains a hydroxyl functional group which can align itself in either of two directions, parallel to the ring. Histidine and tryptophan have asymmetric aromatic rings that can be found in two different configurations approximately perpendicular to the C_α-C_β bond. Tryptophan was unique in that additional conformations were found where the indole ring could converge to geometries nearly parallel to the C_α-C_β bond. However these conformers tended to be significantly higher in energy than the corresponding perpendicular conformations and were not considered further because of their negligible Boltzmann population. Therefore for each of these three amino acids two distinct sub-rotamers were modeled for a total of six possible conformers of each in the zwitterionic form.



tyrz – t_I tyrz – t_{II} hisz – t_I hisz – t_{II}

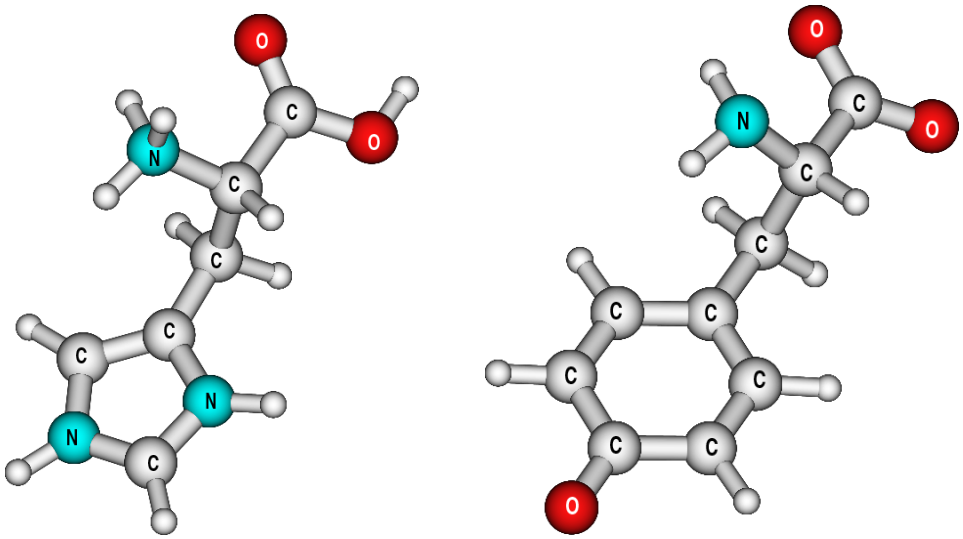


trpz – t_I trpz – t_{II}

Figure 5-2: Selected optimized examples of the sub-rotamers of the tyrosine, histidine and tryptophan zwitterions.

In addition to investigating the optical activity of these compounds near neutral pH, it is worth while to consider them at low pH, as a sizeable amount of data are available on these molecules measured in strongly acidic conditions. Consequently the specific rotation of the protonated forms of these molecules also has been considered. Protonation of the COO^-

functional group takes place preferentially on the oxygen more distant from the NH_3^+ group; this proton has a strong tendency to locate itself between the oxygen atoms. While other protonation locations are possible, an earlier study has indicated that they are high enough in energy relative to the ground state that these isomers are not significantly populated at room temperature.[58] Histidine is unique in this set of molecules in that it becomes doubly protonated at low pH. But since this second protonation can only occur at one position on the ring nitrogen atom, this does not lead to additional conformations, either. Therefore, just as with the zwitterions, six cationic (protonated) structures have been modeled for tyrosine, histidine and tryptophan and three structures were modeled for phenylalanine. An example of a doubly protonated molecule of histidine is depicted on the left side of Figure 5-3.



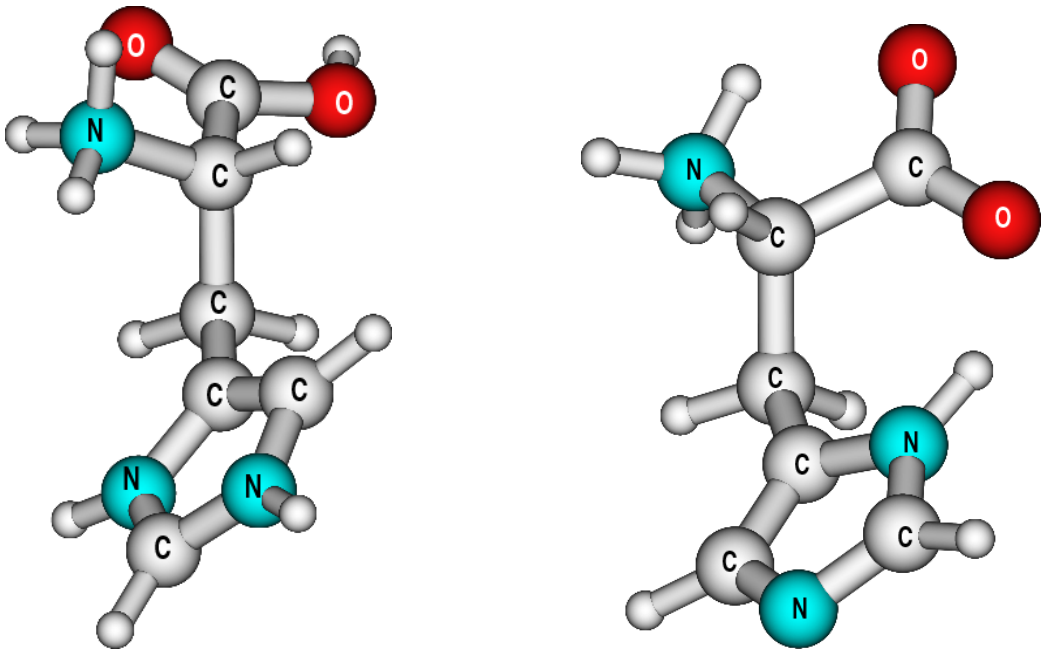
$his^{++}t_I$ $tyr^{--}t_I$

Figure 5-3: Select conformers of doubly protonated histidine and doubly deprotonated tyrosine

Among the amino acids studied, tyrosine is unique in that it can become doubly deprotonated at high pH. Encouraged by the availability of experimental data on the optical activity of tyrosine at high pH (and the limited availability of experimental data for this molecule

near neutral pH) the specific rotation of dianionic (doubly deprotonated) tyrosine was also modeled. Compared to the zwitterion the conformational space of the tyrosine dianion simplifies as the phenolic hydrogen is removed, but becomes more complex as any one of the three amino protons can be removed. All nine possible resulting conformers of the tyrosine dianion were considered, but the configuration depicted on the right side of Figure 5-3 was by far the lowest in energy and therefore dominates the analysis of the optical rotation. In this conformation the aromatic ring is trans to the carboxylate group and the lone pair of the amino nitrogen is directed away from this ring.

For the sake of brevity not every optimized calculated structure has been depicted in this work. There are however two configurations not already shown that are worth noting. For differing reasons, two conformers of histidine converged to structures where the aromatic ring was more parallel than perpendicular to the C_α-C_β bond. Their structures are shown in Figure 5-4.



$his^{++}t_{II}$ $hiszg_{I}$

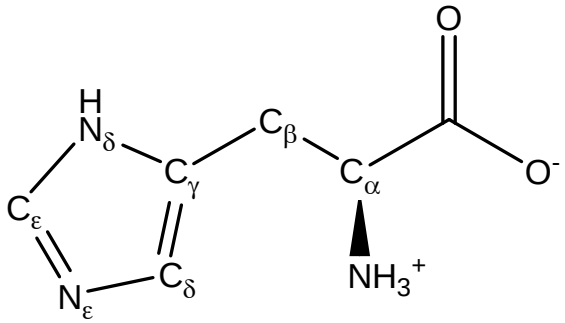


Figure 5-4: Top - The "parallel" conformers of histidine (left: dianion, right: zwitterion). The length of the hydrogen bond between the ring and carboxylate in the zwitterionic structure is 1.89 Angstrom. Bottom – atom labeling for histidine.

The misfit histidine dication conformer seems to result from repulsion between the $NH3^{+}$ group and the imidazole ring. However ordinary steric repulsion alone is not sufficient to explain this geometry, since the analogous histidine zwitterion did optimize to a perpendicular conformation. The steric interaction between the previously mentioned groups remains virtually the same in the zwitterionic and dicationic forms of histidine. The charge on these groups is what changes upon protonation. In the dicationic form of this amino acid the amino group and

the imidazole ring bear a significantly higher positive charge than they do in the zwitterionic form. An increased repulsion between positively charged functional groups could force the aromatic ring of the histidine dication away from its otherwise preferred perpendicular conformation into a more parallel one. The effect of this Coulomb repulsion on the geometry of the histidine molecule is shown in Table 5-1.

Table 5-1: Select functional group charges and geometric data of "trans" rotamers of histidine. Charges reported represent the sum of the computed Mulliken charges for each moiety.

molecular charge	amine group charge	imidazole ring charge	amino N - δ N distance (angstroms)	Cα-Cβ-Cγ-Cδ dihedral angle (degrees)
-	-0.11	-0.16	3.341	-90
z	0.29	0.15	3.539	-77
+	0.69	0.88	3.623	-72
++	0.76	0.88	3.857	-52

Coulombic repulsion is directly proportional to the product of the charges of two bodies and inversely proportional to their distance from one another. Based on the charge analysis the Coulomb repulsion should be the lowest in the anionic state and become progressively higher as the molecule is protonated. This repulsion may be relieved by the two groups moving farther away from each other, which happens in this case. The distance between the amino nitrogen and the aromatic ring atom closest to it, the nitrogen at the δ position, progressively increases as the charge repulsion increases. As this part of the imidazole ring is repelled, the aromatic ring rotates, maintaining its rigid, planar form, and the dihedral angle about the C_β-C_γ bond deviates from its ideal 90 degree configuration. This deviation grows progressively larger as the columbic repulsion increases, and at the doubly-cationic protonation state the repulsion large enough that the aromatic ring is pushed beyond the 60 degree maximum that distinguishes "perpendicular" from "parallel" configurations, giving rise to the parallel configuration seen on the left of Figure 5-3.

The imidazole ring in the zwitterionic histidine conformation depicted on the *right* side of Figure 5-4 has also drifted away from the otherwise preferred perpendicular configuration. However here it is not repulsion but attraction between functional groups that causes this deviation. In this structure the ring has rotated to bring the NH group into close proximity to the COO^- group, allowing the two groups to hydrogen-bond. The formation of this stabilizing intramolecular interaction causes the aromatic ring to deviate toward a parallel configuration in this instance.

This hydrogen bonding has been a potential source of error in the past, as the computational technique employed significantly overestimated the strength of intramolecular hydrogen bonding in zwitterions.[58] In those calculations the aug-cc-pVDZ basis was used, which includes diffuse functions. Diffuse functions have been regarded as important in modeling hydrogen bonding. While this basis was adequate for the geometry optimization and energy calculations of the small amino acids modeled in our earlier work, it proved problematic when dealing with the larger molecules in this current work. Here the TZVPP basis, which does not include diffuse functions, was used for geometry optimizations and energy calculations. It was chosen in part due to its established compatibility with the COSMO solvent model because the atomic radii were optimized for use with this basis which should therefore yield the most accurate solvation energies. The aug-cc-pVDZ basis was retained for the response calculations for which it is known to perform well. The use of different basis sets for optimizations and property calculations is well established in the field of computational chemistry, and efficiency regarding computing time by itself is enough to justify this practice. But there is another benefit from using a non-diffuse basis for these calculations.

Hydrogen bonding is an interaction between the orbitals of different atoms, be they on the same molecule (intramolecular) or between adjacent molecules (intermolecular). This bonding can take place over a significant distance, often two angstroms or more; this is why diffuse functions are needed to adequately describe it. In a system of hydrogen bonding molecules, such as an aqueous solution of an amino acid, an equilibrium exists between the tendency for a solute molecule to hydrogen bond with itself and its tendency to bond with solvent. When a continuum solvent model such as COSMO is used in lieu of explicit molecules to simulate solvation, this competition between inter- and intra-molecular hydrogen bonding becomes biased toward the latter. With no explicit water molecules for the amino acid solute to hydrogen bond to, it has an over-tendency to hydrogen bond with itself. Therefore, the amino acid conformations which include intramolecular hydrogen bonds become more favored energetically, and their computed Boltzmann populations become too high. This in turn can adversely alter the computed specific rotation of the solution, which depends in part on the computed populations of those conformers. We have investigated this issue in detail in a previous paper and concluded that the aug-cc-pVDZ + COSMO level indeed tends to overemphasize the internal hydrogen bonding in the zwitterionic amino acids.[58] Using a less-diffuse basis hinders hydrogen bonding. Since in the absence of explicit solvent molecules this (intramolecular) hydrogen bonding is too great, these two opposing errors partially compensate for each other giving results that are somewhat closer to experiment. This first became apparent from calculations on the smallest amino acid, glycine. The previously mentioned earlier work with the COSMO/B3LYP/aug-cc-pVDZ method indicated that the zwitterionic form of glycine was favored over the neutral form by 4.6kJ/mol. Here with the TZVPP basis it is calculated at 7.5kJ/mol. Although this is only marginally closer to the experimental value of 30.4kJ/mol,[23]

it represents a significant relative change of the energy difference between the structures. We note in passing that switching to a basis that is not augmented with diffuse functions more than cuts in half the amount of "outlying charge" from the glycine zwitterion. This is the amount of electron density that escapes the COSMO model's outermost cavity and is a potential source of error.[22] Although it is still not a perfect solvent model, the COSMO/B3LYP/TZVPP method has adequately modeled the fact that the zwitterionic form of the amino acids is more stable than the neutral form in aqueous solution.

The basis set effects on the computationally derived mole fractions can be seen in Table 5-2. Here all geometries were computed with the TZVPP basis. The only variable that contributed to the differing Boltzmann populations between the two basis sets employed is the COSMO-corrected single point energy.

Table 5-2: Computed room temperature populations of amino acid rotamers with differing basis sets compared to experimental data. Computed populations for tyrosine, tryptophan and histidine rotamers represent sums of the populations of two sub-rotamers per rotamer.

	Computed Population TZVPP	Computed Population aug-cc-pVDZ	Experimentally Derived Population
	Phenylalanine Zwitterion		
g	0.03	0.03	0.24[a], 0.24[b]
h	0.21	0.19	0.27[a], 0.28[b]
t	0.76	0.78	0.50[a], 0.48[b]
	Phenylalanine Cation		
g^+	0.26	0.24	0.24[a], 0.28[b], 0.36[c], 0.32[f]
h^+	0.19	0.20	0.27[a], 0.26[b], 0.27[c], 0.26[f]
t^+	0.55	0.56	0.50[a], 0.46[b], 0.37[c], 0.42[f]
	Tyrosine Cation		
g^+	0.25	0.23	0.20[e]
h^+	0.20	0.20	0.40[e]
t^+	0.55	0.57	0.40[e]
	Tyrosine Dianion		
g^{--}	0.03	0.04	0.16[e]
h^{--}	0.03	0.02	0.38[e]
t^{--}	0.95	0.94	0.46[e]
	Tryptophan Zwitterion		
g	0.01	0.01	0.32[d], 0.26[f]
h	0.17	0.15	0.15[d], 0.23[f]
t	0.82	0.84	0.53[d], 0.51[f]
	Histidine Zwitterion		
g	0.44	0.63	0.18[g]
h	0.44	0.31	0.31[g]
t	0.11	0.07	0.51[g]
	Histidine Dication		
g^{++}	0.55	0.52	0.34[g]
h^{++}	0.11	0.13	0.28[g]
t^{++}	0.34	0.35	0.37[g]

a- Kainosho and Ajisaka[66] b- Fujiwara et al.[54] c- Hansen et al.[67]
d- Dezube et al.[68] e- Juy et al.[69] f- Reddy et al.[70] g- Merrett et al.[71]

For the phenylalanine, tyrosine and tryptophan molecules where intramolecular hydrogen bonding is not an issue the computed populations are nearly the same with the two basis sets. The agreement between these computed populations and those derived from experiment can be described as qualitatively correct. While the numerical agreement between experiment and theory is not perfect, both agree that for these three molecules the stericly favored t rotamers are more populated than their g and h counterparts at room temperature.

Consistent with our assumptions about controlling the intramolecular hydrogen bonding with the basis set, the results for histidine are different. Unlike the other aromatic amino acids, histidine shows a tendency to form intramolecular hydrogen bonds; the possibility for bonding exists between the carboxylate anion and the NH group at the δ position of the aromatic ring. Such an interaction is favored in the g conformer depicted on the right side of figure 5-4, where the two participating functional groups have been able to move into close proximity unhindered by the $NH3^+$ group. Here the populations computed with the differing basis sets differ. While the population computed with the TZVPP basis overestimates the experimentally derived population of this hydrogen bonded conformer (at the expense of the t conformers in which the two groups are most distant and cannot bond), the diffuse aug-cc-pVDZ basis overestimates it even more. This is consistent with the theory that both methods overestimate the extent of intramolecular hydrogen bonding, but the less-diffuse basis errors to a lesser extent. Therefore for the remainder of this work all Boltzmann populations reported will be computed from energies calculated with the TZVPP basis.

Computation of the sodium D-line specific rotation

Basis sets were also examined as they applied to the computed specific rotation. For nearly every conformation of every molecule investigated, the computed sign of the specific

rotation was the same with the aug-cc-pVDZ and the TZVPP basis sets. The exceptions are the conformations that have specific rotations of too small a magnitude to be accurately calculated by this method. This should not be read to conclude that diffuse functions are not important for modeling specific rotation in general; often times they are. But for this particular set of molecules the difference in chiroptical response between the two basis sets was not very significant.

The experimental and computed specific rotations of the four aromatic amino acids in select ionization states are shown in Table 5-3. At first glance, the agreement between theory and experiment appears somewhat disappointing, as the theory used herein yields a specific rotation that agrees in sign with experiment six out of nine times. But this is not unexpected when dealing with rotations that are so small in magnitude. Stephens and coworkers have shown the limitations of TDDFT for assigning absolute configuration of molecules with very small rotations.[4] When a small specific rotation value is the result of a delicate weighted average of several conformers having positive and negative optical rotations of large magnitude, then even a small error in computing the Boltzmann populations of these conformers could reverse the sign of the specific rotation. Furthermore, in cases where the measured specific rotation is exceedingly small, variations in laboratory conditions combined with the inherit limits of precision in the instrumentation can result in the sign of the experimentally measured rotation to itself be ambiguous, as was the case with the tryptophan cation discussed later in this paper. If it were the purpose of this work to assign configurations, using a shorter wavelength of light closer in energy to the lowest electronic excitation energies of the molecules in question would be advantageous because this would result in rotations much higher in magnitude and give a more

accurate assignment. Here the sodium D-line light of 589.3 nm has been selected because most of the experimental data in the literature has been obtained at this convenient wavelength.

Table 5-3: Computed and experimental data for the specific rotation of the aromatic amino acids in select ionization states. All specific rotation values are in (deg·cm^3)/(g·dm).

			aug-cc-pVDZ		TZVPP		
				Boltzmann		Boltzmann	experimental
conformer	ΔD (kJ/mol)	Boltzmann factor	specific rotation	factor x specific rotation	specific rotation	factor x specific rotation	specific rotation
			phenylalanine zwitterion				
t	0.0	0.76	6.7	5.1	18.2	13.8	
h	3.0	0.21	-150.9	-32.4	-120.6	-25.9	
g	7.6	0.03	162.4	5.0	142.8	4.4	
Average specific rotation				-22.3		-7.7	-35.1
			phenylalanine cation				
t^+	0.0	0.55	25.8	14.1	18.0	9.8	
g^+	1.8	0.26	221.8	58.1	210.4	55.1	
h^+	2.5	0.19	-69.1	-13.2	-66.7	-12.7	
Average specific rotation				59.0		52.2	-7.4[a]
			tyrosine zwitterion				
t_I	0.0	0.34	9.2	3.1	20.6	7.0	
t_{II}	0.2	0.31	-7.1	-2.2	4.1	1.3	
h_I	1.8	0.16	-132.9	-21.6	-104.5	-17.0	
h_{II}	3.3	0.09	-147.2	-12.5	-120.9	-10.3	
g_I	4.5	0.05	169.0	8.7	150.1	7.7	
g_{II}	4.7	0.05	167.7	8.0	145.3	6.9	
Average specific rotation				-16.5		-4.3	-10.0 to -12.3[b]
			tyrosine cation				
t_I^+	0.0	0.34	-12.2	-4.1	-15.0	-5.1	
t_{II}^+	1.1	0.21	-3.6	-0.8	-8.4	-1.8	
g_I^+	2.2	0.14	237.8	32.7	226.1	31.1	
h_I^+	2.4	0.12	-57.4	-7.0	-55.4	-6.8	
g_{II}^+	2.7	0.11	229.9	25.6	218.2	24.3	
h_{II}^+	3.6	0.08	-71.4	-5.4	-71.8	-5.5	
Average specific rotation				41.0		36.3	-10.6

			tyrosine dianion				
t_I^{--}	0.0	0.82	-36.3	-29.6	-25.5	-20.8	
t_{II}^{--}	5.8	0.07	-100.8	-7.3	-72.0	-5.2	
t_{III}^{--}	6.3	0.06	55.6	3.2	7.9	0.5	
h_I^{--}	9.4	0.02	-64.4	-1.0	-75.6	-1.2	
g_I^{--}	9.9	0.01	145.8	1.9	97.9	1.3	
g_{II}^{--}	9.9	0.01	101.8	1.3	72.1	0.9	
h_{II}^{--}	10.6	0.01	-60.3	-0.6	-25.7	-0.2	
g_{III}^{--}	14.4	<.01	116.6	0.2	68.6	0.1	
h_{III}^{--}	15.2	<.01	-71.8	-0.1	-38.0	-0.1	
Average specific rotation				-32.0		-24.8	-13.2
			histidine zwitterion				
g_I	0.0	0.44	115.5	51.1	108.3	47.9	
h_I	0.0	0.43	-196.6	-85.1	-147.3	-63.8	
t_I	3.4	0.11	-51.4	-5.4	-54.3	-5.7	
t_{II}	9.2	0.02	-72.6	-0.7	-42.1	0.0	
h_{II}	9.7	0.02	-67.7	-0.5	-63.4	-0.5	
g_{II}	12.4	<.01	213.4	0.5	196.9	0.5	
Average specific rotation				-40.2		-22.0	-38.95
			histidine dication				
g_I^{++}	0.0	0.28	120.3	33.9	117.5	33.1	
g_{II}^{++}	0.1	0.27	153.7	41.2	143.4	38.5	
t_I^{++}	0.3	0.25	22.0	5.5	8.4	2.1	
t_{II}^{++}	2.8	0.09	-73.8	-6.5	-76.9	-6.8	
h_I^{++}	3.0	0.08	-74.9	-6.1	-67.1	-5.5	
h_{II}^{++}	5.2	0.03	1.1	0.0	-8.8	-0.3	
Average specific rotation				67.9		61.1	13.34
			tryptophan zwitterion				
t_I	0.0	0.76	-22.0	-16.8	-30.6	-21.7	
h_I	4.2	0.13	8.8	1.2	14.2	1.7	
t_{II}	6.5	0.05	101.8	5.0	123.6	5.6	
h_{II}	6.8	0.04	-318.5	-14.3	-294.5	-12.3	
g_I	10.4	0.01	260.1	2.5	226.3	2.0	
g_{II}	15.0	<.01	-46.3	-0.1	-48.2	-0.9	
Average specific rotation				-22.5		-30.4	-31.5

				tryptophan cation			
t_I^+	0.0	0.74	-42.7	-54.1	-82.6	-57.2	
h_I^+	5.3	0.08	89.6	7.2	84.0	6.3	
g_I^+	5.6	0.07	395.4	28.3	367.4	24.5	
g_{II}^+	7.0	0.04	-7.1	-0.3	-5.9	-0.2	
t_{II}^+	7.2	0.04	67.5	2.4	73.1	2.5	
h_{II}^+	7.5	0.03	-209.3	-6.5	-215.0	-6.3	
Average specific rotation				-22.9		-32.4	2.4

Experimental data are from the *Merck Index*[42] except where otherwise noted.
a – experimental data for phenylalanine cation from the work of Greenstein and Winitz.[72]
b – experimental data for tyrosine zwitterion from world wide web.[73]

Of the possible sources of error, inaccuracy in the computed energy likely plays a large role. This energy is used to calculate the Boltzmann population of each conformer of a particular molecule. These factors are then multiplied by the specific rotation of each conformer and the products are added to give the weighted average specific rotation that should be comparable to the experimentally observable value. As can be seen in Table 5-4, the specific rotations of some of these conformations are an order of magnitude larger than the average specific rotation that should be modeled. Since this relatively small weighted mean is produced from a number of conformers of relatively large specific rotations of differing signs, one can see how even a small error in a Boltzmann weighting factor can cause the computed average specific rotation to change sign.

Another important conclusion that the data lead to is that the method performs about equally regardless of the protonation state of the molecule being modeled. In an earlier work we have shown that TDDFT performed well in modeling the specific rotation of zwitterionic, cationic and anionic amino acids in solution.[58] Here the specific rotation of an amino acid in its dicationic form, histidine, as well as an amino acid in its dianionic form, tyrosine, have been modeled with an equal margin of error. This is particularly gratifying in the case of the tyrosine dianion, due to the added difficulty of modeling electronic transitions in molecules with the more

diffuse electronic structures found in anions. Different conclusions have been reached regarding whether or not TDDFT is adequate for modeling the optical rotation of similar anionic systems.[28, 58] For this molecule the theory employed yields an average specific rotation of -32.0, which compares reasonably well with the experimental value of -13.2. The general overestimation of the magnitude of rotation is considered typical of TDDFT due in part to known underestimation of excitation energies typically found for many organic molecules.[45] However, the correct modeling of the specific rotation of the molecules in this study depends not only on correctly modeling the excitation energies and rotatory strengths, but also on computing the correct relative energies of the conformers and, in addition a correct modeling of the solvation of the molecules being studied.

Contributions from different chromophores to the specific rotation

The primary purpose of this work has not been merely to compare computed specific rotations of the aromatic amino acids with experiment, but also to examine the interplay of the two different types of chromophores that are unique to the aromatics. The specific rotation of amino acids is generally thought to result from an $n \rightarrow \pi^*$ transition localized on the carboxylate chromophore. In aromatic compounds there also exists the possibility for a $\pi \rightarrow \pi^*$ transition involving the bonding and anti-bonding orbitals associated with the aromatic ring(s).

The effect of the carboxylate chromophore common to all the amino acids can be seen in the computed specific rotation of the various rotamers. The primary rotamers differ from one another by rotation about the C_α-C_β bond. The functional group that is responsible for perturbing the symmetry about the carboxylate group, in this case an aromatic ring, may be found in three orientations illustrated in Figure 5-1. For the g rotamers where the perturbing functional group is gauche to the carboxylate group on its stericly unhindered side the specific rotations are

predominately positive. Conversely for the h rotamers when the perturbing group is on the opposite gauche side of the carboxylate group, the side that is stericly hindered by the ammonium group, the specific rotation is predominately negative. For the t rotamers where the perturbing group is trans to the carboxylate chromophore and most distant from it, the specific rotation is on the whole neither predominately negative nor positive when all the t conformations of all the molecules studied are considered. However the specific rotations of the t rotamers are generally smaller in magnitude than their g and h counterparts. This is consistent with the idea that when the perturbing group is most distant from the perturbed chromophore, the perturbation of the symmetry of the chromophore is less significant, resulting in a smaller specific rotation.

For tyrosine, histidine and tryptophan two subrotamers per rotamer can also be considered. These differ from one another by a rotation of approximately 180 degrees about the C_β-C_γ bond and are designated simply as "I" and "II", with the former being lower in energy. No differing subrotamers of phenylalanine were found due to the symmetry of the phenyl ring.

For tryptophan in particular, a consistent difference in the specific rotation of pairs of subrotamers is observed. In all cases investigated, 3 pair of zwitterionic geometries and 3 pair of cationic geometries, each pair of subrotamers were found to have specific rotations of opposite sign. Here it is the C_β-C_γ dihedral angle and not the C_α-C_β angle that is most important for determining the specific rotation of the conformer. In other words for tryptophan the orientation of the rest of the molecule with respect to the carboxylate chromophore has less of an effect on the specific rotation that the orientation of the rest of the molecule with respect to the indole chromophore. All conformations in which the rest of the tryptophan molecule is oriented on one side of the plane of the indole ring have a positive specific rotation and all those in which the perturbation is on the opposite face have a negative specific rotation. This is the chiroptical

response to be expected from an indole chromophore, which itself has C_s symmetry that can be perturbed to give either a positive or negative response depending on which side of the symmetry plane the perturbing atoms are oriented, in agreement with the well known concept of sector rules.[15] The conclusion regarding the specific rotation of tryptophan is that the aromatic chromophore is primarily responsible for the observed specific rotation and the carboxylate chromophore has a secondary effect. This is consistent with experimental[74] and theoretical[63] data which show significant CD excitations associated with the indole chromophore at energies lower than 250 nm. CD excitations that are largest in magnitude and closest to the energy where specific rotation is measured should have the greatest effect on its sign and magnitude, which is easily rationalized by the form of the denominator in the sum-over-states expression of optical rotation. The equation reads:

$$[\alpha]_\omega = k \cdot \omega^2 \cdot \sum_j \frac{R_j}{\omega_j^2 - \omega^2} \tag{5-1}$$

where k is a constant, ω is the frequency, ω_j an excitation frequency and R_j the corresponding rotatory strength for excitation number j. The sum formally runs over the complete set of excitations of the molecule.

Histidine like tryptophan contains an aromatic ring chromophore. This imidazole ring would belong to the C_{2v} point group were it not for the fact that the selective protonation of the nitrogen at the δ position and the attachment of the rest of the molecule to the ring at a position that does not straddle the two nitrogen atoms reduce the site symmetry to C_s, which is then further perturbed to a chiral point group by the rest of the molecule. As with tryptophan, rotation about the C_β-C_γ bond can have a significant effect on the specific rotation. But unlike with tryptophan the specific rotations of the pairs of subrotamers for histidine do not always have

opposing signs. The specific rotation depends less on the C_β-C_γ angle than on the C_α-C_β angle. The arguably higher symmetry of the imidazole ring compared to the indole ring might partially explain this difference. But an explanation for which there is observable physical evidence is that the energy for the lowest electronic transition for the imidazole ring is higher than that of the indole, as the first CD excitation for histidine is not observable until around 220 nm.[74] The first excitation energy of an imidazole ring is expected to be higher in energy than that of an indole ring since the imidazole has fewer π electrons and fewer atoms over which they are conjugated. Assuming similar magnitudes of the rotatory strengths, since the first $\pi \rightarrow \pi^*$ CD excitation occurs at a higher energy in histidine than in tryptophan, the histidine aromatic chromophore should have less of a dominating effect on the measured specific rotation, according to the equation (1). Our computations indicate that in histidine both the aromatic and the carboxylate chromophores have significant effects on this optical activity.

For the tyrosine cation, different pairs of subrotamers do not yield significantly different specific rotations. The pairs of subrotamers of tyrosine only differ in the position of the phenolic hydrogen, which always orients itself in one of two positions in the plane of the aromatic ring. For the tyrosine dianion, this proton is removed and rotation of 180 degrees about the C_β-C_γ bond yields degenerate structures. The dihedral angle about the C_α-C_β bond appears the most important for the tyrosine conformations studied, with the g rotamers always having a positive specific rotation, the h rotamers having a negative rotation and the t rotamers showing neither tendency.

For phenylalanine, the effect of the phenyl chromophore on specific rotation cannot be modeled with the method used in this work. The six-membered aromatic ring in phenylalanine when unperturbed has D_{6h} symmetry, meaning that the lowest electronic excitation is symmetry

forbidden. Experimentally vibronic coupling allows this weak transition to be observed in both ORD[75] and CD,[63] where the vibronic fine structure is clearly visible. But since vibronic coupling is not included in the method for computing specific rotation, the effects of the electronic transition in the aromatic chromophore cannot be analyzed. Such effects can be modeled, in principle, in a relatively straightforward manner in the CD,[76] but would require excessive computational resources.

Computation of the optical rotatory dispersion in the near UV

In order to make an assignment of absolute configuration of a molecule, measurement and modeling of the specific rotation at 589.3 nm may be a convenient choice, but it is not the most accurate method. To make a comparison at higher frequencies where specific rotation is larger would likely add improvement, particularly in situations where differing conformations of a molecule have specific rotations of opposing signs at long wavelengths but agree in sign at shorter wavelengths closer to an excitation. For example the computational method used thus far modeled the wrong sign for the specific rotation of the phenylalanine and tyrosine cations, but if the wavelength of 300 nm is used, which is close to but still lower than the first excitation energy of these molecules, then the experiment and the theory yield rotations that agree in sign. At this wavelength the specific rotation of protonated phenylalanine and tyrosine were measured as +91 and +132 degrees, and calculated to be +658 and +866, respectively. The overestimation of the magnitude of the rotation is to some extent the result of using a method that does not include damping for excited states[77]; the exaggeration becomes more severe as the wavelength used is closer to an excitation energy, where the computed specific rotation would yield a singularity.

Comparing measured and modeled specific rotation at a single frequency can sometimes be useful in assigning absolute configuration. But to make this comparison over a large range of

wavelengths, by computing an ORD curve and comparing it to experiment, would be the preferred method.[78] Transparent-region ORD curves are available in the literature for all of the molecules for which our theory does not model the correct sign of the sodium line specific rotation.[43] Of these the one for tryptophan is the most interesting due to its distinctive trough feature near 340nm, so it is chosen here as a representative example.

The comparison of the ORD of tryptophan between 300 and 600 nm is depicted in Figure 5-5. The overall agreement between the two is quite good; the characteristic shape of the curve has been faithfully reproduced. Interestingly, *these* experimental data appear to agree with our calculations which indicate that the tryptophan cation has a rather large negative specific rotation near 589nm. In fact, we found a number of sources in the literature that indicated that solutions of tryptophan in hydrochloric acid are dextrorotatory[42, 72] and some that indicated that it is levorotatory[43, 79] at this wavelength. This serves to further emphasize the point that comparing computed and measured specific rotation at a single wavelength is not the preferred method for assigning absolute configuration. Matching the shape of an ORD curve, as has been done in Figure 5-5, provides a far more reliable means for matching chiroptical response to structure.

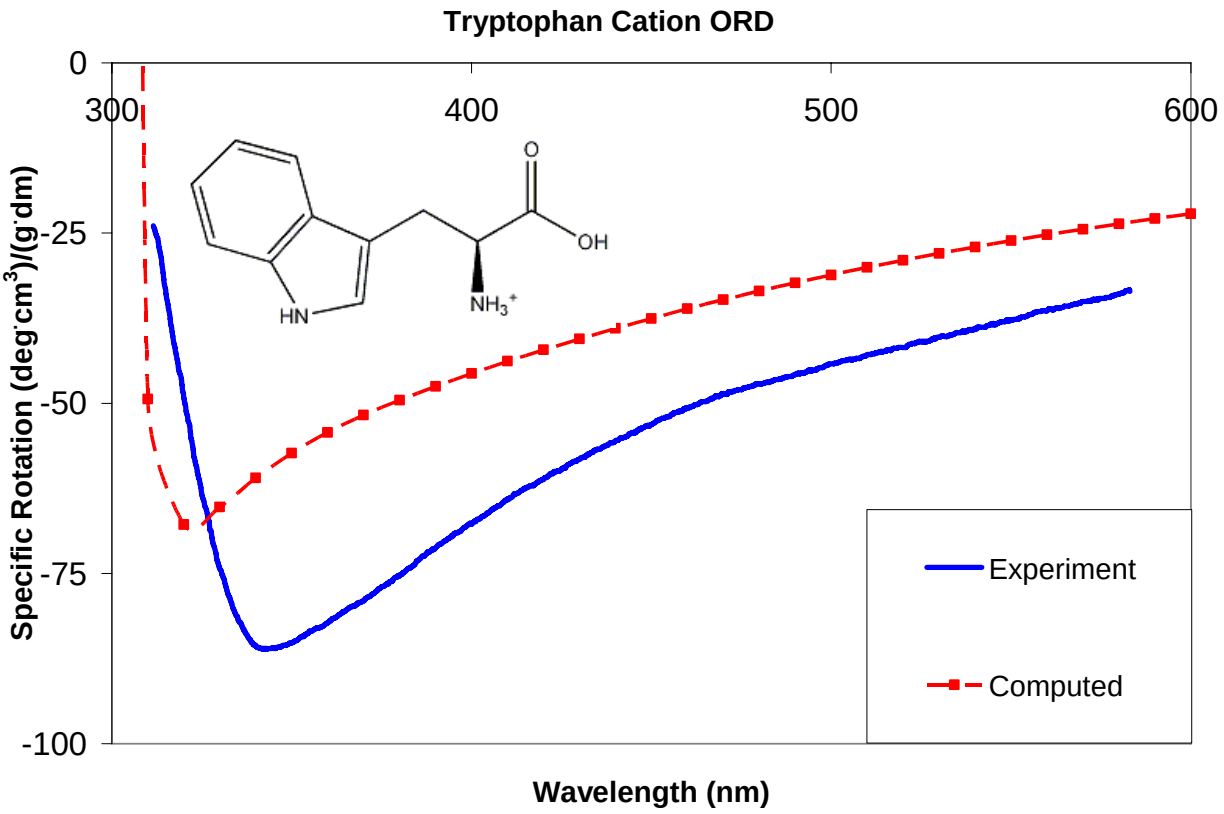


Figure 5-5: Computed and experimentally measured ORD of protonated tryptophan. Computed data points represent Boltzmann averages of optical rotations from six conformers. Experiential data are from the work of Djerassi.[43]

The ORD that was measured and modeled in Figure 5-5 covers only wavelengths longer than the first excitation wavelength of the molecule. Based on their CD spectra both tyrosine and tryptophan are expected to have anomalous ORD features in the 225 – 300 nm region that should also be readily measurable. In fact near-UV region dispersion curves of tyrosine at a variety of pH conditions are available in the literature.[60, 80] As these conditions correspond to some of the protonation states of tyrosine that were already modeled here, we have considered the ORD for this molecule as well.

The ORD of tryptophan in Figure 5-5 was successfully modeled by carrying out multiple linear response calculations. This covered only a region of wavelengths to which the molecule is

transparent, so the lifetime broadening of the excited states was not of major concern. But when one calculates anomolous ORD with a program that does not include any damping in the linear response calculations, such as the one used in this work, singularities occur at all of the excitation wavelengths. Graphics depicting these effects on computed ORD curves can be found in the literature.[11, 77]

At present, a program to calculate by direct linear response resonant ORD with hybrid functionals and the COSMO solvent model is not available to us. The method used here to model resonant ORD is as follows: First the lowest 100 excitations of a CD spectrum for each conformation of each molecule studied were calculated. This is not meant to imply that 100 excitations are either necessary or sufficient to model the ORD in the region of interest; in fact we later found that truncating the series to five excitations did not significantly change the shape of the resultant ORD in the resonance region. The number 100 was chosen as an arbitrary cutoff to keep the computational time required to less than one processor week per calculation. There is no specific rule regarding how many excitations one should calculate for this method, and some truncation error is inevitable no matter where the cutoff is set.[65],[81] Next the CD intensity was simulated using an empirical Lorentzian broadening with half-width at a half peak height of 0.19eV Then these CD spectra were transformed into ORD curves via a numerical Kramers-Kronig transformation as described by Polavarapu.[65] Finally the resulting ORD curves of the individual conformers were Boltzmann averaged to produce the computed dispersion curve that is reported for each ionization state of tyrosine. The resulting plots are compared with experiment in Figure 5-6. This method was recently used to successfully model the anomolous ORD of protonated proline through its first Cotton effect.[81]

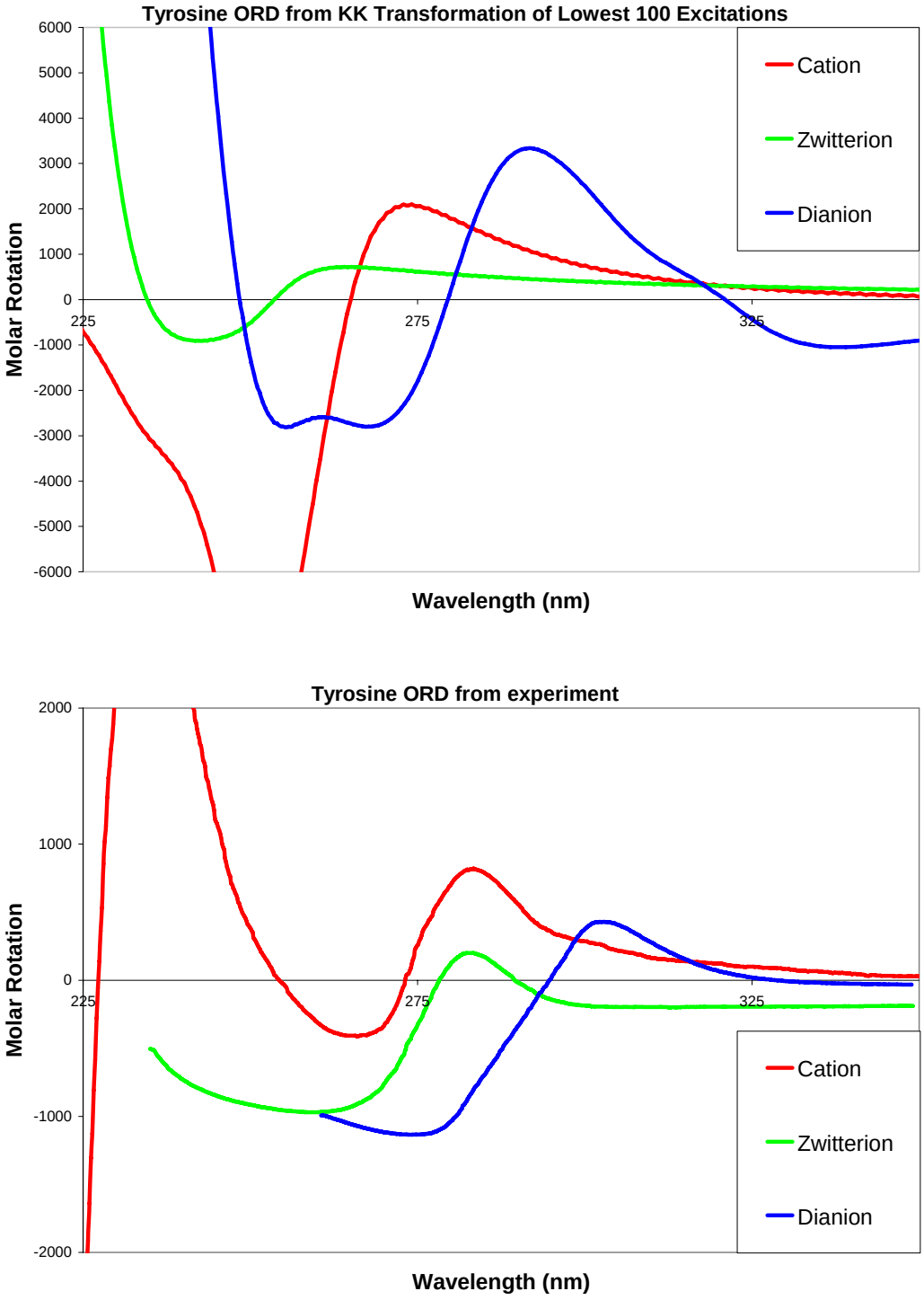


Figure 5-6: Computed and experimental ORD curves for tyrosine in various states of protonation. Experimental data for the cationic form are from Iizuka and Yang.[60] Experimental data for the zwitterionic and dianionic forms are from Hooker and Tanford.[80]

As is shown in Figure 5-6, for all of the protonation states the theory correctly models the sign of the first Cotton effect. Furthermore, it correctly reproduces the experimentally observed trend that the first Cotton effect of the zwitterion occurs at a higher energy than that of the cation or dianion. The intensity of the effects is not correct, but this is largely affected by the empirical broadening factor used. For the sake of simplicity a constant broadening factor was used, while in reality the broadening may be expected to increase with energy.

A limitation of this method is seen in the lower energy region away from the excitation. In this long wavelength area the ORD becomes increasingly dependent on the higher lying excitations. Since only a finite number of excitations can be included in the KK transformation, the truncation error becomes more of an issue in this low energy area of the ORD where such excitations play a larger rule. This is why KK transformation of CD spectra is not the preferred method for computing ORD in the transparent region. As Polavarapu has stated earlier, direct linear response should instead be used to model ORD in the long wavelength region,[82] as was done in this work for the calculations at 589.3nm and for the transparent region ORD of the tryptophan cation. KK transformations perform best for wavelengths that are in the vicinity of well separated electronic excitations, where, per the sum-over-states equation (1), the dispersion is dominated by these individual excitations. The purpose of using the Kramers-Kronig transformation here is to investigate the ORD in such an anomolous region. A comparison of the ORD in the vicinity of the first Cotton effect would be sufficient to match the chiroptical response with the correct enantiomeric structure. Just as others have used the KK transformation to model the anomolous ORD for several neutral organic molecules,[65] the method has proven useful here for investigating the ORD of the various protonation states of tyrosine as well. It appears that the modeling of resonant ORD by means of the KK transformations of a truncated

CD spectrum well compliments linear response calculations of the ORD and specific rotation in the transparent region.

CONCLUSIONS

The specific rotation and optical rotatory dispersion of solutions of the aromatic amino acids in various protonation states have been modeled. The Boltzmann distribution of the low energy conformations of the molecules was in qualitative agreement with experimentally derived distributions. The agreement was poorest for histidine due to difficulties with modeling of its intramolecular hydrogen bonding. Correct modeling of the specific rotation of the molecules at 589nm proved a challenge since the rotations were relatively small in magnitude. However some insight was gained about the effects of the different aromatic chromophores on this specific rotation. Comparison of computed and measured ORD over a broad range of wavelengths is clearly a more reliable method for assigning absolute configuration than comparison of specific rotation only at 589nm. Kramers-Kronig transformation of computed CD spectra to ORD curves proved useful in modeling the sign and relative energy of the first Cotton effect for tyrosine in various protonation states.

Acknowledgment

We wish to thank Dr. Mykhalo Krykunov for his assistance with the Kramers-Kronig transformations. We also acknowledge the University at Buffalo Center for Computational Research (CCR) for maintenance of our computing resources. J.A. is grateful for financial support from the ACS Petroleum Research Fund and from the CAREER program of the National Science Foundation (CHE-0447321).

CHAPTER 6: THE CLOUGH-LUTZ-JIRGENSONS RULE

Kundrat, Matthew D.; Autschbach, Jochen. Computational modeling of the optical rotation of amino acids: A new look at an old rule for pH dependence of optical rotation. Journal of the American Chemical Society; **2008**; *130*(13) pp 4004-4414.

Kundrat, Matthew D.; Autschbach, Jochen. Specific rotation of amino acids: A new look at an old rule. Poster presented at 232nd ACS National Meeting, San Francisco, CA, Sept. 10-14, 2006, PHYS 449.

ABSTRACT

The molar rotation of a solution of a natural alpha amino acid is changed in the positive direction by addition of a strong acid. Three decades ago, an attempt to rationalize this old rule, named for Clough, Lutz and Jirgensons (CLJ), was made by assigning circular dichroism octants for overlapping carbonyl n to pi* transitions. Modern quantum chemical methods allow us to take a new look at this phenomenon. Time-dependent density functional theory was used to model the electronic structure and transitions responsible for CLJ. We show that sector rules originally developed for circular dichroism (CD) can be applied to the optical rotation in this case, but with some restrictions, and with great caution, due to the change of the overall charge of the acids upon protonation and the distortion of the COO^- chromophore in the zwitterions. We have prepared sector maps based on first-principles computations to study the correspondence between CD and optical rotation for zwitterionic and protonated L-amino acid chromophores. The CLJ effect is correctly obtained from the computations for all 12 amino acids studied in this work.

INTRODUCTION

Already in the early days of chemistry it was known that solutions of certain compounds (now designated as chiral), many biological in origin, rotate a plane of polarized light by a specific number of degrees per the concentration of the compound and the path length of the light. Since the advent of polarimetry other analytical techniques have evolved that are capable of linking molecular chiral structure with observable physical responses. Circular dichroism[15] (CD), both electronic and vibrational (infrared and Raman), has proven a useful technique for probing the configuration of optically active compounds. X-Ray diffraction[83] has become an invaluable tool for determining the absolute configuration of compounds where crystals can be obtained. Recently Raman optical activity[84] has emerged as a promising technique for assigning the absolute configuration of chiral molecules. However for reasons of its ease of use and inexpensiveness, the simple polarimeter remains among the most widely applied probes of chirality.

Since the inception of polarimetry scientists have strived to make a rational connection between the sign and magnitude of chiroptical response and molecular structure. It is important to note that *general* rules to predict the sign (or the magnitude) of the chiroptical response from the molecular structure without performing first-principles computations do not exist. Over the twentieth century several methods have been proposed to overcome this problem. Well-known examples are sector rules for the CD of specific chromophores, of which the carbonyl "octant rule"[85-87] is a subset, as well as the exciton chirality method[88] (also for CD), to name a few. These methods are invaluable tools for linking a molecule's CD to its absolute configuration. Unfortunately, the relation of the sign and magnitude of the optical rotation (OR) to molecular structure is less well understood except for simple cases where one can single out a chromophore

that is almost exclusively responsible for the OR and for which sector rules may apply. One empirical rule for OR that caught our attention was developed in the early twentieth century[89-91], and can be stated as follows: If upon acidification of an aqueous solution of an amino acid its specific rotation becomes more positive, the amino acid is of the "L" absolute configuration. If the opposite is true, then it is of the "D" configuration. This rule, named for Clough, Lutz and Jirgensons[72] (CLJ), is not without exception, but it has been shown to be reliable in the assignment of the absolute configuration of a multitude of amino acids, and has found occasional use as recently as a few years ago.[92]

Since the turn of the twenty-first century, much of the effort to link the sign and magnitude of the chiroptical response measurements to local and global absolute configuration has centered on computational chemistry. This has been recently referred to as a "renaissance in chiroptical spectroscopic methods" which has sparked renewed interest in the field.[18] The current state of the art method for modeling specific rotation for most but the smallest molecules is Time Dependent Density Functional Theory (TDDFT), and the algorithms for performing such calculations are available in several popular program packages.[30, 93-96] For recent reviews of computational techniques and applications, see the works of Polavarapu[1, 18],Crawford[6], Pecul and Ruud,[5] and Autschbach[19]. Many of the papers published in this field of research center around benchmarking and improving upon these methods.[3, 21, 29, 97] [4, 11] There have been many recent advances in the modeling of optical activity, a few examples of which will be mentioned here. Ruud and Zanasi[98], Kongsted et al.[99]and Mort and Autschbach[99, 100] have studied the zero point vibrational contributions to chiroptical response properties. The temperature dependence of chiroptical responses has also been modeled.[100] Mennucci et al[8], Pecul et al[101] and Mukhopadhyay and coworkers[102] have reported methods to model

the effects of solvents on optical rotation. Ruud and Helgaker[37], as well as Crawford and coworkers[44] have applied coupled-cluster methods to computing chiroptical properties, which should yield higher accuracy than the current standard functionals in TDDFT as more powerful computers and new algorithms make it more practical. Such progress means that theory continues to improve and makes the combination of molecular modeling with experimental measurement of specific rotation a more useful tool for the assignment of absolute configuration.

While focusing on improving the accuracy of molecular modeling is important in its own right, sometimes it is useful to take a look back at how the theories of today relate to the theories of the past. Obviously the CLJ rule cannot anymore be considered a major tool for the assignment of absolute configurations of amino acids. However, it is *one of the rare examples* where it appears to be possible to relate the sign of the optical rotation (here: a trend for closely related structures) to the absolute configuration. As such, the CLJ rule is of fundamental interest because, as already mentioned, the relationship(s) between the sign of the optical rotation and the molecular structure remains one of the great enigmas in stereochemistry. One of the goals of computational research in this field is to uncover these relationships which will ultimately lead to practical rules (perhaps similar to the sector rules of CD) for the easy prediction and rationalization of the sign of the OR based on a molecule's configuration. Such rules should have a firm foundation in first-principles theory. Another aspect of this study is the following: There are serious difficulties in modeling the OR of conformationally flexible molecules with small-magnitude ORs (for which amino acids are good examples).[20, 28, 58, 103] Their computational modeling is further complicated by the need for treating solvent effects. However, we will show that the CLJ effect itself is reproduced quite well in the computations. The analysis of the origin of the CLJ effect will demonstrate that exceptions to the rule can be easily

rationalized. Therefore, if an effect of similar type as CLJ can be exploited in studies of other chiral molecules we believe it would greatly improve the predictive power of computation-based assignments of absolute configurations in cases where conformational averaging of ORs causes unacceptable uncertainties in the computational results. Finally, for the carbonyl and the carboxyl chromophore in amino acid zwitterions and their protonated forms we show by mapping out "chiral sectors" that the problem of understanding OR – structure relationships can in some cases be reduced to an analysis of the CD.

We begin with a computational benchmarking on a test set of twelve common L-amino acids to see how well TDDFT can replicate the CLJ effect. Next we focus more closely on alanine, the smallest chiral amino acid, to show why, from a quantum mechanical prospective, the CLJ rule exists the way it does. The reasons the rule is obeyed for many amino acids and disobeyed for a few are discussed. Finally we take a look back to an explanation of the CLJ rule from 3 decades ago, to investigate to what extent this rationale fits with data obtained from first – principles theory.

COMPUTATIONAL METHODS

The computational methods used in this work are detailed in a previous publication,[20] where TDDFT based computations of optical rotations of amino acids were exhaustively benchmarked.[58] A brief summary follows: All data were computed with the Turbomole[30] quantum chemical software, version 5.7.1. The calculations were performed with the Becke three parameter B3-LYP and BHLYP[31] hybrid functionals as implemented in the Turbomole code. Molecular geometries were optimized with the default doubly polarized valence triple zeta (TZVPP) basis set from the Turbomole library; all energies used for Boltzmann averaging were computed with this basis. Response calculations were performed with the aug-cc-pVDZ

basis.[64] All optimizations and response calculations included the COnductor-like Screening MOdel (COSMO)[33] of solvation applied to the ground state. For molecules for which multiple conformers exist, the energies used to compute relative conformer populations at 293K included COSMO corrected electronic energies as well as zero point energies; additional terms needed to compute relative Gibbs free energies were not included since it is not clear how the computed (gas-phase) corrections for ΔG relate to *solvated* molecules. Except where otherwise noted, molar rotations were calculated at the wavelength of the sodium D line (589.3nm). All molar rotations are reported in units of deg·cm^2/(dmol). The center of mass has been used for the coordinate origin for all response calculations. While all results are formally origin dependent, this dependence is minimized in variational methods when large basis sets such as aug-cc-pVDZ are used; see our earlier work and the references cited therin.[58] As a practical test, moving the gauge origin 10 angstroms from the center of mass of an alanine cation produced a change in molar rotation of only 3 deg·cm^2/(dmol). (Using a different program with a formally gauge-origin independent method yielded changes an order of magnitude smaller. The residual gauge-origin dependence can be attributed due to finite convergence thresholds and numerical imprecision in solving the linear response equations).

RESULTS AND DISCUSSION

I. Modeling the Molar Rotation of the Zwitterionic and Cationic (Protonated) Amino Acids in Solution

Before further discussion of the CLJ rule, it must be established that the computational method employed here can reproduce the effect. For this benchmarking twelve optically active L-amino acids were selected: alanine, cysteine, histidine, isoleucine, leucine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. They were chosen primarily because they do not have excessively large conformational spaces, so to limit the number of structures for

which computations were needed. The chiroptical response of each amino acid was computed for the protonation states they adopt in neutral and acidic solutions. All of the amino acids studied here take the zwitterionic form in pure water and become protonated on the carboxylate group at low pH. (Previous work by Pecul et al.[28] mentioned a related change of the optical rotation between protonated and non-zwitterionic proline and alanine in the gas phase, but the work did not directly address the aqueous forms responsible for CLJ) This protonation is found to preferentially occur on the carboxylate oxygen trans to the amino group[58], which is illustrated in figure 6-1. This resulted in one protonated form of each amino acid being found for every zwitterionic rotamer. Histidine is unique among this set in that it becomes doubly protonated in strong acid. The second protonation occurs at the imidazole ring. As a consequence one deprotonated form was considered for each zwitterionic form.[20]

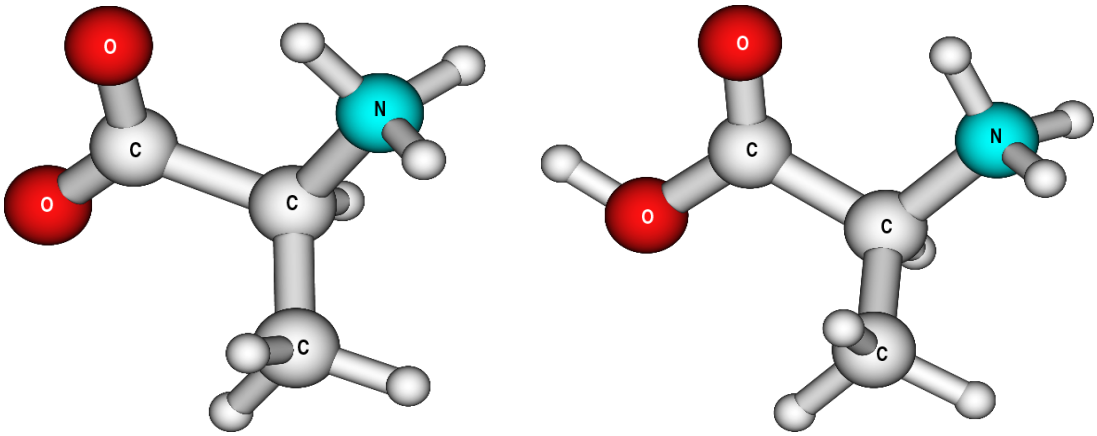


Figure 6-1: Optimized Zwitterionic (left) and Cationic (right) structures of alanine.

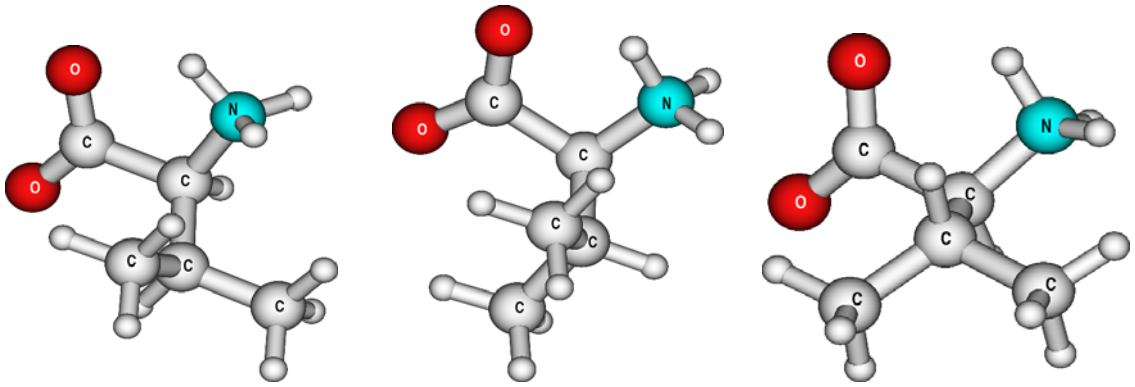
For alanine since only one zwitterionic structure could be found and only one of its protonated structures has a significant population at room temperature, modeling its molar rotation is a relatively straightforward process from the perspective of conformational averaging. Modeling of the larger amino acids is more complicated, since they can be found in multiple conformations at ambient temperature. Of the other 18 common naturally occurring chiral amino

acids, proline is unique in that its ring structure restricts the side chain to two energetically favored conformations. For the rest of the amino acids, the number of possible rotamers that must be considered for a particular molecule depends mainly on the number of carbon – carbon bonds around which rotation may occur. For valine and phenylalanine, rotation about the C_α-C_β bond results in three structures that must be modeled. For the aromatic compounds histidine, tryptophan, and tyrosine, six low energy conformers were found, generated by combining the threefold rotation about the C_β-C_α axis with the twofold rotation about the C_γ-C_β bond. Cysteine, isoleucine, leucine, serine and threonine all contain two threefold axes of rotation which means that nine possible rotamers of each had to be considered in this conformational search, though steric issues dictated that fewer than nine optimized structures were found for some of these. The rest of the amino acids commonly found in mammalian proteins have even greater conformational flexibility than the aforementioned molecules, so for them an *ab-initio* conformational averaging of the optical rotation of all possible rotamers was deemed impractical, at present. Therefore for the purpose of studying the CLJ rule this study is limited to the twelve amino acids already listed.

Identification of every energetically accessible conformer of a molecule is critical to the correct modeling of its molar rotation. Ascertaining the relative populations of these conformers is also extremely important, since the molar rotation of a compound that can be measured experimentally represents a weighted average of this chiroptical response of the molecule in all of its possible conformations. Unlike many other linear response properties, optical rotations of different conformers can differ in magnitude *and* sign. Obtaining reliable Boltzmann averages of amino acid rotamers has been identified as a source of error in the past; especially in cases where

the average can be biased by intramolecular hydrogen bonding, which is difficult to model correctly.[20, 58]

The importance of Boltzmann averaging can be illustrated by using the valine zwitterion as an example. This molecule can adopt three possible conformations in solution, which are illustrated in figure 6-2. The naming designations of the rotamers, t, g, and h, refer to whether a group (one of the two methyl groups in the case of valine) is *trans* or *gauche* to the carboxylate group, or it is sterically *hindered* between the carboxylate and amino group. The three rotamers differ from each other only in the angle about the C_β-C_α, bond, but this distinguishing factor is enough to cause the conformers to yield molar rotations of differing sign. As such, the molar rotation that is modeled for a valine zwitterion depends strongly on the Boltzmann factor that is calculated for each of these rotamers. Errors in these mole fractions will likely result in an erroneous computed molar rotation Boltzmann average.



h,t	**g,h**	**g,t**
0.0kJ/mol	+4.3kJ/mol	+5.3kJ/mol
-52.6	57.8	-60.6

Figure 6-2: Rotamers of zwitterionic valine, along with their computed relative energies at the B3LYP/TZVPP level and molar rotations at BHLYP/aug-cc-pVDZ

The average molar rotations of this as well as the other eleven amino acids modeled are summarized in table 6-1. Details about the geometries used, as well as their individual computed

molar rotations and Boltzmann populations are available in the supporting information. The response properties were originally modeled using the popular B3LYP hybrid DFT functional. In light of concerns about the frequent underestimation of excitation energies by this functional and the consequential overestimation of molar rotation, calculations with another common functional, BHLYP, were carried out for comparison.[104] Results with both functionals are presented. For both sets of response calculations the aug-cc-pDVZ basis set was employed, which is well established for calculating these chiroptical properties.[4, 21]

Table 6-1: Computed and measured molar rotations [ϕ] at 589.3nm for selected L-amino acid solutions. Δ[ϕ] is the CLJ effect.

	Computed Molar Rotation			# of Conformers
Molecule	B3LYP	BHLYP	Experimental	CLJ obeyed for
alanine zwitterion	5.4	-1.5	1.6	
alanine cation	27.2	27.8	13.0	1/1
Δ[ϕ]	21.7	29.3	11.4	
cysteine zwitterion	-51.9	-46.2		
cysteine cation	9.8	11.3	7.9[a]	7/8
Δ[ϕ]	61.7	57.5		
histidine zwitterion	-62.4	-45.7	-59.8	
histidine dication	105.4	88.3	18.3	4/6
Δ[ϕ]	167.7	134.0	78.1	
isoleucine zwitterion	-3.9	-7.4	16.3	
isoleucine cation	84.7	74.9	51.8	9/9
Δ[ϕ]	88.7	82.3	35.5	
leucine zwitterion	-48.8	-52.1	-14.4	
leucine cation	84.5	74.6	21.0	9/9
Δ[ϕ]	133.3	126.7	35.4	
phenylalanine zwitterion	-36.8	-41.4	-57.0	
phenylalanine cation	97.5	81.8	-7.4	3/3
Δ[ϕ]	134.3	123.2	49.6	
proline zwitterion	-127.3	-101.5	-99.2	
proline cation	-93.1	-65.3	-69.5	2/2
Δ[ϕ]	34.2	36.2	29.7	
serine zwitterion	-6.3	-5.6	-7.9	
serine cation	6.7	12.1	15.9	7/7
Δ[ϕ]	13.0	17.7	23.8	
threonine zwitterion	-4.6	-12.6	-33.9	
threonine cation	52.9	48.8	-17.9	7/7
Δ[ϕ]	57.5	61.4	16.0	
tryptophan zwitterion	-46.0	-54.1	-68.8	
tryptophan cation	-46.8	-23.4	13.0	4/6
Δ[ϕ]	-0.8	30.7	81.8	
tyrosine zwitterion	-29.9	-34.8		
tyrosine cation	74.3	57.8	-19.2[a]	5/6
Δ[ϕ]	104.2	92.6		
valine zwitterion	-39.5	-39.2	6.6	
valine cation	90.9	79.4	33.1	3/3
Δ[ϕ]	130.4	118.6	26.5	
TOTAL				61/67

Experimental values from Greenstein and Winitz[72] except where otherwise noted. $\Delta[\phi]=[\phi]_{cation}-[\phi]_{zwitterion}$ in deg·cm^2/dmol

a- Merck Index, 12th ed.[42]

The first comparison to be made in table 6-1 is between the computed and measured signs of the molar rotation values for the zwitterionic and cationic forms of the L-amino acids being studied. For the zwitterionic forms, the computed and measured molar rotations agree in sign for ten of the molecules when the B3LYP hybrid functional is used, and for nine out of twelve molecules with the BHLYP functional. For the cationic (protonated) forms, theory and experiment agree for eight out of eleven molecules for both functionals, with tryptophan being deemed inconclusive due to disagreement amongst experimental values in the literature. The fact that theory and experiment disagree in the sign of the molar rotation for several of these molecules is not unexpected. Many of these values are relatively small in magnitude,[4] and often the result of the partial cancellation of larger optical rotations of conformers as illustrated in figure 6-2.

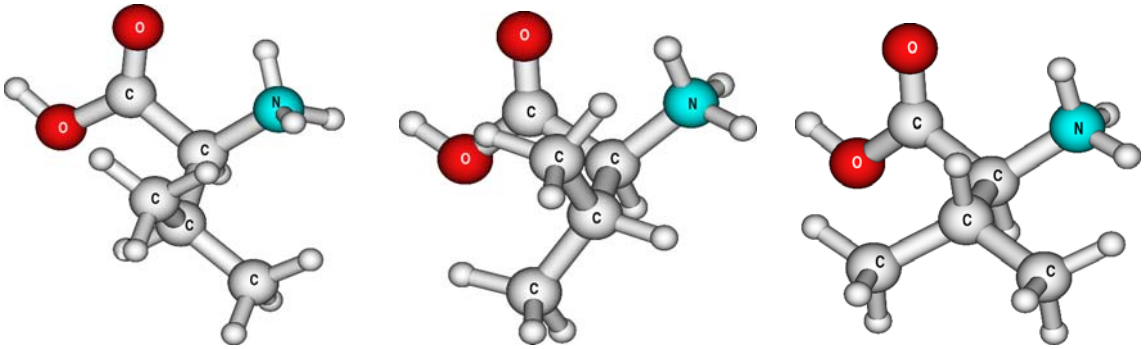
While the correlation of the signs of theoretical and experimental molar rotations yields somewhat mixed results, the comparison of the *change* in molar rotation when a particular amino acid is protonated/deprotonated is far more promising. According to Greenstein and Winitz, the molar rotations for all of the amino acids studied become more positive when they are protonated with hydrochloric acid than when they are in their zwitterionic forms in aqueous solution, and they were the first to name this phenomenon after Clough, Lutz and Jirgensons.[72] When the computed differences in molar rotations are tabulated, we find that with the B3LYP functional this rule is obeyed for eleven out of twelve of these molecules. (We will comment on the outlier later.) With the BHLYP functional the CLJ effect is successfully modeled for all of the molecules in this test set.

The reason that the difference in molar rotations may be modeled with greater reliability than the absolute values of these rotations themselves is obviously due to a balanced cancellation

of errors. Potential sources of error in TDDFT calculations may include: imprecision in calculated molecular geometries, insufficient treatment of solvent effects, neglect of vibrational/thermal effects, overlooking possible non-linear concentration effects, residual gauge origin dependence, approximations in the exchange-correlation potential and kernel, basis set truncation, and inaccurate Boltzmann weighting of conformers. For a discussion of some of these sources of error, see the work of Stephens et al,[29] as well as other works. [5, 82, 104-106] When attempting to model a molar rotation that is relatively small in magnitude, the combination of these errors may cause the sign of molar rotation that is modeled and the one that is measured to disagree. But when modeling two very similar molecules using an identical model chemistry and taking the difference of the results, many of these sources of error can subtract out, leaving behind a computed difference that is meaningful. Since the correct CLJ behavior is obtained for all of the amino acids modeled here it is reasonable to conclude that the *change* in optical rotation is obtained for the correct reasons. This finding suggests that modeling the change in optical rotation, for instance upon protonation (if possible), for a molecule other than an amino acid may be useful for assigning its absolute configuration even in cases where the optical rotation's magnitude in itself is too small to be a reliable measure. However, further studies would certainly be required to confirm this hypothesis.

More insight regarding the utility of the CLJ rule may be gained by looking at the rightmost column of table 6-1. This column lists the number of rotamers that each amino acid may be found in, and how many of the corresponding zwitterion/cation pairs the CLJ rule is obeyed for in the computations. In this case no difference was seen regardless of which hybrid functional was employed, so separate columns for each are not listed.

Consider valine again as an illustrated example. The three rotamers of the valine cation are depicted in figure 6-3. Note that each rotamer structure has a corresponding zwitterionic structure already shown above in figure 6-2. We find that protonation causes a positive change in molar rotation for all of the rotamers: for the h,t rotamer $\Delta[\phi]=+50.4$, for the g,h rotamer $\Delta[\phi]=+75.6$, and for the g,t rotamer $\Delta[\phi]=+94.1$. (To the extent that the relative rotamer populations remain the same upon protonation, these $\Delta[\phi]$ values are comparable to the overall measured $\Delta[\phi]$.) Since all rotamers of valine, as well as isoleucine, leucine, phenylalanine, proline, serine, and threonine obey the CLJ rule, one wishing to use it to assign absolute configuration could chose *any* rotameric form to model and still make the correct assignment. This stands in contrast to assigning absolute configuration based on molar rotation alone, where all low energy rotamers must be modeled, and their resulting molar rotations need to be averaged in the correct proportions to achieve the correct result. As already pointed out, the Boltzmann averaging might introduce additional uncertainties about the quality of the computed results.



h,t	**g,h**	**g,t**
+3.8kJ/mol	0.0kJ/mol	+0.5kJ/mol
-2.5	133.4	33.5

Figure 6-3: Rotamers of protonated valine, along with their computed relative energies and molar rotations (BHLYP/aug-cc-pVDZ//B3LYP/TZVPP)

There are however amino acids for which the CLJ rule is not universally obeyed for all of the conformers. These include cysteine, histidine, tyrosine, and tryptophan. One feature that

these all have in common is that all possess unsaturated functional groups (π orbitals), or, in the case of cysteine, lone pair orbitals on its sulfur atom. These orbitals can take part in electronic transitions in the near UV, and as we have shown previously for the aromatic amino acids, these low energy excitations may have a great effect on the observed specific rotation.[20] The effect of such chromophores is discussed in more detail in the following section.

II. The Relationship between Circular Dichroism Excitations and Molar Rotation

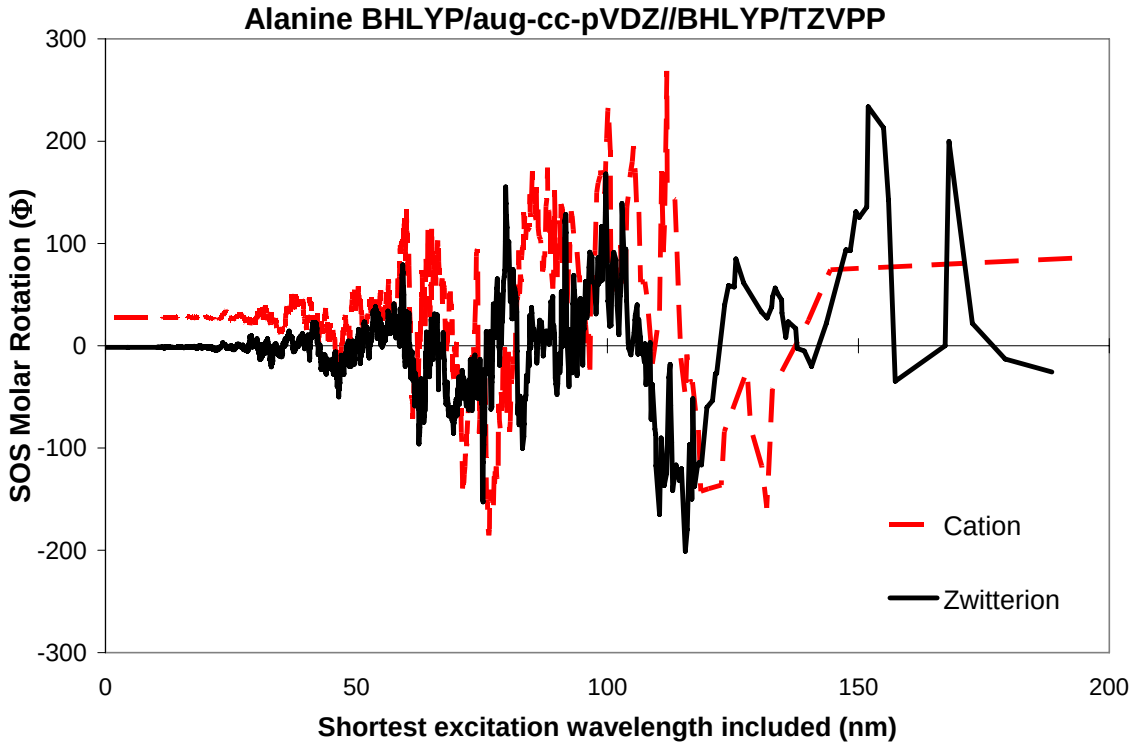
The direct linear response method for computing optical rotation that was used in the first section does well to model transparent-region chiroptical properties, but without further analysis it reveals little about the electronic effects responsible for these effects. To further explore why Clough-Lutz-Jirgensons' rule is obeyed in most cases and why in a few others it breaks down, we decided to investigate the electronic transitions that give rise to molar rotation at 589.3 nm. According to the Kramers-Kroning relationship,[65] or the Sum Over States (SOS) equation for molar rotation (equation 6-1 below, where typically $\omega<\omega_{01}$), one may expect the lowest energy CD transitions to be the most responsible for optical rotation in the transparent region because of the relatively large denominators.

$$[\phi]_\omega = 91.43028 \cdot \omega^2 \sum_{n\neq 0} \frac{R_{0n}}{\omega_{0n}^2 - \omega^2} \qquad (6\text{-}1)$$

Here [ϕ] is in units of deg·cm^2/(dmol), ω is the angular frequency of light (2π·ν) of light, is the *n*'th excitation frequency out of the ground state with R_{0n} the rotatory strength of the transition in units of 10^{-40} esu^2·cm^2. The sum runs over all excited states and continuum states. We define the incomplete summation up to excited state no. *k* as

$$[\phi]_\omega^k = 91.43028 \cdot \omega^2 \sum_{n=1}^{k} \frac{R_{0n}}{\omega_{0n}^2 - \omega^2} \qquad (6\text{-}2)$$

In figure 6-4 we have plotted $[\phi]_{\omega}^{k}$ for the prototypical chiral L-amino acid, alanine. The first excitations in the respective zwitterionic and cationic forms are largely responsible for the CLJ effect. The top part of this figure displays the truncated SOS molar rotation, Eq. (2), of alanine in the two forms of interest as a function of the wavelength of the highest-energy electronic excitation, *k*, which is included in a calculation of the truncated SOS molar rotation according to equation (6-2). The bottom of figure 6-4 depicts the same data, except here as a function of *k*, the number of states included in $[\phi]_{\omega}^{k}$.[107]



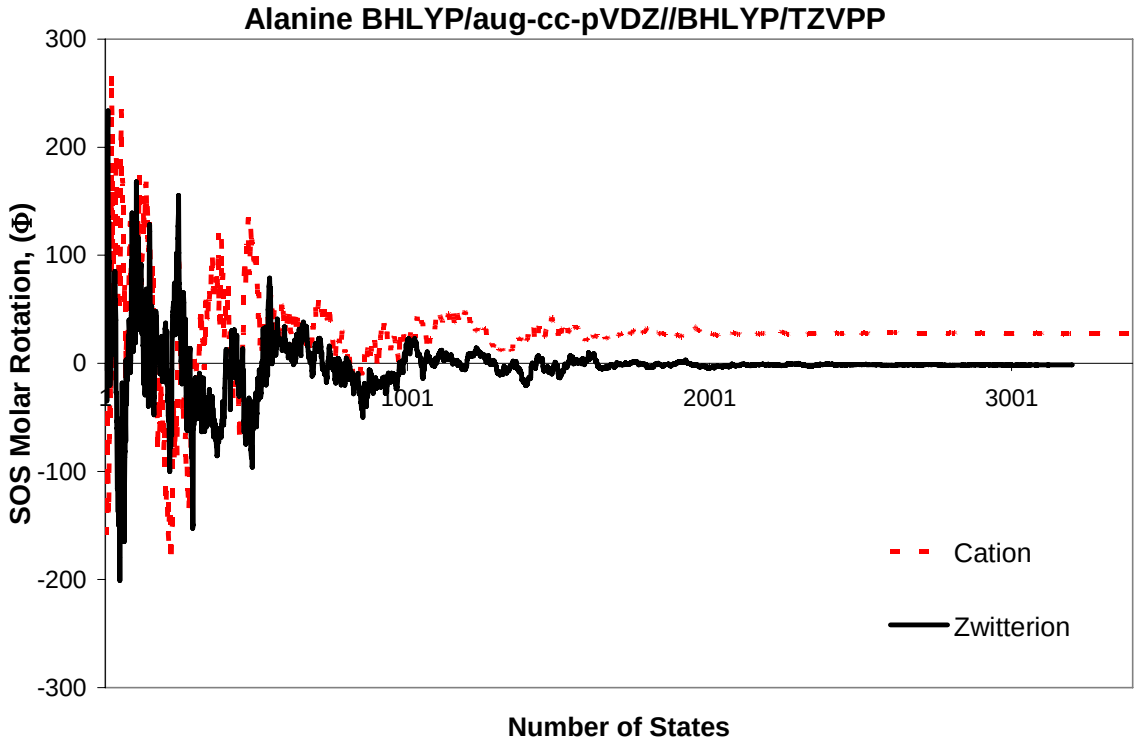


Figure 6-4: The molar rotation of alanine computed from the truncated sum over states, $[\phi]_{\omega}^{k}$ of equation 6-2, versus the absorption wavelength for state number *k* (top) and versus the number of states, *k* (bottom).

If all excited states possible for a given basis set (whether physically meaningful or not) are included in the summation it ultimately yields the same results as the linear response computation.[107] The sum converges on the right side of the second graph (the left side of the first graph) as the number of states approaches the limit of the basis set. The advantage of using the SOS equation is that it allows us to investigate not just what the molar rotation of a compound is, but also which excitations are responsible for that response, according to the SOS analysis. Here it is shown that the lowest electronic excitation, at the far right side of the first graph, causes a pronounced CLJ effect in the molar rotation. That is, at this first step of the SOS summation, the molar rotation of the cationic form of the L-amino acid is seen to be more positive than that of the zwitterion. Higher lying excitations dampen this effect. The sum rule,

[108] $\Sigma_n R_{0n} = 0$, indicates that such damping is more likely than not, since the sum of the rotatory strengths beyond the first must equal the opposite of that of the first excitations so that their sum can be zero. Experience shows that the R_{0n} values strongly oscillate as n increases.[81] However, this R_{0n} damping only diminishes the magnitude of the CLJ effect observed to result from the first CD excitations, and the sign of the $\Delta[\phi]$ for the zwitterion → cation reaction remains the same in the end: it is positive for L forms of the amino acids. The analysis of the alanine CD spectrum therefore strongly suggests that the structural origin of the CLJ rule for α-amino acids can be rationalized by examining the trend for the lowest CD transition upon protonation. This OR / CD relationship will be investigated in more detail below. As a general disadvantage of the SOS analysis we note the abundance of large contributions in the sum which makes it difficult to single out a few important terms without applying a "bias" of physical reasoning that assigns the low lying excitations a particular significance. However, exclusion of this excitation in the SOS would mean that the CLJ effect is *not* obtained.

Computing the entire sum of electronic excitations is an arduous task, which becomes impractical with our chosen basis set for the larger molecules in our test set. However calculating the molar rotation resulting from just the lowest CD transition is straightforward, and as figure 6-4 has shown it is this transition that appears to be very influential regarding the CLJ effect under investigation. Therefore, the portion of molar rotations resulting from these lowest energy CD transitions was modeled for the rest of our test set of molecules. The results are summarized in table 6-2. We note here that all CD transitions are computed with the length gauge dipole representation, however computations of partial molar rotations with the velocity gauge dipole always agreed in sign and agreed in magnitude with those with the length gauge, therefore origin dependence does not effect our conclusions.

Table 6-2: Computed longest excitation wavelengths, λ, and partial molar rotations $[\phi]_{589.3nm}$ from this excitation for selected L-amino acid solutions. Δ = cation value – zwitterion value. Molar rotation [ϕ] is in deg·cm^2/(dmol).

	Excitation Wavelength, λ (nm)		Molar Rotation caused by first excitation [ϕ]		Lowest Energy
Molecule	B3LYP	BHLYP	B3LYP	BHLYP	Chromophore
alanine zwitterion	211.4	188.6	-147.8	-25.7	carboxylate
alanine cation	203.9	192.6	106.7	85.5	carboxylic acid
Δ	-7.4	4.0	254.5	111.2	
cysteine zwitterion	218.0	204.6	6.2	31.1	thiol
cysteine cation	230.8	204.5	31.9	16.0	thiol
Δ	12.8	-0.1	25.7	-15.1	
histidine zwitterion	229.7	209.1	-93.5	-38.6	imidizole
histidine dication	211.1	196.5	683.3	302.6	imidizole
Δ	-18.6	-12.6	776.8	341.2	
isoleucine zwitterion	210.0	188.6	-162.6	-36.4	carboxylate
isoleucine cation	206.6	194.6	180.1	123.8	carboxylic acid
Δ	-3.4	6.0	342.7	160.2	
leucine zwitterion	208.3	188.1	-103.0	-19.0	carboxylate
leucine cation	207.5	196.1	208.5	150.4	carboxylic acid
Δ	-0.8	8.0	311.5	169.4	
phenylalanine zwitterion	234.7	223.1	-5.9	-3.7	phenyl
phenylalanine cation	235.1	222.9	17.1	-0.7	phenyl
Δ	0.4	-0.2	23.0	3.0	
proline zwitterion	211.3	188.6	-60.5	-24.0	carboxylate
proline cation	204.9	193.2	102.9	73.6	carboxylic acid
Δ	-6.4	4.6	163.4	97.6	
serine zwitterion	209.9	188.3	-128.2	-36.5	carboxylate
serine cation	207.2	195.2	132.4	79.6	carboxylic acid
Δ	-2.7	6.9	260.6	116.1	
threonine zwitterion	207.3	186.7	-89.2	-24.0	carboxylate
threonine cation	208.6	196.7	190.0	108.2	carboxylic acid
Δ	1.3	10.0	279.2	132.2	
tryptophan zwitterion	277.4	254.3	-55.0	-37.6	indole
tryptophan cation	280.1	247.4	-8.2	-44.1	indole
Δ	2.7	-6.9	46.8	-6.5	
tyrosine zwitterion	253.2	236.3	24.0	15.7	phenol
tyrosine cation	258.0	235.4	207.6	32.6	phenol
Δ	4.8	-0.9	183.6	16.9	
valine zwitterion	215.4	189.4	-162.8	-41.8	carboxylate
valine cation	212.0	197.2	149.1	113.2	carboxylic acid
Δ	-3.4	7.8	311.9	155.0	

When the rightmost column of table 6-2 is compared with the corresponding column in table 6-1, the pattern becomes obvious. For all the molecules in the set for which the CLJ rule is not obeyed for all rotamers, there exists a chromophore that gives rise to a lower energy electronic transition than that of the carboxylate/carboxylic acid present in all zwitterionic/cationic amino acids. For most of the molecules for which CLJ is obeyed for all rotamers, one of the carboxylate/carboxylic acid chromophore excitations is the lowest in energy. The only molecule with an aromatic chromophore for which the CLJ rule is obeyed for all conformers is phenylalanine, which has been shown to have a very weak π to π^* transition in its phenyl group, as the nearly D_{6h} symmetry of the group makes this transition strongly dipole forbidden.

The data set for the 12 L-amino acids further supports the conclusion that Clough-Lutz-Jirgensons' rule can be rationalized by the study of low lying electronic transitions in the carboxylate and carboxylic acid chromophores. Furthermore, they indicate that any chromophore that has electronic transitions that are low in energy can potentially interfere with this effect, eventually causing the rule to break down. For cysteine, histidine and tyrosine, our calculations hint at the beginning of a breakdown since the rule is disobeyed for one or more pairs of conformers, though for the ensemble average of conformers the rule is still obeyed. Experimental data indicate that CLJ is valid for cysteine and histidine, but no published information is available for tyrosine. For tryptophan, for which we earlier showed[20] that the indole group, not the carboxylate chromophore dominates optical rotation at the sodium D-line, our computations yield somewhat ambiguous results. Response calculations with both the BHLYP and B3LYP functionals indicate that the CLJ rule is obeyed for only four out of six rotamer pairs, but disagree on whether or not the rule is obeyed when these results are averaged

to give the result that should be physically observable, see table 6-1. Experimental data are equally ambiguous on the last question, with some values indicating that tryptophan should obey the CLJ rule and some indicating that it should violate it. For discussion of this issue, see our previous work.[20]

In summary, it is apparent from the data that the n to π^* transition of the carboxylate/carbonyl chromophore can be regarded as responsible for the CLJ rule. Without this excitation the CLJ effect would not be obtained. For all the zwitterionic alpha L-amino acids studied, the circular dichroism of this transition is negative in sign, resulting in a negative contribution to their specific rotations at 589.3nm. For all the cationic (protonated) L-amino acids studied the CD of this transition is positive, causing a positive perturbation in their specific rotations. Ultimately, for all of the L-amino acids where the lowest energy electronic transition is centered on this chromophore, the specific rotation will tend to be more positive in the protonated form than in the zwitterionic state, which is exactly what is observed experimentally.

III. How the SOS-based explanation of the CLJ rule relates to empirical reasoning based on overlapping CD sectors

One semi-empirical approach to rationalize circular dichroism that was developed decades ago is the octant rule. For a summary of the history and development of the octant rule see Lightner's chapter in reference 1. This rule was first developed for ketones[109] then adapted for use in lactones.[110] It is this octant rule for lactones which Jorgensen (not to be confused with Jirgensons from CLJ) first used to rationalize the CLJ rule,[16] by using an argument similar to ours relating the lowest – energy CD transition to the change in molar rotation.

The way in which Jorgensen applied the sector rule in his CLJ explanation is illustrated in figure 6-5. The principal idea was to represent the COO^- group of a zwitterion as an overlapping system of two C=O chromophores and to superimpose the respective sectors. Focusing on just the left part of the illustration will reveal how the area about a C=O group is separated into octants by planes. Two of the planes are dictated by the C_{2v} site symmetry of a carbonyl group. The other surface, which need not even be planar, varies throughout differing permutations of the octant rule. Moffit et al. originally assigned it as a plane bisecting the C=O bond[109]; here Jorgensen assigned it as a plane bisecting the carbon atom normal to its bond with the oxygen.

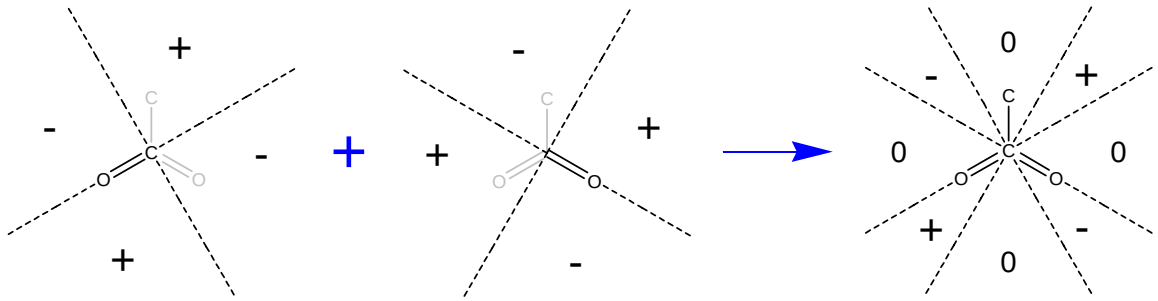


Figure 6-5: Formation of Jorgensen's sector rule for the carboxylate chromophore.[16] Plus and minus signs shown refer to the sign of the sector above the plane of the paper, signs of corresponding sectors below the page are the reverse.

The interference pattern on the right side of figure 6-5 represents Jorgensen's vision of the sectors that rationalize the optical activity of an amino acid in its zwitterionic form. This pattern is a simple supposition of the octants for two overlapping carbonyl chromophores, to form a carboxylate chromophore. For an L-amino acid, the perturbing group that makes the compound chiral is always centered above the plane of the paper in the upper right corner of the circular sector pattern, as is shown in figure 6-6. (Note that the acidic proton that differentiates between the two structures is in the plane of this page, which is a nodal plane in the sector rule, and as such it has no direct effect on the optical rotation in the chromophore.) For the protonated

form, only one carbonyl group is present and so this perturbing group is oriented in a positive octant, and gives a positive contribution to specific rotation. For the zwitterionic form, depicting the chromophore as two C=O groups yields a more complicated set of sectors, and in this case the perturbing group falls into a zone that is far less positive than it was in for the protonated form.

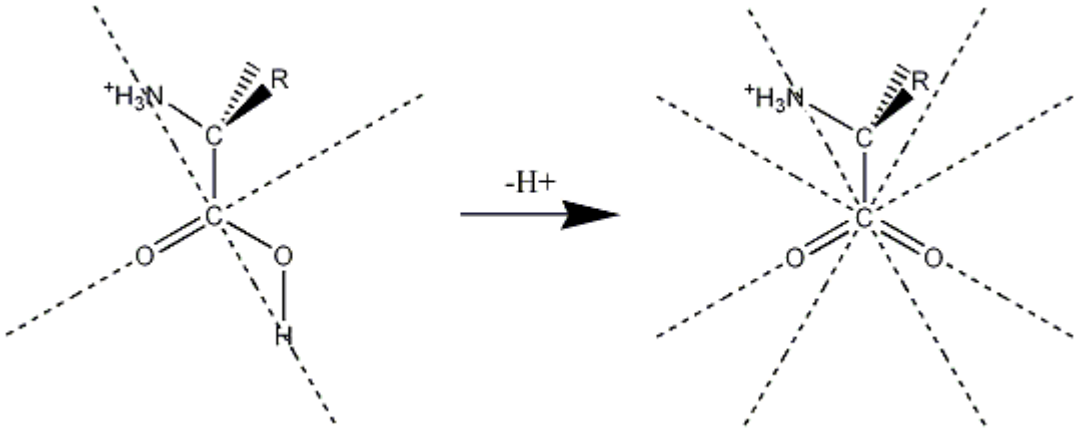


Figure 6-6: How protonated (left) and zwitterionic (right) L-amino acids fit into Jorgensen's sectors

Given the fact that the optical rotation can be written as a sum over all the CD transitions, Eq. (6-1), and the fact that the denominator in the equation is the smallest for low-energy transitions and ORs in the transparent region it appears natural to reduce the OR – structure problem to the CD – structure problem and try to apply CD sector rules. However, there are several important caveats: For one, *all* CD transitions contribute to the OR, and CD intensities can grow very large for high lying excitations, thereby overpowering the diminishing effect from the denominator in Eq. (1)[81]. Without establishing by reliable theoretical methods that the OR – CD connection can indeed be made for the lowest-energy transition Jorgensen's argument is not sufficiently strong.

Empirical rules based on chromophores are not absolute, and have been known to fail to accurately model chiroptical response properties on occasions where more generally applicable

first-principles methods such as TDDFT have succeeded.[111] Jorgensen's overlapping sector rational seems appealing, though the shape of those sectors need not be as strictly symmetrical as implied in figure 6-5 unless the third plane dividing the sectors is precisely where Jorgensen assigns it. We decided to look at this assignment more closely to see if those sectors could in fact be "mapped out" using TDDFT by choosing a perturbing moiety to move about the carbonyl and carboxylate chromophore to investigate the variation of chiroptical as a function of the perturber's position.

Another assumption which is central to the rationalization of the CLJ rule that would need to be firmly established is Jorgensen's treatment of the carboxylate chromophore as a simple set of overlapping C=O chromophores, and that this model yields the correct behavior upon amino acid side chain perturbation. As we will show below one needs to consider the overall change in charge of the molecule upon protonation which complicates the model.

All the amino acids that exhibit the CLJ effect can be conceptually regarded as derivatives of glycine. Although the side chain that induces such an amino acid's chiroptical response varies considerably amongst the amino acids, what remains constant is the position of this perturbing group relative to the carbonyl / carboxylate chromophore: for all L-amino acids, this group is on the same side of the carbonyl / carboxylate mirror plane. We have investigated whether the CLJ effect could be induced in the glycine molecule by placing a perturbing moiety in this appropriate position. A glycine zwitterion and cation with a generic perturber in the appropriate position are illustrated in figure 6-7.

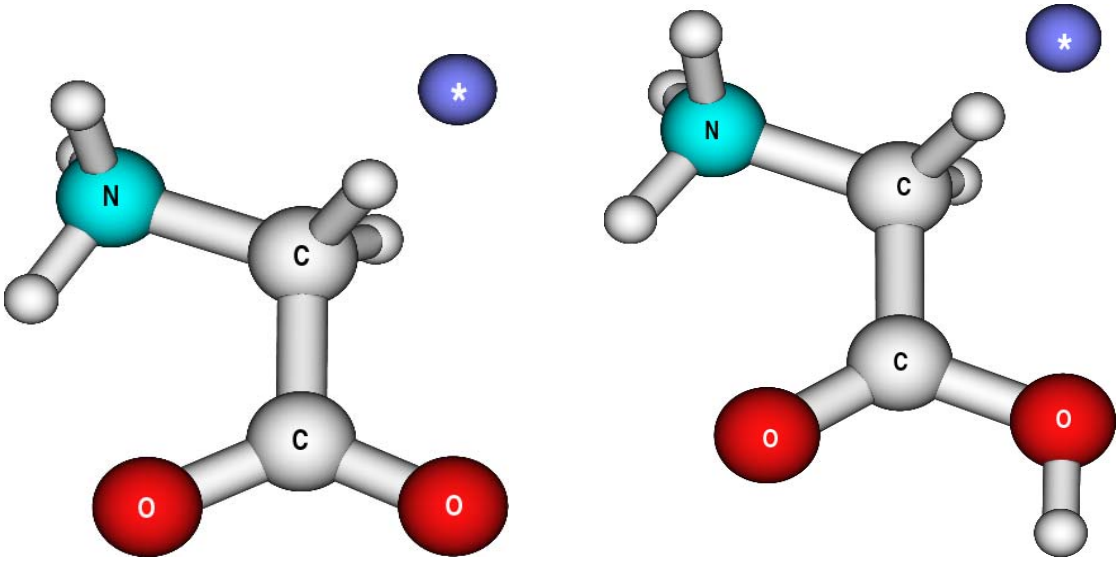


Figure 6-7: A glycine molecule with a chiral perturber in a position designed to mimic the chiroptical response seen in the L-amino acids. For our example we have placed noble gas atoms and point charges in the position marked with an asterisk (*) to observe the resulting chiroptical response.

In order to make the model computationally feasible, the perturbing group which is responsible for the dissymetry of the amino acid should be as simple as possible. The first perturber we used was a neon atom, placed 2.2 angstroms from the α-carbon, with a C-H-Ne angle of 180 degrees to simulate a neutral group causing steric interaction. The results of the CD spectrum for such perturbed glycine systems, along with the CD spectra of natural chiral L-amino acids are depicted in figure 6-8.

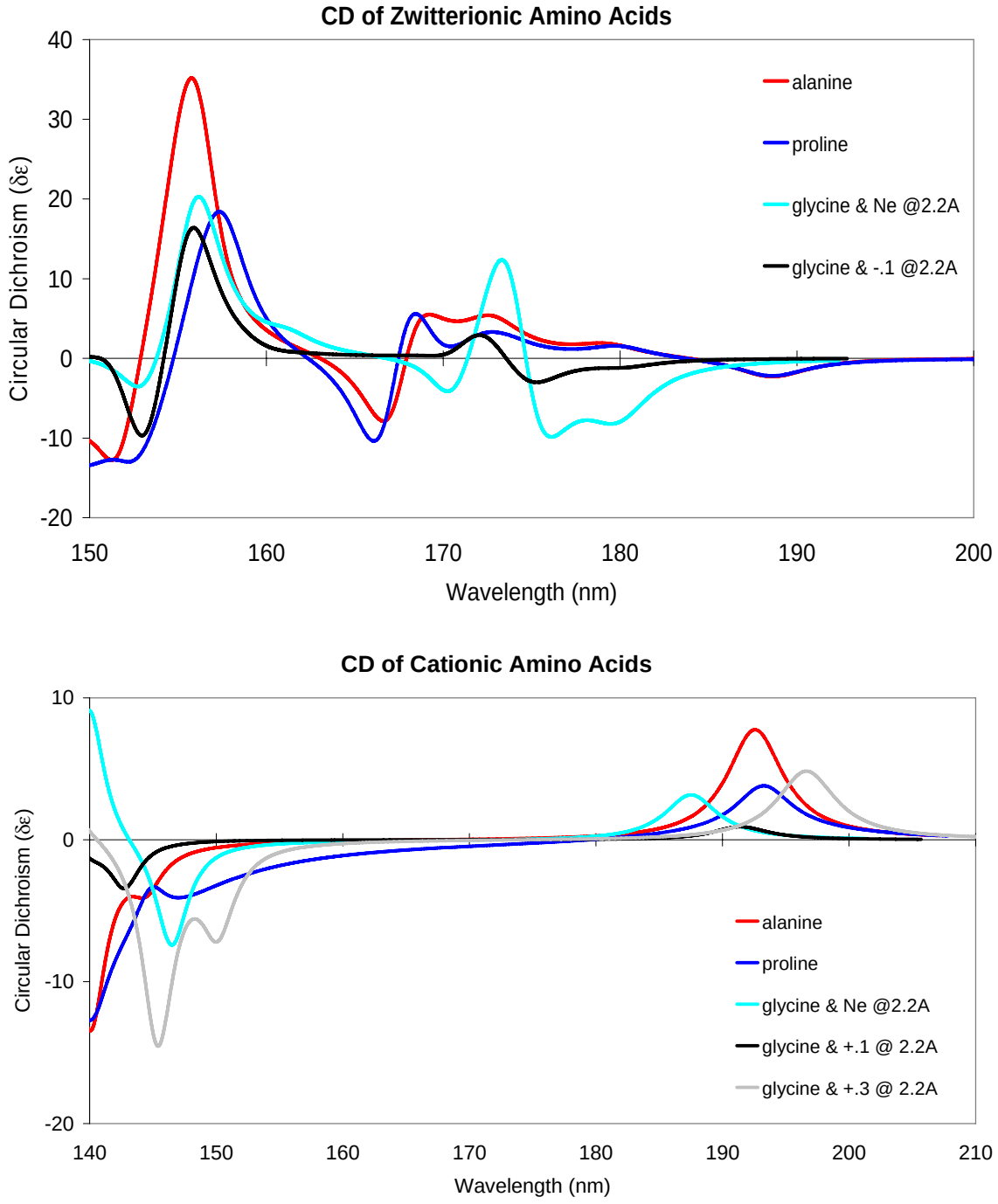


Figure 6-8: Computed near ultraviolet circular dichroism spectra of glycine perturbed by point charges and by neon atoms as well as that of alanine and proline. See text for details. Spectra for zwitterionic (top) and protonated structures (bottom), respectively. An empirical Gaussian broadening of 0.09eV was used for the plots.

For both the zwitterionic and cationic amino acids, the chiroptical response manifested in the lowest energy circular dichroism transition that is responsible for the CLJ effect can be successfully reproduced by modeling the appropriate form of a glycine molecule perturbed by a neon atom. For each CD spectrum the sign of the first Cotton effect of the perturbed glycine matches that of the corresponding chiral amino acids, of which alanine and proline are the examples displayed here. We found similar results were obtained with a helium atom used as the perturber (not shown). As the data from the test set of 12 L-amino acids suggest, the identity of the perturbing group causing the overall sign of the OR and of the CLJ effect is not important, so long as it does not give rise to electronic excitations that are lower in energy than those of the carbonyl/carboxylate chromophores. From a computational perspective this suggests that this perturbed glycine model can be simplified even further, by using a simple point charge. Point charges can also model inductive (electron withdrawing or electron donating) effects from the side chains on the carbonyl and carboxylate chromophores. For zwitterionic glycine a point charge of -0.1, when placed in the region where functional groups are attached to glycine's prochiral carbon in the L-amino acids, can induce virtually the same chiroptical response in the lowest energy CD transition of that molecule as is seen in the natural chiral amino acids. The first CD transition of the glycine cation can also be perturbed to appear similar to those of the chiral amino acids, but for this protonated form a *positive* point charge is needed to simulate the positive chiroptical response seen in the protonated forms of the chiral amino acids. For the cationic forms of the amino acids a glycine perturbed by a positive charge of +0.1 produced the correct sign of the first CD transition, although this transition was somewhat weak compared to the corresponding transitions seen in the chiral amino acids, alanine and proline. However adjusting the magnitude of the charge allowed for a better match of the intensity, and as can be

seen in figure 6-8 a perturbing charge of +0.3, when placed in the appropriate position relative to glycine, affords a lowest energy CD response of a magnitude more comparable to those of the alanine and proline cations.

One important feature present in the CD spectra as well as in the partial sum of optical rotations of Fig. 4 is that for the protonated species the first CD transition is well separated from the higher lying excitations. For the zwitterions, this is not the case. That is, when comparing the zwitterionic to the protonated forms not only the signs of the lowest-energy CD transitions change but also the energies of the next several CD bands decrease considerably leading to several positive and negative CD bands in the range down to about 150 nm. In Fig. 4 it is shown for alanine that the partial rotations from these excitations largely cancel. But their influence and that of higher lying excitations may be seen in sector patterns for the optical rotation that may lead to different qualitative appearances as those for the CD.

After identifying a simple perturber that can reproduce a CD transition pattern in glycine similar to those of chiral aliphatic amino acids (a CD pattern that reproduces the CLJ effect in a SOS equation) we have mapped out the chiroptical response to obtain sectors for optical rotation and CD derived from first-principles computations. To plot the CD and optical rotation sectors, we moved the perturbing charge by a step size of 0.1 angstrom over a glycine molecule and calculated the resulting lowest energy CD transition as well as the molar rotation at 589.3nm each time. The perturbing charge was held at a fixed distance above the glycine molecule's plane at 1.3 angstroms, which is approximately the distance that the perturbing side chain is located above the chromophore plane in the chiral amino acids that exhibit the CLJ effect. The results of these calculations are depicted in figure 6-9. The regions where the perturbing charge

induces a positive first CD transition and a positive molar rotation are depicted in white; the negative regions are colored gray.

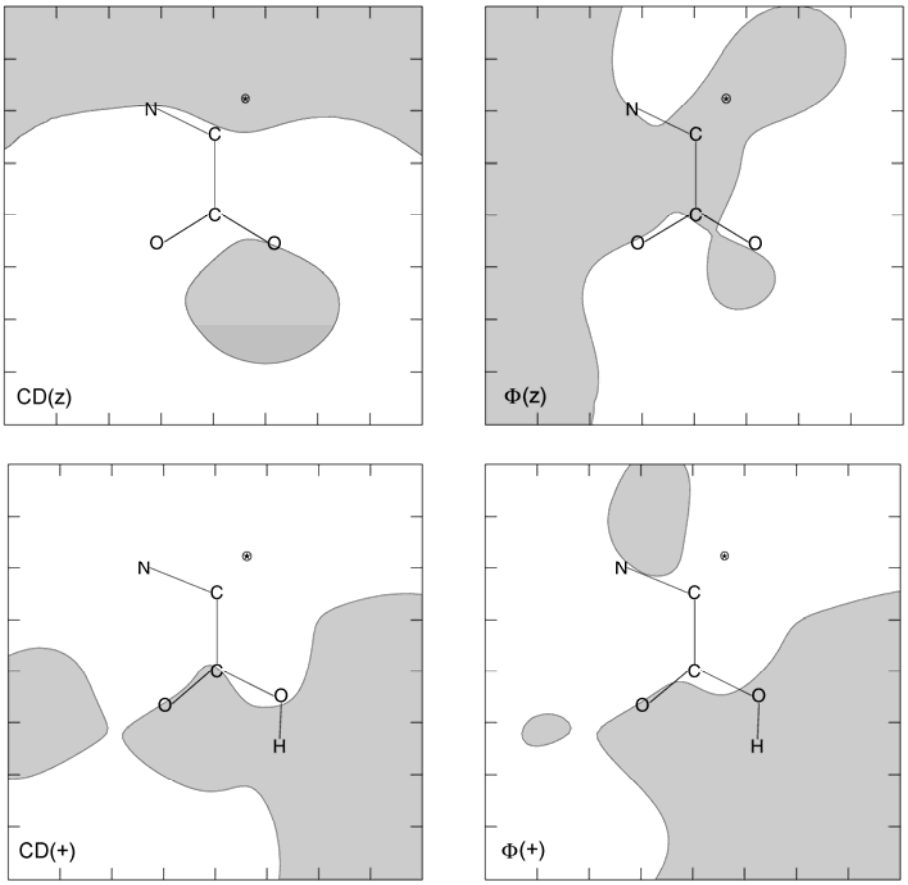


Figure 6-9: TDDFT computed "sectors" of the neutral zwitterionic (z, top) and protonated (cationic, +, bottom) glycine molecule. The left two pictures illustrate the sign of circular dichroism of the corresponding lowest energy CD transitions; the right two illustrate sectors for the molar rotation. Positive and negative areas are represented by light and dark shading, respectively. Note the greater correlation between the molar rotation and first CD sectors for the cationic form of the amino acid than for the zwitterionic form. The chromophore is in the plane of the page while the sectors are 1.3 angstroms above. The perturbing group was a point charge of +0.3 for the acid and -0.1 for the carboxylate. The approximate point of attachment of the side chain groups for L-amino acids is designated by an asterisk.

One immediately notes the correlation in the –COOH and $–COO^-$ regions between the circular dichroism sectors and molar rotation sectors for the glycine zwitterion and cation. This correlation forms a critical part in the overlapping-sectors explanation of the Clough-Lutz-Jirgensons rule, since the CLJ rule relates the *molar rotations* of the protonated and deprotonated

forms of an amino acid, while the sector rule for lactones which Jorgensen invokes is a rationale for the sign of the first *circular dichroism* excitation. Unfortunately, Jorgensen's model simply assumes this correlation exists, even though it is not necessarily the case for both the forms of the amino acids. For the protonated form of glycine, a strong relationship between the molar rotation and first CD excitation is in fact readily apparent in the illustrations on the right side of figure 6-9. Also, as already pointed out, for the electron deficient, cationic forms of the amino acids there is a large separation between the first and second electronic circular dichroism transitions, with the first occurring over 190nm and the second not until under 150nm; see figure 6-8.

The sectors for the zwitterionic forms of amino acids differ from those of the corresponding cations. There is some connection between the signs of the CD and molar rotation for the zwitterions (figure 6-9, left), but this correlation is not as good as that of the cations which we attribute to the aforementioned lowering of the energies of the higher excited states. As such, Jorgensen's assumption that molar rotation *sectors* are necessarily equal to CD sectors is undermined notwithstanding the fact that we indeed found from the SOS analysis that the CLJ effect can be attributed mainly to the lowest excitations from the C=O and COO^- chromophores. From Figs. 4 and 8 it is clear that CD excitations beyond the lowest energy one are apt to have more of an effect on the molar rotation of the amino acid zwitterions than on the cations and therefore have an impact on the computed molar rotation sector patterns. In the region where the perturbing side chain is located (* in Fig. 6-9) we note that the signs of the CD and the OR are the same for the zwitterionic and the protonated species, respectively. Therefore, for the perturbed glycine model the correct sign for the CLJ effect would be obtained from using the computed sector maps for the optical rotation as well as for the partial rotation from the lowest-

energy CD transition. However, we remind the reader that the sectors for the zwitterions and the protonated form, respectively, have been obtained with opposite signs of the perturbing charges in order to yield the same behavior as found for chiral amino acids whereas Jorgensen's model does *not* require opposite sector patterns. We will discuss this issue in more detail below.

Whether the computed sectors for the glycine zwitterion closely resemble those of Jorgensen's overlapping sector model, shown in figures 6-5 and 6-6, depends on the choice of the nodal plane for the C=O group. The computed sectors for the glycine zwitterion do bare some similarity to those of overlapping carbonyl octants, but only if the third nodal plane bisects the C=O bond, as Moffit et al. originally assigned it.[109] Because we have shown that the correlation between the shape of CD and molar rotation sectors for the amino acid zwitterions is not particularly good, the similarity between Jorgensen's overlapping *circular dichroism* sectors and our computed *molar rotation* sectors may be a coincidence. TDDFT calculations that we performed on a perturbed formate anion (not shown) for which the overlapping sector model should apply as well yielded little correlation between the sign of the first CD transition of a perturbed carboxylate anion and the sign of that system's molar rotation. Also, the C_{2v} symmetry implicit in Jorgensen's carboxylate model has obviously been quite strongly perturbed in the glycine chromophore, i.e. the two CO bonds are visibly different in the calculated sectors which further renders a direct application of an overlapping sector model difficult. The CD sectors seem to exhibit a more strongly pronounced octant effect about the carbon-oxygen bond that is farther from the amino group. This makes sense since this C=O bond should have more double bond character than the bond closer to the amino group, since the latter may be participating in intramolecular hydrogen bonding to some extent. The distortion of the

overlapping octant behavior is significant and extends into the region on the sector map where the side chain perturbation is located.

Despite the computational results showing that Jorgensen's simple overlapping-sector model is not straightforwardly applicable to the optical rotation we have nonetheless argued that the CD of the lowest transition can be made responsible for the CLJ effect. This may appear like a contradiction but it is not. Jorgensen's assumption of modeling CLJ with the circular dichroism was correct in the sense that the SOS equation indeed yields a (very strong) CLJ effect for just the first excitation. However, the sectors as computed from first principles only loosely resemble those of Jorgensen, and only near the C=O and COO^- chromophores, not in the region where the side chain actually appears as the perturber (indicated by an asterisk in Fig. 6-9). In particular, if we consider the sign of the optical rotation obtained from the TDDFT sector maps in the region indicated by the asterisk then the correct sign of the CLJ effect (upon subtraction) is only obtained if the sectors about the C=O bond in the protonated form appear not as octants but as *anitoctants*. This finding agrees with the choice of the sign of the perturbing charge necessary to induce a CD pattern in glycine that has the same sign as those for chiral amino acids (as previously discussed, see Fig. 8). In the past an "antioctant rule" has been applied to perturbing groups that are strongly electron withdrawing, like fluorine.[112] In this case of amino acids the perturber is an aliphatic or aromatic R group, which in the zwitterionic form exhibits the normal octant effect. But in the protonated form, where the molecule has an overall positive charge, we see that this perturbing group may induce an overall effect more akin to an electron withdrawing perturber because it is competing for negative charge with other groups in the electron deficient species. Seeing an organic perturbing group switch from causing an octant to antioctant effect, or in other words switching from having a *consignate* to a *dissignate* contribution to chirality is

not unheard of. In fact the methyl group, the perturbing group in the simplest amino acid to obey the CLJ rule, has been known to switch from dissignate to consignate based on the polarity of the solvent used.[87] Here the perturbing groups, be they methyl (alanine), isopropyl (valine), isobutyl (leucine), et cetera all appear to change from consignate to dissignate upon protonation of their respective zwitterions. This change in behavior of the perturbing group based on the charge of the molecule to which it is attached was not considered in the empirical overlapping C=O sector model. Therefore, this simple model does not agree with the results of the TDDFT-based analysis despite the fact that in both cases the lowest CD transition is made mainly responsible for the overall occurrence of CLJ. The computations performed in this work along with the sector maps derived from these computations strongly suggest that a consignate – dissignate change along with the change of the overall charge is an important factor in the reasoning behind the CLJ rule.

SUMMARY AND CONCLUSIONS

The Clough-Lutz-Jirgensons (CLJ) effect has been successfully modeled by TDDFT for a set of 12 L-amino acids. A sum-over-states analysis of the molar rotation shows that the carboxylate / carboxylic acid chromophore is largely responsible for the effect. This explains why alpha amino acids with no other chromophores always obey the CLJ rule, while those with additional chromophores with excitations that interfere with those of the carboxyl/carboxylate group sometimes do not. In addition, TDDFT has been used to map out CD and molar rotation sectors for the amino acids and to show the effect of side chain position on these chiroptical response properties. Within the framework of the sector rules, the change in optical rotation upon protonation of an amino acid zwitterion results not only from a change in the geometry of

the sectors upon protonation but also from a change in the action of the perturbing group: in the zwitterion this group appears to act quite similar to an electrostatic or steric repulsion, whereas in the cation the perturbing group is best modeled via a slight electrostatic attraction, resulting in a consignate contribution to chirality in the deprotonated form of amino acids and a dissignate contribution in the protonated forms. This fact that a perturbing group can have two differing effects on optical activity depending on the overall charge of the molecule was not indicated in the earlier rationale of the CLJ rule, but becomes apparent from the results of the TDDFT computations. We hope that this knowledge will aid future investigations into better linking chiroptical response and molecular structure. For instance, similar effects as CLJ may be used for the assignment of absolute configurations of molecules other than amino acids using optical rotation measurements along with TDDFT computations. Such a procedure of comparing the *change* of optical rotation for closely related species to assign their unknown absolute configurations with the help of computations may be particularly useful for situations where the optical rotation itself is too small to yield a reliable configurational assignment or where conformational averaging adds significant uncertainties to the computed results.

Acknowledgment

We wish to thank the University at Buffalo Center for Computational Research (CCR) for support. This work has received financial support from the CAREER program of the National Science Foundation (CHE-0447321). JA thanks Prof. D. Dyckes for pointing out the existence of the CLJ rule several years ago.

Supporting Information Available: Optimized structures and energies of the 12 amino acids studied in this work. This material can be obtained free of charge through the American Chemical Societies' publications website at http://pubs.acs.org

CHAPTER 7: MODELING THE CHIROPTICAL RESPONSE OF GLYCINE AND ALANINE USING EXPLICT SOLVATION AND MOLECULAR DYNAMICS

Kundrat, Matthew D.; Autschbach, Jochen. Time-dependent density functional theory calculation of specific rotation using molecular dynamics and explicit solvation. Poster *and* talk presented at 233rd ACS National Meeting, Chicago, IL, March 25-29, 2007; PHYS 566.

Kundrat, Matthew D.; Autschbach, Jochen. Modeling the effects of an achiral solvent on the chiroptical response properties of chiral amino acids. Talk presented at 40th Midwest Theoretical Chemistry Conference, Ann Arbor, MI, June 26-28 2008.

Kundrat, Matthew D.; Autschbach, Jochen. Ab-initio and density functional theory modeling of the chiroptical response of glycine and alanine in solution using explicit solvation and molecular dynamics. Journal of Chemical Theory and Computation *Submitted.*

ABSTRACT

We investigate ways in which simple point charge (SPC) water models can be used in place of more expensive quantum mechanical water molecules to efficiently model the solvent effect on a solute molecule's chiroptical responses. The effect that SPC waters have on the computed circular dichroism of a solvated glycine molecule are comparable to, albeit somewhat weaker than that of full fledged quantum mechanical waters at the coupled cluster CC2 level of theory. The effects of SPC waters in fact correlate better with QM-CC2 waters than quantum mechanical waters computed with density functional theory (DFT) methods, since they do not promote spurious charge transfer excitations that are a known deficiency with most popular density functionals. Furthermore, the near zero order scaling of point charge waters allows multiple layers of explicit solvation to be modeled with negligible computational cost, which is not practical with CC2 or DFT level. As a practical example, we model the molar rotations of glycine and alanine, and track their convergence.

INTRODUCTION

Molecular modeling of molecules in solution poses a continuing challenge. The amino acids which comprise our proteins are of great interest, and modeling them in solution phase is drawing increasing interest. Glycine, being the smallest of this class of molecules is a natural target for ab-initio computations. Many recent papers have been published on the solvation of this molecule, several of which we cite here.[39, 41, 113-122] Alanine, the second smallest amino acid and the smallest which is chiral, has also gathered significant attention.[120, 123-125]

A recent focus of our research has been modeling the chiroptical response properties of amino acids in solution.[20, 58, 81, 126] These interactions of chiral molecules with polarized light include specific rotation / molar rotation at fixed wavelengths, optical rotatory dispersion and circular dichroism.[15, 112] We, as well as others[8, 49, 120, 127] have employed continuum solvation models[40] in our earlier studies of these properties. Such methods, among the most popular being Tomasi's Polarizable Continuum Models (PCMs) and Klamt's COnductor like Screening MOdel (COSMO), do well to model bulk solvent effects. However they fare more poorly for short, explicit solvent – solute interactions such as hydrogen bonding, which is important for aqueous solutions of highly polar molecules such as glycine. For these, a more detailed solvent model is called for, such as one incorporating several solvent molecules explicitly. The exploration of such explicit solvation on chiroptical properties has only recently begun.[102, 128, 129] Here we continue this exploration with a look at how differing explicit solvation models effect the chiroptical response properties of the glycine and alanine molecules, and at the performance of simplified models that, unlike the continuum models, treat the solvent molecules as discrete entities. Replacing a continuum such as COMSO or PCM by discrete

solvent molecules comes at the added cost of averaging over many solvent – solute configurations. For this purpose, we employ molecular dynamics calculations.

This work begins with a look at how water molecules congregate around a glycine molecule that they are solvating, and how those solvent effects decrease as a function of solute – solvent distance with various model systems. It next probes more deeply into how the solvent molecules are perturbing the molar rotation and circular dichroism response of the glycine molecule, paying particular attention to how the solvent effects the lowest energy electronic excitation of the solute. It then explores the differences and similarities of using quantum mechanical water molecules and those comprised simply of point charges for solvation modeling, discussing the benefits and shortcomings of both methods. Finally the average molar rotations resulting from thousands of TDDFT calculations of glycine and alanine will be compared with experiment.

COMPUTATIONAL METHODS

Many of the computational methods used in this work are detailed in a previous publication,[20] where TDDFT based computations of optical rotations of amino acids were exhaustively benchmarked.[58] All quantum mechanical data were computed with the Turbomole[30] quantum chemical software, version 5.7.1. The calculations were performed with the Becke three parameter B3LYP and BHLYP hybrid functionals[31] as well as the CC2 coupled cluster method[130] as implemented in the Turbomole code. Dunning's aug-cc-pVDZ basis set[64] was used. For some calculations the COnductor-like Screening MOdel (COSMO)[33] of solvation was applied to the ground state. Molar rotations were calculated at the wavelength of the sodium D line (589.3nm). All molar rotations are reported in units of

deg·cm^2/(dmol). The center of the mass of the glycine molecule has been used for the coordinate origin for all response calculations. While results attributed to the length representation of the electronic dipole operator are formally origin dependent, this dependence is minimized in variational methods when large basis sets such as aug-cc-pVDZ are used; see our earlier work and the references cited therin.[58] Origin dependence however does not vanish upon basis set saturation for the CC2 method.[17, 131] However, the small difference in the results using the origin dependent length operator and the independent velocity operator did not effect our conclusions.

Geometries used in the quantum mechanical calculations were generated with the GROMACS[132] molecular dynamics program, version 3.3.3. Molecular dynamics of a glycine molecule were run in a cubic periodic solvent box measuring 25x25x25nm containing 509 Simple Point Charge (SPC) water molecules, resulting in a system with an average density of 1.0 g/cc. To avoid possible artifacts from the cubic shape of the periodic box, no water molecules greater than 12.5 nm from the center of the solute were included in the subsequent quantum mechanical calculations. The all-atom OPLS-AA[133] force field was used for the simulations, which were carried out at 300K with a time step of 1fs. Snapshots of the simulations for subsequent computations of Circular Dichroism (CD) and optical rotation were taken every 10ps, which was a sufficient duration for energetic and chiroptical response calculations of adjacent configurations to be uncorrelated.

RESULTS AND DISCUSSION

Glycine solvation as a function of distance

Explicit solvation at the quantum mechanical level entails significant computational costs. With this in mind, any development of such a solvation model should first consider how many explicit solvent molecules are needed in a system in order to affect full solvation on the solute. Fortunately, with respect to the solvation of the amino acid glycine, we can benefit from the experience gained in previous studies. Several articles have been published on the topic, much of which is summarized in the recent work of Aikens and Gordon.[114] They concluded, inter alia, that "Eight water molecules do not appear to fully solvate the glycine molecule." As we are interested in the effects that water molecules have on the chiroptical response properties of amino acids such as glycine, we must further investigate just how many of these molecules are needed to "fully solvate" glycine, so that we may have some idea of how many water molecules should be included in system that aims to model the full effects of explicit water molecules on this solute.

Before delving into the energetics of solvation, let us first consider the structure that the water molecules form around the glycine solute. To look at how solvent molecules orient themselves about the solute a radial distribution plot was created from 4096 frames of a glycine in water dynamics simulation. The results are depicted in figure 7-1.

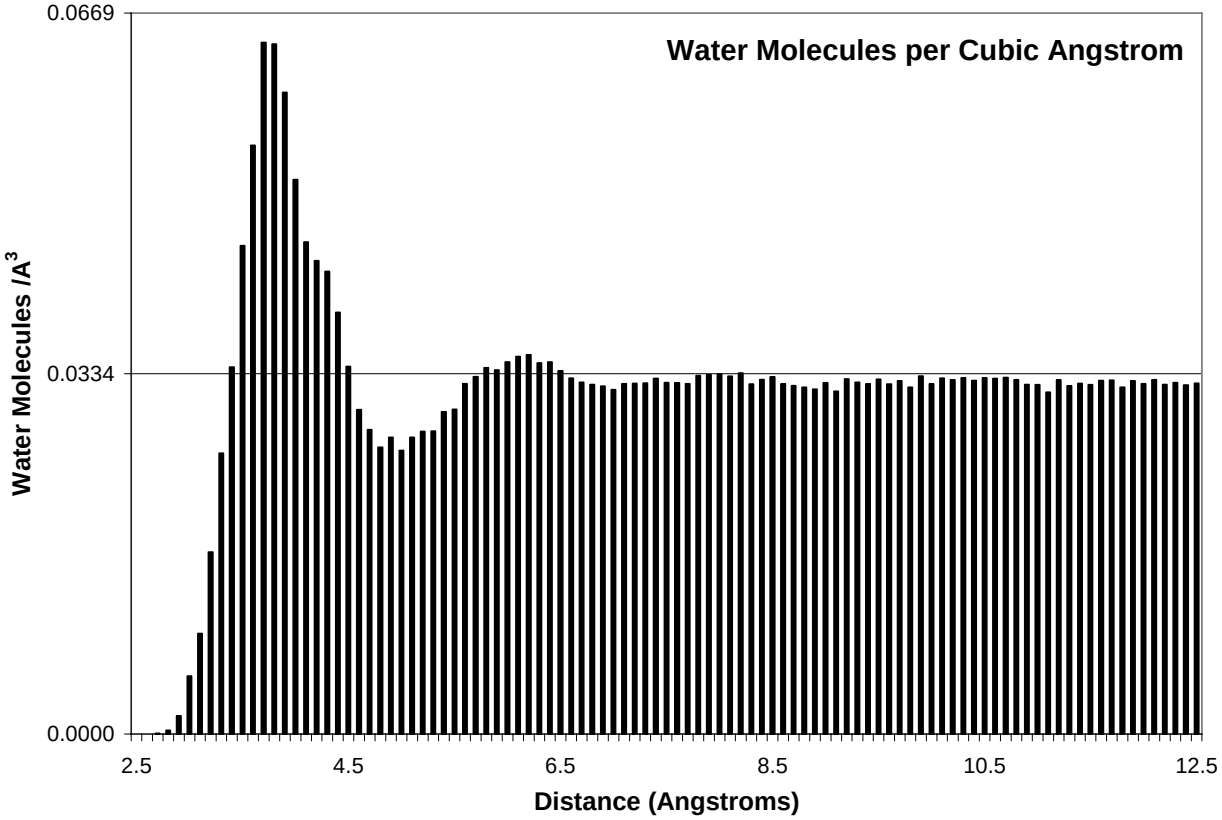
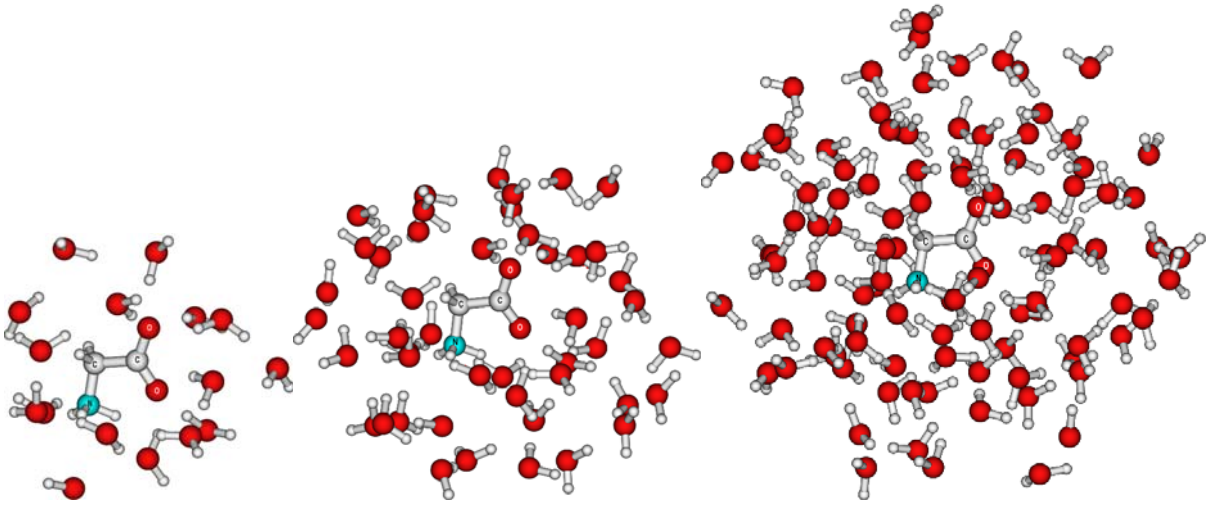


Figure 7-1: Molecular Dynamics simulation of glycine in water: Probability of finding a water molecule (per cubic angstrom) as a function of distance from the solute. Solute position is defined as the center of solute mass; solvent position is defined as the location of its oxygen nucleus. Radial shells with a width of 0.1 angstrom were used to generate the plot. For reference, pure water with a density of 1.00g/cc has 0.0334 water molecules per cubic angstrom.

From the plot it is apparent that the density of the solvent reaches a maximum around 3.7 angstroms from the solute center, and then dwindles to a minimum at 5.0 angstroms. Integration of the area between the solute and the minimum at 5.0 angstroms reveals approximately 16 water molecules are contained in this innermost solvation shell. Integration through the next minimum in solvent density at 7.0 angstroms adds another 30 water molecules in a second solvation sphere, giving a total of 46 waters needed to fill two solvation spheres. Beyond this the water appears to have little structure with respect to the solute, and maintains a density close to that of pure water. If we set a cutoff for a third solvation shell at 9.0 angstroms, it would mean

including approximately 100 water molecules total in the first three shells. A typical frame of a molecular dynamics simulation of glycine including one, two and three solvent shells of water is illustrated in figure 7-2.



glyz+16 H_2O glyz+46 H_2O glyz+100 H_2O

Figure 7-2: Representative configurations of a glycine molecule surrounded by 16, 46 and 100 water molecules which approximate one, two and three shells of solvation, respectively.

To have a uniform reference for the solvation energy, all energies for the interaction between the solvent and the solute were obtained from quantum mechanical calculations. In the molecular dynamics framework used here, such calculations are inherently more time consuming than the dynamics calculations used to generate the configurations, so the number of water – glycine interactions as well as the number of sampling frames had to be limited. Here we chose to look at the individual interactions of each of the nearest 256 water molecules with the glycine solute. Since this entailed performing that number of costly ab-initio calculations for every frame examined, only the first eight sampling frames of the dynamics simulation were used for this example. In the interest of possibly saving computational resources in the future, each water molecule was modeled once in the full quantum mechanical (QM) set, with a full compliment of electrons and basis functions, and once as a set of negative (oxygen) and positive (hydrogen)

point charges. The charges for this model were those from the 3 point TIP3P type waters.[24] For each calculation the interaction energy was calculated by subtracting the sum of the energy of the glycine molecule and the water molecule from that of the glycine – water system (no correction was made for basis set superposition error). Note that this is only really partial solvation energy, since only a single molecule of solvent is being considered. The results are depicted in Figure 7-3.

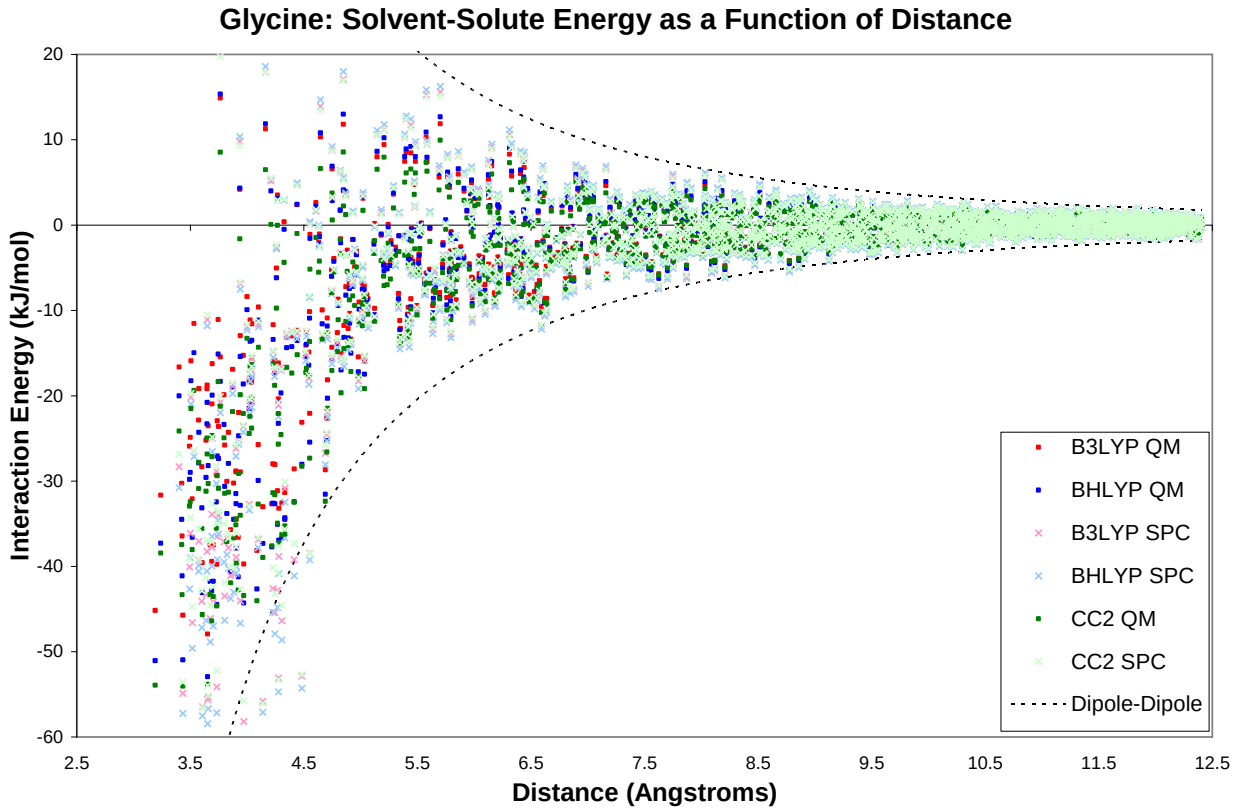


Figure 7-3: The energy of solvation caused by one water molecule upon glycine as a function of the distance between the oxygen atom of the water and the center of mass of the glycine. ($E_{interaction} = E_{solute+solvent} - E_{solute}$, with $E_{solute} = 0$ when SPC waters are used) The closest 256 water molecules considered at in each of eight frames of a molecular dynamics simulation. Calculations at the B3LYP, BHLYP and CC2 are shown in red, blue and green respectively. Full quantum water results are shown in dark colors, while those from SPC point charge waters are in lighter colors. The range of variation in energy expected from a dipole – dipole interaction is depicted by the black line, computed using a dipoles of 12 and 2.35 debye for glycine and water, respectively.

The energetic results are consistent with what is expected from the radial distribution of the water molecules. The water molecules in the innermost solvation shell, those within five angstroms of the solute center, are those with the greatest effect on the energy. This effect is largely negative, which indicates the stabilization of the highly polar glycine zwitterion by the polar water molecule. Note that this stabilizing effect is seen when point charge water molecules are used as well as when full QM waters are modeled. This indicates that the dipole – dipole interaction, which can be effectively modeled by the SPC waters, plays a prominent role in the solvation.

For waters farther from the solute, the solvation effect is lessened. The energetic effect of waters in the second solvation shell, those between approximately 5 and 7 angstroms, is far smaller, on average less than half the magnitude of those in the inner shell. Beyond that distance solvation effects dwindle further. At 9 angstroms from the solute center, about where the 100^{th} water is located the effect upon solvation has become negligible compared to the level of accuracy inherent in the ab initio calculations.

Glycine: Solvent effects on molar rotation

The recent focus of our research has been the modeling of the chiroptical properties, including molar rotation, of amino acids in solution. Incorrect modeling of the Boltzmann populations of amino acid conformers, caused in part by an insufficient solvent system has been considered as one possible source of error. This is especially true since these conformers can have both large positive and negative rotations, which should match experimental results only upon calculation of a correct weighted average of their responses. Another important part of this current model, which was absent in earlier works, is the direct effect that explicit solvent

molecules have on optical rotation. It is important to study there effects separately, which makes the glycine and alanine zwitterions attractive candidates for study since each has only one local minimum structure.[58] Glycine draws particular interest not only because of its low molecular weight, but also since it is symmetrical, making it a good test case as a system which can exhibit a chiroptical response depending on its configuration, responses which we know must average to zero over time.

To investigate the aspects of explicit solvation a linear response calculation of molar rotation at 589.3nm of a glycine – water system was performed on each of the frames considered in the previous section. Again both full QM and SPC waters were used, and the B3LYP and BHLYP hybrid functionals were considered. The CC2 method was not used here because of the lack of a CC2 implementation for optical rotation in the software used for this work.

While the glycine molecule is itself an achiral molecule, a snapshot of a solvated glycine can exhibit a chiroptical response for a combination of two reasons: First, the glycine molecule itself can be found in geometries where its plane of symmetry is broken. It has an equal probability of being captured in both dextrorotatory and levorotatory conformations. For the limited sample size of eight configurations, we were fortunate to find that the unsolvated glycine molecule had a positive molar rotation exactly four times, and an equal number of negative molar rotations. The second reason that a solvated glycine molecule may exhibit a non-zero molar rotation is the asymmetrical orientation of a water molecule with respect to the glycine. This is the phenomenon of interest; in order to isolate this solvent effect on molar rotation response calculations were carried out glycine both with and without the presence of a water molecule and subtracting the two molar rotations. (If the water molecule could adopt a chiral configuration, which it can in the presence of other water molecules[134], we would have to

consider its molar rotation as well; however an isolated water molecule has only three atoms which are inherently coplanar, and thus can have no intrinsic chirality regardless of how its geometry is distorted by vibration.) The results of these calculations are shown in Figure 7-4.

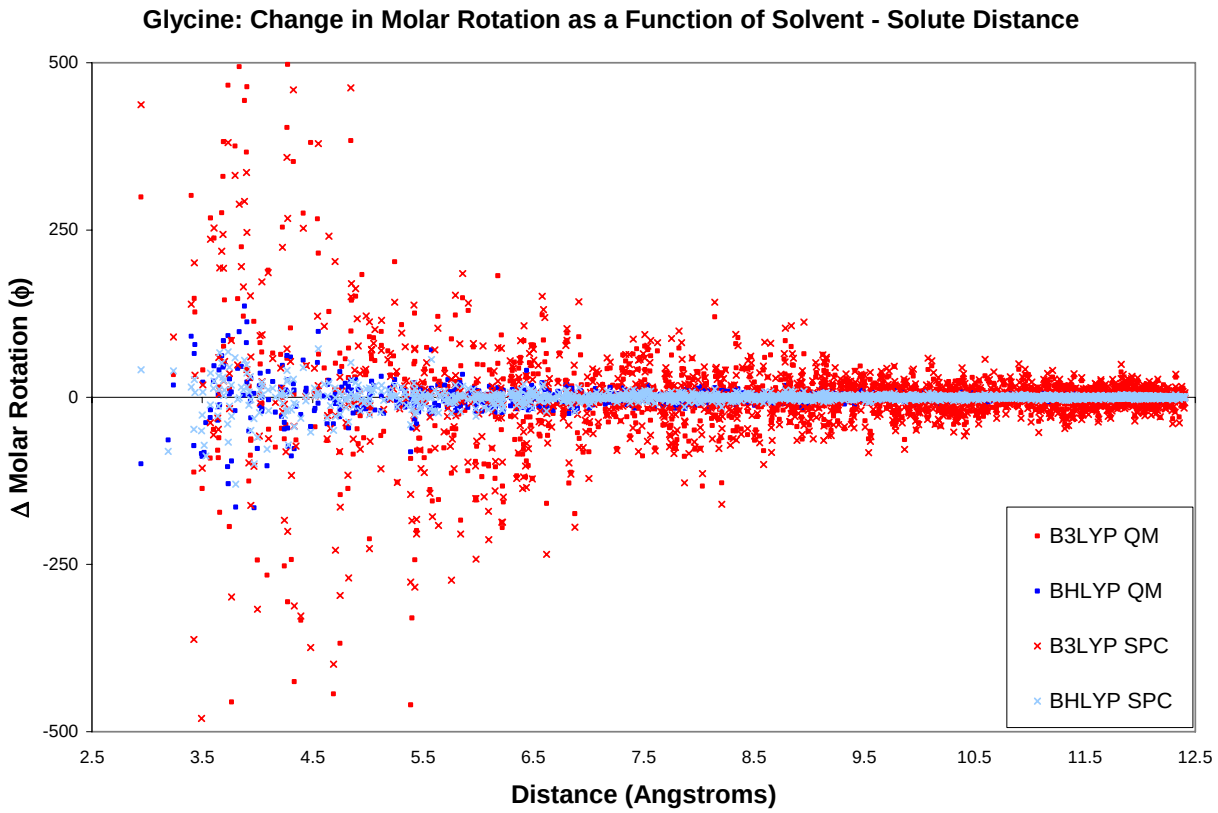


Figure 7-4: The change in molar rotation (ϕ) caused by a single water molecule as a function of its distance from the glycine center. ($\Delta\phi=\phi_{solute+solvent} - \phi_{solute}$ with $\phi_{solute} = 0$ when SPC waters are used) Calculations were performed with the aug-cc-pVDZ basis at the B3LYP (red) and BHLYP (blue) levels of theory using QM (dark colored) and SPC (light colored) water molecules.

As expected, a water molecule has a pronounced effect on the molar rotation of a glycine – water system, and the magnitude of that effect varies inversely with the distance between the solvent and solute molecules. The magnitude of this change in molar rotation remains rather consistent regardless of whether a point charge or quantum mechanical water causes the perturbation. A drastic change however can be seen when the amount of exact exchange in the hybrid DFT functionals is varied. The change in molar rotation caused by solvation is much

greater in the B3LYP hybrid than in the BHLYP; the cause of this change will be discussed in the next section. The magnitude of the solvent effect on the molar rotation illustrates the challenge of successfully averaging this quantity over the course of a molecular dynamics simulation in order to minimize the statistical error.

Effects on First Circular Dichroism Excitation

The direct linear response method for calculating molar rotation used in the previous section performs the task very efficiently. However, it tells us little about *why* a molar rotation is what it is. Here we are interested in what electronic transitions are responsible for the molar rotation. This can be done by computing molar rotation via the sum over states method. In a prior study on aliphatic amino acids such as glycine, we confirmed that the lowest lying electronic excitation, the n to π^* transition of the carboxylate moiety, is largely responsible for the molar rotation observed at 589.3nm.[126] As we are interested now in the effect an explicit solvent molecule has on this molar rotation, it is therefore beneficial to consider the effect this solvation has on the critical first electronic CD.

Using the same configurations as in the previous section, calculations of the rotatory strength and wavelength of the first CD excitation were performed at the B3LYP, BHLYP, and CC2 levels of theory using QM and SPC water molecules. In each case the change of the CD response caused by the solvent molecule was computed by subtracting from the CD of each water – glycine system the CD caused by an unsolvated glycine molecule in the same geometry. The lowest electronic excitation of water is far higher in energy than that of the glycine, and so the first CD transition that is observed in the model is always centered on the solute. The solvent serves merely to perturb that transition within the solute. The results of these calculations are plotted in Figure 7-5.

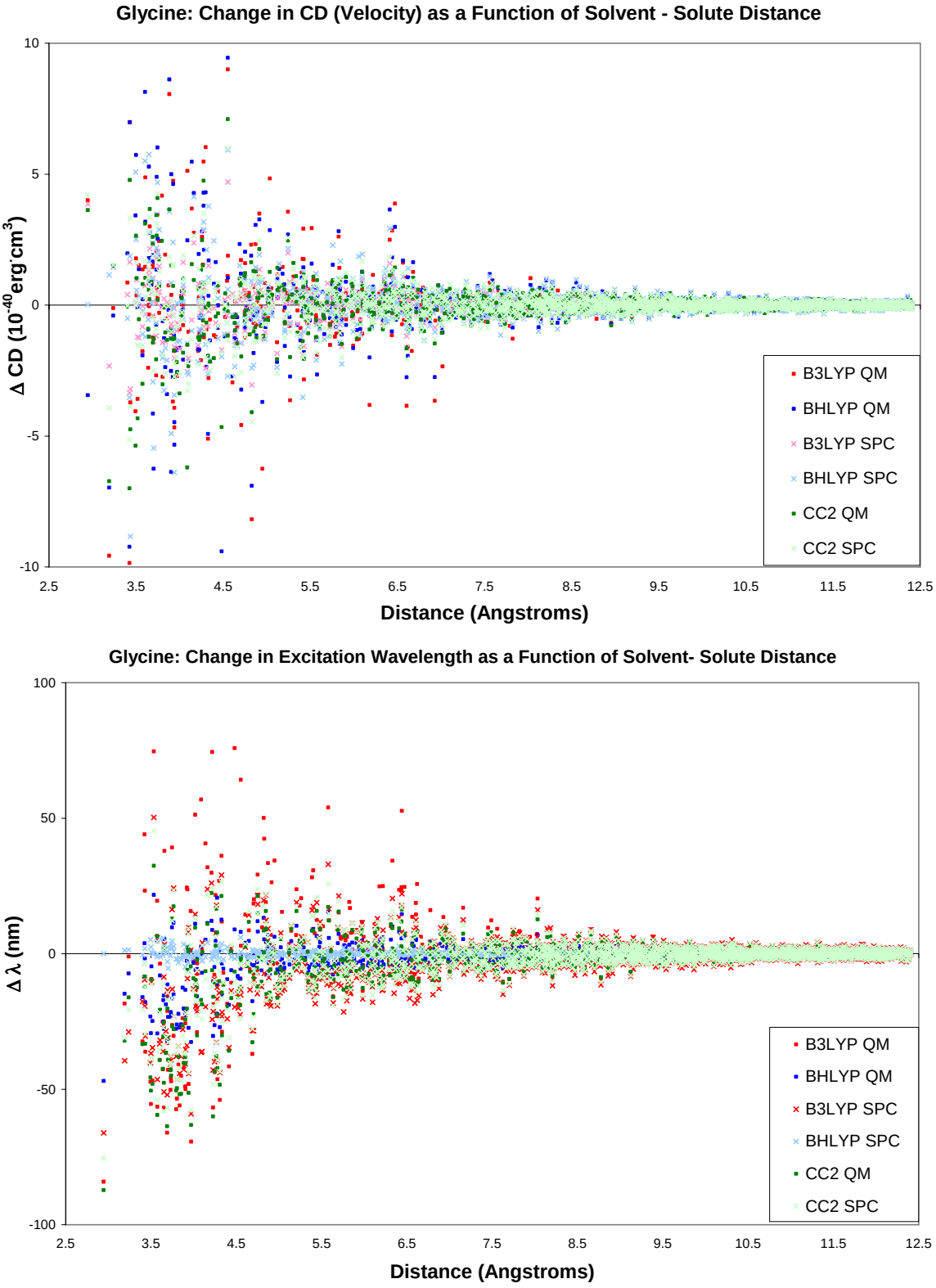


Figure 7-5: The change in Circular Dichorism of the first electronic transition as a function of water distance from the solute center. Change in ellipticity (top) and wavelength (bottom) are plotted.

Just as with the energy and molar rotation, the effect of a water molecule on the first CD excitation decreases with distance, as expected. The change in ellipticity caused by the water has an equal probability of being positive or negative, as one would expect for a chiral molecule being perturbed by an achiral solvent. The change in the wavelength of that transition however tends to be negative at all levels of theory. That is: a nearby water molecule tends to induce a blue shift of the first CD transition.

This blue shift is far more pronounced at the B3LYP and CC2 levels of theory than with the BHLYP functional. In the gas phase model, this excitation takes place at 430, 393 and 287 nm, respectively. These are all far too low in energy compared to solution phase measurements of aliphatic amino acid CD, which indicate that this dichroism has a maximum in the range of 200 to 215 nm.[74, 135] Solvation increases this excitation energy and thus lowers its wavelength at all levels of theory, more so for B3LYP and CC2 than BHLYP. The contribution of a CD excitation to the molar rotation depends inversely on the difference between the wavelength at which the molar rotation is computed or measured (in this case that of the sodium D-line, 589.3nm) and the wavelength at which the excitation occurs, per the Kramers-Kronig relationship. This leads to the explanation of why solvation has a more pronounced effect on the molar rotation of the glycine molecule with the B3LYP hybrid functional than with the BHLYP. Solvation shifts the first electronic excitation wavelength farther from the Sodium D line to a greater extent with the B3LYP functional than with the BHLYP, where it was already much farther away to begin with.

The correlation between differing levels of theory

When developing a model chemistry with efficiency in mind, it is helpful to compare the result computed at various methods with those calculated at the most robust level of theory available on a relatively simple system. Here the system of glycine and one water molecule has served as a microcosm of a glycine molecule solvated by multiple water molecules. We have at our disposal to perform multiple calculations on this system at a correlated wavefunction based level of theory (CC2/aug-cc-pVDZ). However as more waters are added, performing the coupled cluster calculation on an increasingly large system will become impractical and eventually impossible, due to the scaling of this method with respect to system size. This is illustrated quite clearly in figure 7-6, where CD calculations (first electronic excitation only) on a system with a glycine and 7 water molecules take over one day *per configuration*, and hundreds of such configurations are required to achieve an averaged result that is a reasonable representation of the dynamic system. With larger systems that include more solvent molecules, modeling with more efficiently scaling methods such as density functional theory, or the "zero order scaling" point charge waters becomes a necessity. Employing these point charge waters such calculations take approximately 22 and 8 *minutes* per configuration at the CC2 and DFT methods, respectively, regardless of how many waters are included. As such, it makes sense to take a close look at how closely more efficient model chemistries compare with the coupled cluster model.

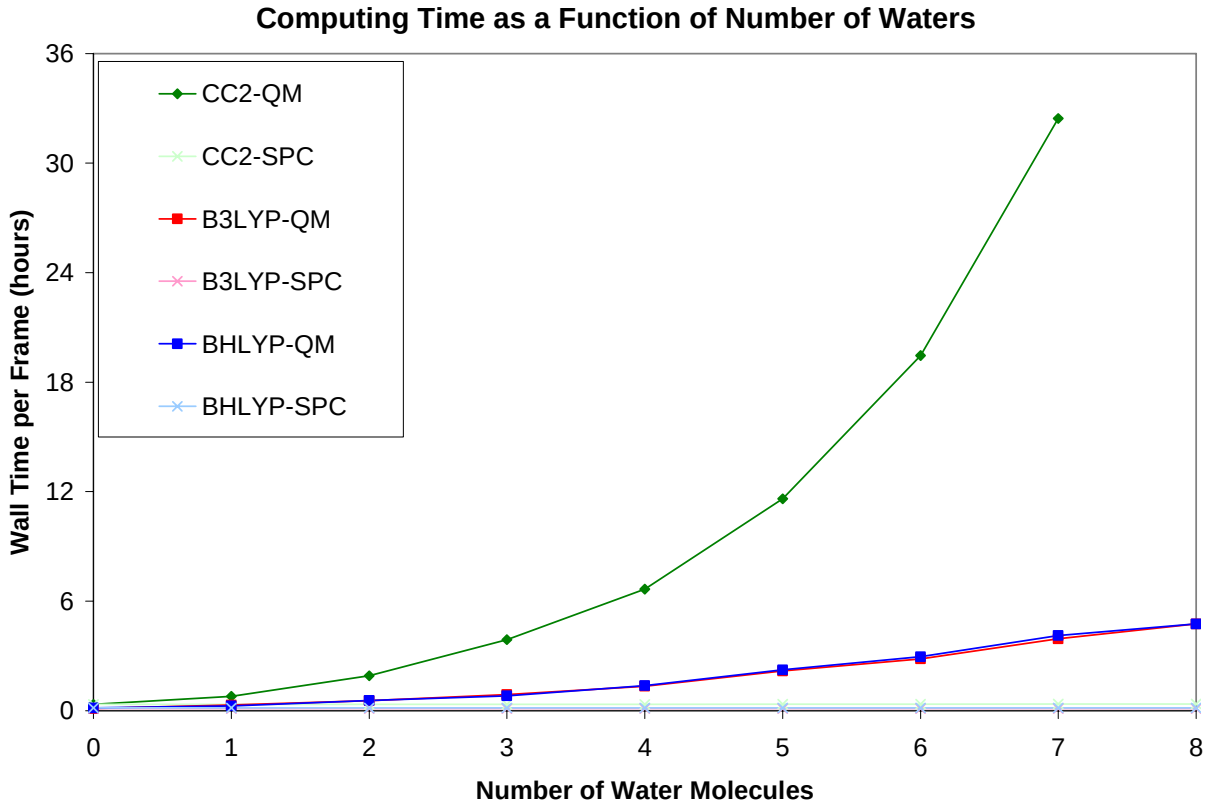


Figure 7-6: Time required to complete a CD calculation on a single frame of a glycine – water system as a function of the number of water molecules. Numbers reported are average values from 128 frame simulations. One core of an AMD 2.2 GHz 64-bit dual core Opteron processor was used for each calculation,

In order to do this, we plotted to correlation of the change in the lowest ECD excitation (both rotatory strength and wavelength) caused by various types of water molecules versus those changes caused by a QM CC2 water. The correlation of the change in wavelength is depicted graphically in the top of figure 7-7. The correlation of change in rotatory strength is shown in the bottom graph. Each data point corresponds to a distinct frame of a glycine in water molecular dynamics simulation; 128 frames were considered. For each frame the glycine molecule was modeled at the QM level, and which ever single water molecule that happened to be closest to the glycine center at that point in time was included as well, either at the QM level or as a set of point charges.

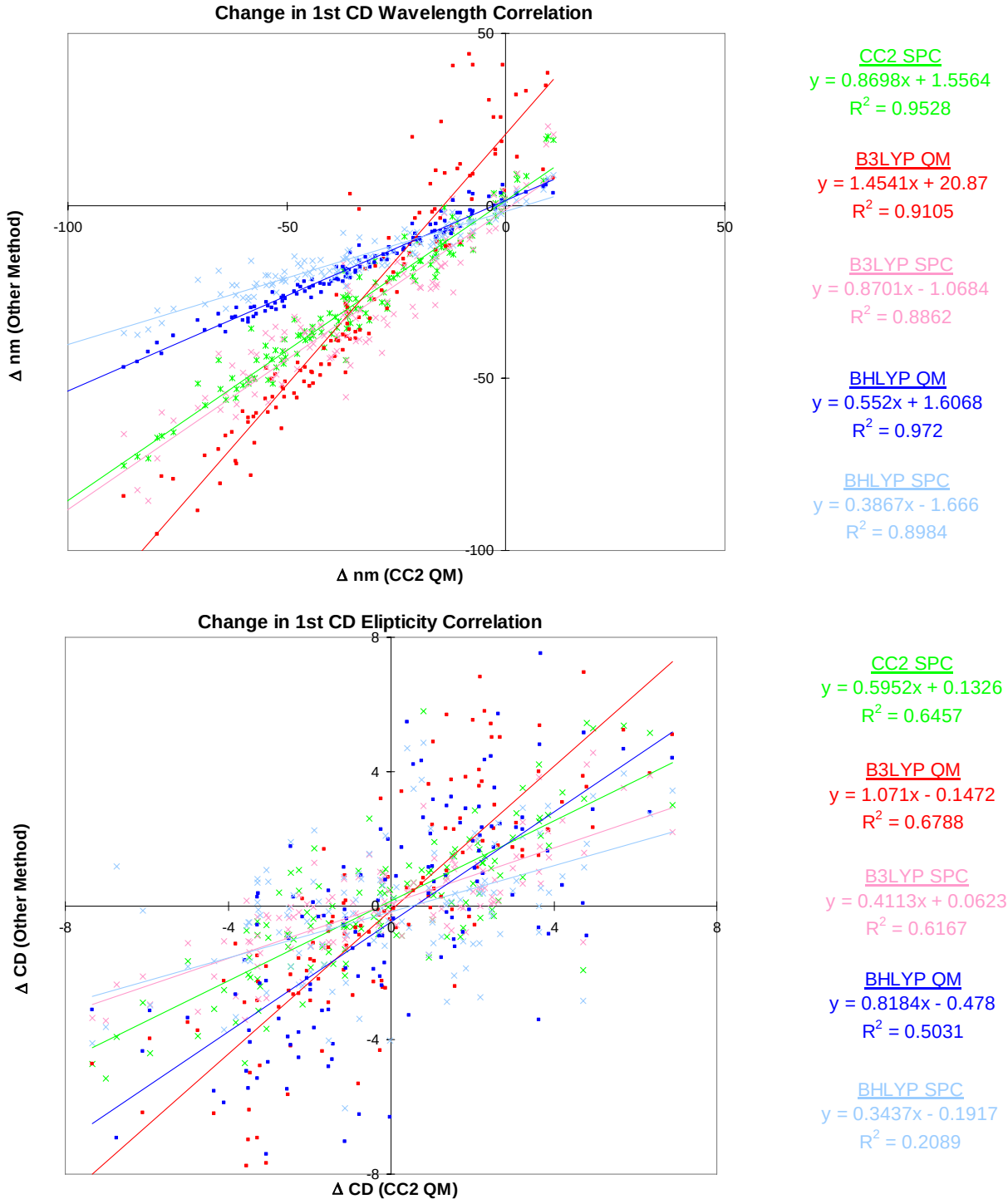


Figure 7-7: Correlation of ellipticity (top) and wavelength (bottom) of the first excitation caused by the closest water molecule to the glycine solute at various levels of theory with full CC2/aug-cc-pVDZ. 128 configurations were considered.

The perturbations to the first CD excitation of glycine caused by water molecules modeled with all of the less costly methods correlate positively with that caused by a quantum water included at the coupled cluster level. For all of the methods except QM B3LYP, the slope of the correlation in wavelength is less than one, which indicates that the change in excitation wavelength, nearly always a blue shift, caused by the water is smaller in magnitude for the other methods than for full QM CC2. The SPC-CC2 method shows arguably the best correlation with QM-CC2, with a regression line slope of .87 and an R^2 value over .95; the SPC-B3LYP also correlates well, with a slope of .87 and an R^2 of nearly .87. The BHLYP methods, both full QM and SPC, tended to have their CD less affected by water than the other methods, principally since the first excitation with this DFT hybrid is already significantly blue shifted compared to those calculated with the other methods. Note the correlation that is examined closely in figure 7-7 is only that of the water molecule that is closest to the amino acid center in each configuration; for water molecules that are farther away, as was already shown in figure 7-5, the correlation between methods appears to improve significantly.

As noted the QM-B3LYP method seems to have the poorest correlation with the QM-CC2 method. It is the only method that shows a change in excitation wavelength that tends to be greater in magnitude than for QM-CC2. Furthermore, it is the only method that gives a significant number of red shifts to the first excitation for water – glycine configurations where the QM-CC2 method yields blue shifts. Adding a water molecule to a glycine zwitterion should induce a blue shift in its first excitation, due to the stabilization of the glycine electronic state by the solvent. Density functional theory, however, is known to have an issue with producing charge transfer excitations which are often unphysically low in energy. This deficiency, along with its particular consequences in supermolecular solvation simulations such as this one has

been discussed in detail recently by Neugebauer and coworkers.[136] Various methods are under development to compensate for this shortcoming.[137-139] This charge transfer problem can be ameliorated somewhat by using hybrid DFT functionals with a greater portion of exact exchange[140], such as BHLYP. But with B3LYP, the combination of a QM solute and a QM solvent sets up a scenario where such an charge transfer excitation can take place between solute and solvent or even between two solvent molecules. As indicated by the red data points in the top of figure 7-5, it happens frequently enough at this level of theory to be an issue. With point charge waters this is obviously not an issue, since they have no electrons or orbitals to participate in such a non-physical electron exchange. Thus, with the B3LYP functional the more simplistic point charge waters appear to better model the solvent perturbation of the first electronic excitation of glycine than more costly QM waters do.

The magnitude of the charge transfer problem is illustrated more dramatically in figure 7-8. As one, two, three and more waters are added to a water glycine system, we expect to see an initial decrease in the wavelength of the lowest electronic transition as the greater number of water molecules stabilize the highly polar glycine zwitterion and increase its HOMO-LUMO gap, with the effect leveling out as the number of water molecules increases. Point charge waters do not cause such unphysical results, and the longest wavelengths computed with these methods converge to slightly shorter wavelengths than those obtained with corresponding continuum based (COSMO) methods. The coupled cluster method does not suffer from the charge transfer problem, so the changes in wavelength at this method is quite similar for QM and SPC waters, regardless of how many are added. BHLYP appears to perform fine as well, though at around eight waters figure 7-8 indicates that the charge transfer problem may be beginning to show itself with this QM method as well; using such a hybrid functional with a large portion of exact

exchange has indeed compensated for part of the charge transfer excitation problem but has not eliminated it completely.

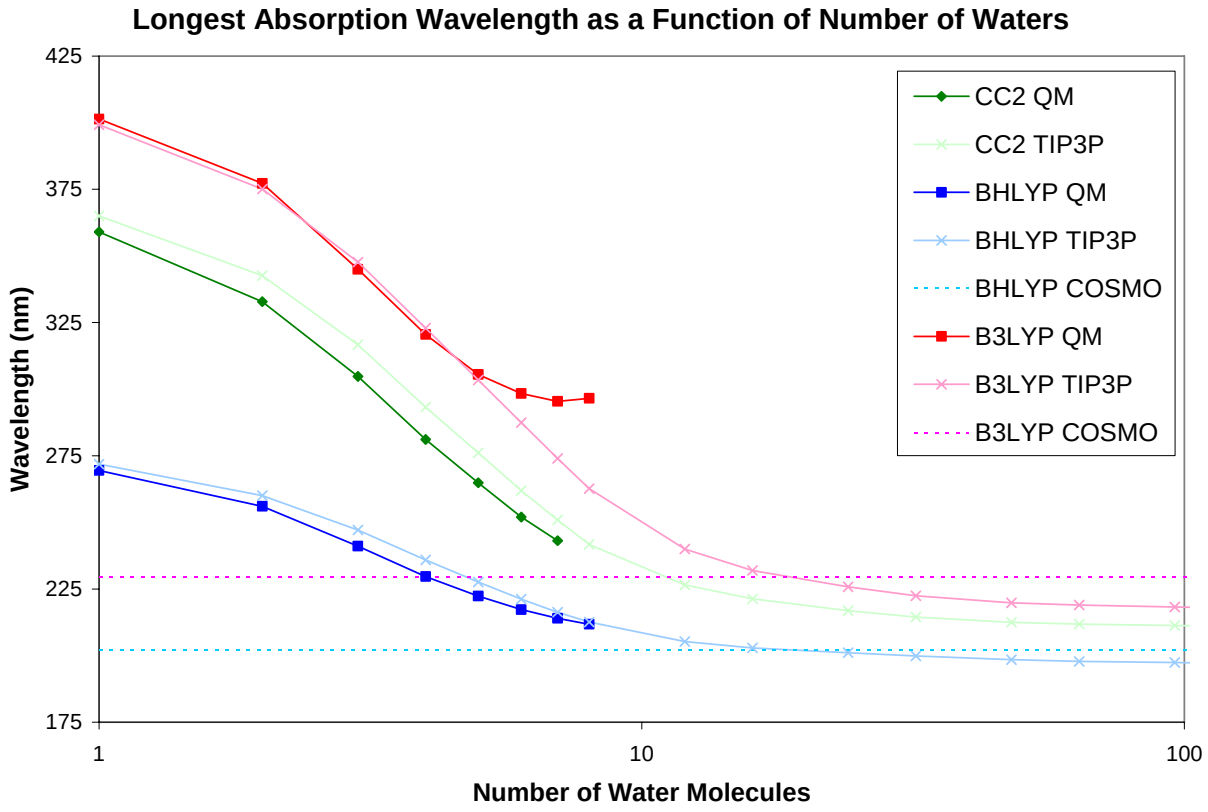


Figure 7-8: The effects of adding multiple water molecules on the 1st excitation wavelength of the glycine-water system. The results represent the weighted average of 128 configurations of the dynamics simulation of the longest absorption wavelength of various glycine – water systems.

As for the change in ellipticity, as with the change in wavelength, there is a positive correlation amongst all of the solvation methods. At the QM-B3LYP level, the 1st CD tends to be larger in magnitude than at QM-CC2, whereas at the QM-BHLYP level, it tends to be smaller. This is in keeping with the excitation energies: the CC2 results tend to fall between those obtained with B3LYP and BHLYP. With point charge waters we consistently see perturbations that are too weak compared to QM-CC2. For example, using the SPC charge waters we see a perturbation to the CD excitation of about 60-80% of that found with full QM CC2 water. This

indicates the limits of the point charge water model; it simply cannot reproduce all of the interactions that take place between a water molecule and the solute such as those involving orbital overlap, polarization[141], and quadrupole and higher order multipole interactions. They can however reproduce the dipole moment of the water molecule, and judging by the slope of the regression lines this seems to be the most important interaction.

Another issue of note is the R^2 values of the regression lines. This indication of correlation is rather poor, particularly between the intensities with the BHLYP hybrid and those at QM-CC2. We can see on the bottom graph of figure 7-7 many instances where the sign of the CD does not agree between these methods. We have reason to believe that this may be caused by insufficient modeling of electron correlation. In an earlier work, we found that increasing the amount of exact exchange in a hybrid DFT functional eventually resulted in the wrong sign of the CD of the first excitation being modeled.[58] For this modeling of one configuration CD of the alanine zwitterion, this sign change occurred at some point between the BHLYP and HF levels of theory, i.e. where between 50% and 100% exact exchange were used.

Mixed SPC / QM solvation

Thus far, we have only considered the solvation of glycine by point charge or quantum mechanical water molecules, but not both at the same time. However a model can be devised in which the water molecules closest to the solute are computed at the QM level, while those farther away are simultaneously considered as sets of simple point charges. To evaluate the merits of such methods, we performed calculations of a system in which glycine is solvated by 256 point charge water molecules. These results were compared to computations in which glycine is solvated by *n* QM waters, where n=1 to 8, and *256-n* SPC water molecules. Waters were

designated as QM based on their proximity to the solute center, with the closest being the first considered as QM, the next closest the second and so on. The results are depicted in figure 7-9.

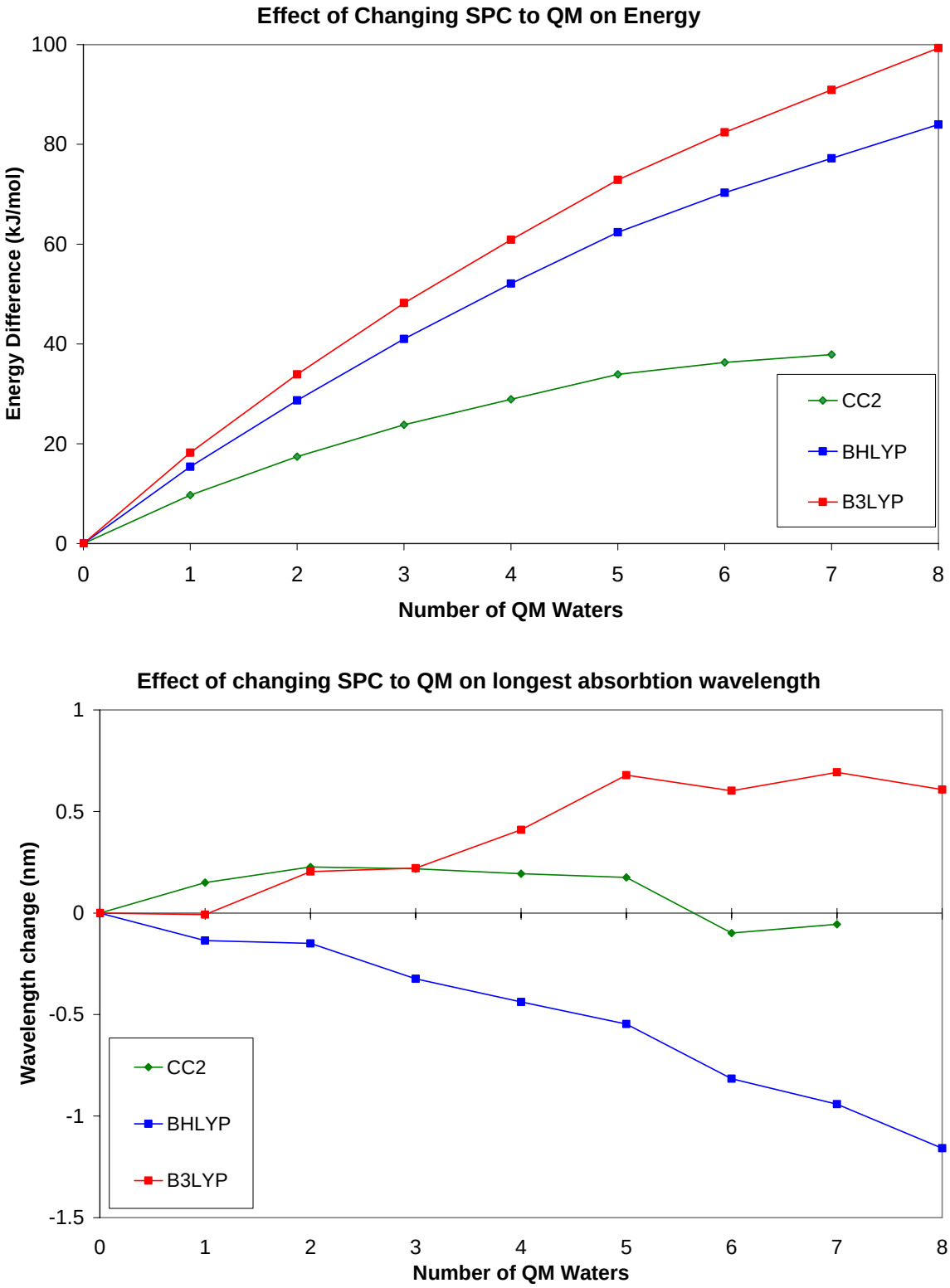


Figure 7-9: The difference in partial solvation energy and longest absorption wavelength of a glycine molecule solvated by 256 point charge water molecules and one solvated by *n* QM waters and *256-n* point charge waters.

The discrepancies in close-distance solvation energies between point charge and quantum waters that was first noted in figure 7-3 is even more apparent in the top of figure 7-9. The water molecules considered in figure 7-9 are all reside in the first solvation shell, where they are apt to be in close contact with the solute, so we expect to see the most pronounced differences here. For the waters in closest contact with the glycine we see that the point charge waters result in a greater stabilization than do the QM waters, which is not unexpected since point charge waters do not take into account somel types of interactions, such as steric repulsion, which are more prominent at short distances. The point charge waters are far more similar to their corresponding QM waters with the CC2 method than with the DFT, though such differences are still significant. However we can see at all levels of theory the significance of using a QM water as opposed to a point charge water diminishes as the waters in question get progressively farther from the solute center.

The effect of using QM as opposed to SPC water molecules for the innermost solvation is far less noticeable on the first excitation wavelength. As is shown in the bottom of figure 7-9, switching from point charge to QM water molecules for even the innermost 8 water molecules of the 256 molecule solvation sphere results in a maximum change of around 1 nanometer in wavelength. At the CC2 level of theory, which should be the most reliable, this difference between QM and SPC waters is quite negligible, ~0.2nm. This indicates that even for the closest held waters of the solvation sphere, a simple point charge model is quite capable to model the solvent effects on this electronic excitation.

Glycine and Alanine: Comparing molar rotations with experiment

This work so far has focused on the chiroptical response properties of glycine, a molecule that may serve as a means for calibrating our method, but which is itself achiral. This serves the purpose of a "blank" for our molecular dynamics method of modeling molar rotation. The molar rotation of glycine must average to zero, and so whatever residual molar rotation that does not average out after computations on multiple frames of the glycine in water system will indicate some idea of how much statistical error is inherit in the method. It seems prudent here to also take this method and extend its application to a similar chiral amino acid, one whose chiroptical response properties ought not to average to zero over time and whose chirality, as exhibited in its molar rotation, can be modeled and compared to that of experiment. Alanine, the smallest *chiral* amino acid, serves as a natural target for such an investigation.

In this section 40960ps molecular dynamics runs of glycine and alanine were performed under the same conditions as the glycine simulations in the foregoing sections. A total of 4096 equally spaced configurations were taken for subsequent molar rotation calculations. Both the B3LYP and BHLYP hybrid DFT methods were employed, using either our simple point charge waters or the COMSO continuum model that we used in prior works.[20, 58, 126] For each selected configuration along the molecular dynamics trajectories a molar rotation was computed, and that rotation was appended to a running average of molar rotations. Two examples, alanine in SPC waters with the BHLYP and B3LYP functionals are shown in figure 7-10. Convergence criteria are based upon the principles of signal averaging, under which the standard deviation of a data set should drop as $n^{1/2}$ per number of points averaged, n, if the points are uncorrelated.[142] We confirmed that for these systems this is the case when we allowed 10ps to elapse between snapshots of the molecular dynamics run.

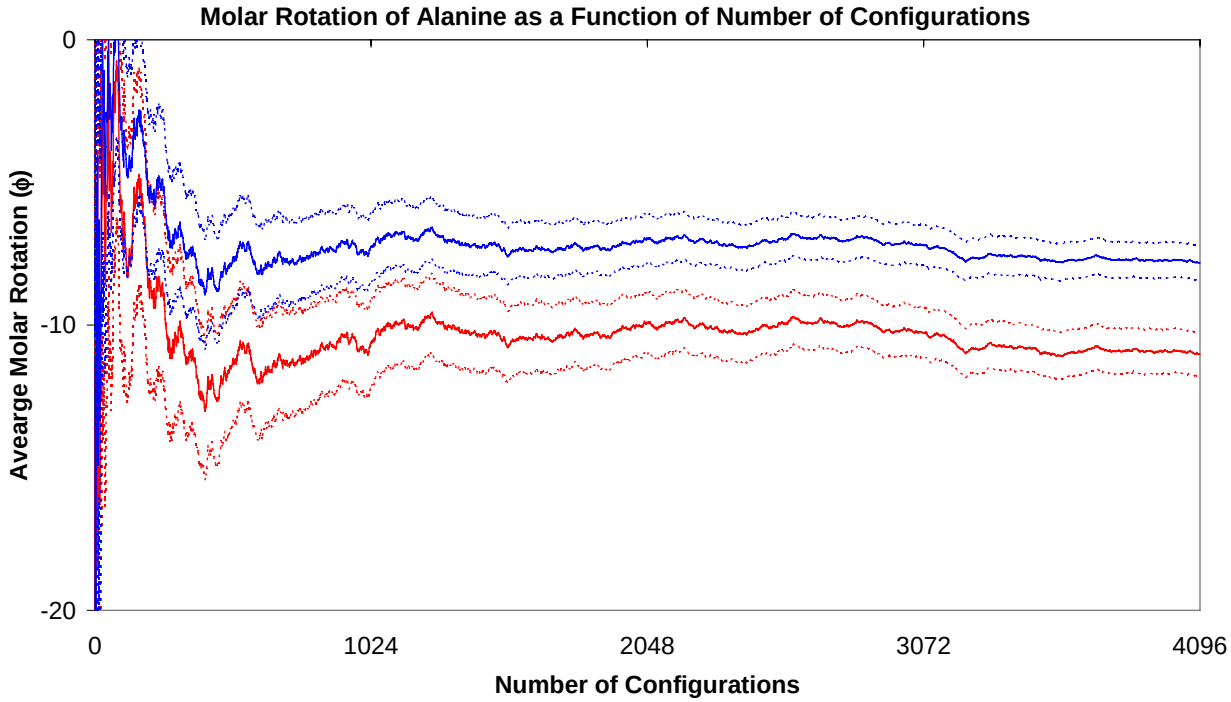


Figure 7-10: The convergence of the molar rotation of alanine as the number of averaged configurations increases. The BHLYP/SPC (blue) and B3LYP/SPC (red) methods were used. The running averages are represented by solid lines, while the error bars, at one standard deviation, are represented by dashed lines.

These results are compared with those obtained using static alanine molecules and the COSMO model, as well as with those from experiment. These data, and data from calculations on glycine are summarized in table 7-1.

Table 7-1: The molar rotation ([ϕ], deg·cm^2/dmol) of glycine and alanine with various methods. Experimental value for alanine is from Greenstein and Winitz.[72] Calculations with static molecules were done by the same method used in our prior work.[126] Error bars represent one standard deviation of statistical error from the dynamics simulation.

			Static - COSMO		Dynamic - COSMO		Dynamic - SPC	
Molecule		Expt.	B3LYP	BHLYP	B3LYP	BHLYP	B3LYP	BHLYP
glycine	ϕ	0	0.2	0.1	-0.9	-0.5	-1.0	-0.8
	+/-				1.4	0.9	1.1	0.8
alanine	ϕ	1.6	5.4	-1.5	-16.7	-10.9	-11.0	-7.8
	+/-				0.8	0.6	0.8	0.6

The data above show that the molecular dynamics / point charge solvation method produces results comparable to earlier results with frozen solute molecules and a continuum

solvent. All results are reasonably close to experiment, and well within the margin of error typical of TDDFT based calculations.[29] Error bars tended to be larger with the COSMO solvent model than with the discrete SPC waters, and higher for the B3LYP hybrid than for the BHLYP. This is due to the relatively lower excitation energies obtained with COSMO and with B3LYP, which result in larger magnitudes of the computed molar rotations.

In addition to comparing our results to experiment, we should also mention the results of D'Abramo et al, who performed dynamics based computations of alanine in water using TDDFT and a "perturbed matrix method" of solvation.[124] Using this method they obtained a specific rotation of about +60 deg·cm^3/(g·dm) for alanine, which corresponds to a molar rotation of +53 deg·cm^2/dmol. It would appear that our method is significantly closer to experiment. However, their rotatory dispersion calculations were performed with a truncated sum-over-states method. In contrast, the linear response method used here does not possess such truncation errors.[82] As such, a direct comparison between results with our point charge solvation model and their perturbed matrix model is not currently possible, though the disagreement is likely due to the truncation error from the sum over states calculation of alanine.[126]

CONCLUSIONS

Simple point charge (SPC) water molecules have been shown to be a computationally efficient alternative to using quantum mechanical waters in modeling the solvent effect on a solute's chiroptical responses. The near zero order scaling of point charge waters allows hundreds of explicit water molecules to be considered at negligible computational cost. The effect that SPC waters have on the computed chiroptical properties of a solvated glycine molecule are comparable to those obtained with the much more expensive CC2 method. When

density functional theory is employed, point charge waters may prove superior to explicit QM waters in simulating solvent effects on such response properties, since the point charge model does not exacerbate the problem of DFT with spurious charge transfer excitations. Calculations on the chiral amino acid alanine with the SPC/DFT method yield results that are in reasonably good agreement with experiment, and marginally better than those with the same geometries and the COSMO/DFT method. Further benchmarking of the point charge / DFT protocol is presently underway.

ACKNOWLEDGMENT

The authors would like to acknowledge support from the Center for Computational Research (CCR) at the University at Buffalo for computational resources. M.D.K wishes to thank Marcel Swart, Marek Freindorf and Mark Rudolph for helpful discussions regarding molecular dynamics programs. J.A. is grateful for financial support from the CAREER program of the National Science Foundation (CHE-0447321).

CHAPTER 8: SUMMARY AND OUTLOOK

My first paper at UB, contained in chapter three, was my first attempt at modeling specific rotation. For these molecules the sign of the computed specific rotation was in agreement with experiment, and the magnitude was consistently somewhat high, consistent with a known shortcoming with TDDFT. It was particularly gratifying that the specific rotation of the proline anion was modeled with that same level of accuracy as its cation and zwitterion forms, since previous authors had concluded that TDDFT was totally inadequate for modeling such anions.

The second paper that I contributed to at this university involved an extension of the modeling the ORD of the proline cation. The solution to this problem was to calculate the circular dichroism spectrum of proline, apply an empirical broadening term, and then use the Kramers-Kronig transformation to obtain an ORD curve without singularities. This method was used to successfully model the dispersion of proline through its first Cotton Effect. The effect of truncation error was also discussed, and it is for this reason that the method used was adequate for modeling the ORD in the absorbing region, but for the lower energy transparent region direct linear response is the preferred method.

Chapter five, which is my third paper at the University at Buffalo, took on the challenge of modeling the dispersive chiroptical properties of the aromatic amino acids. Significant knowledge about the origin of the chiroptical response was obtained in this work. An analysis of the specific rotations of the differing conformers of the aromatic acids led to the conclusion that for tryptophan the contribution from the aromatic chromophore was greater than that of the carboxylate chromophore. For tyrosine and histidine the contribution of the aromatic chromophore was less, due to a greater symmetry of the chromophore and a higher energy of the

$\pi \rightarrow \pi^*$ transition, respectively. The contribution of the benzyl chromophore of phenylalanine, while expected to be small, could not be effectively calculated by the method used since it depends heavily on vibronic coupling not modeled by the computational method used.

Chapter six, my fourth paper published as part of my PhD work appeared in the Journal of the American Chemical Society this year. My presentation on the effect was also selected for the Sci-Mix section of the 2006 San Francisco ACS national convention, another indicator of its acceptance by a broad scientific audience. This was a thorough investigation of the historical Clough-Lutz-Jirgensens rule for amino acids, which states that the molar rotation of a solution of a natural alpha amino acid is changed in the positive direction by addition of a strong acid. The decades old rationale for this rule was made by assigning circular dichroism octants for overlapping carbonyl n to pi* transitions, based on the old sector rules for circular dichroism. Modern time dependent density functional theory allowed me to take a new look at this phenomenon. I devised a method for mapping out these "sectors" using TDDFT. This same quantum mechanical was then used to show how the change in charge of the molecule upon protonation and deprotonation is responsible for reversing these sectors by changing the sign of the critical n to pi* transition in the carboxylate / carboxylic acid chromophore.

The CLJ effect is correctly obtained from the computations for all 12 amino acids studied in this work. This is despite the fact that the absolute values of the molar rotations for the corresponding zwitterionic and cationic forms of these amino acids were *not* all modeled correctly. This suggests a new and possibly useful way to use TDDFT to assign absolute configuration, based not just on the molar rotation of a molecule under on set of conditions, but on the *change* in molar rotation as the conditions of the molecule are altered, in this case by the protonation / deprotonation induced by a change in pH.

In chapter seven, point charge based water molecules were shown to be an effective substitute for computationally more expensive quantum mechanical based waters. Additionally, these simple water models had the advantage of not exacerbating TDDFT's spurious charge transfer problem. The potential use of non-QM solvent molecules for modeling the solvent effects on QM solute molecules has been demonstrated on the small amino acids glycine and alanine.

Outlook

The research of any scientist stands on the foundation built by those who came before him. He strives in his studies to not only advance the body of knowledge in the field for his own good, but also to lay the groundwork for those who will come after him. This is quite important in the rapidly advancing field of computational chemistry, where problems that are insoluble today can be dealt with effectively in the near future as computing technology advances. In this section I outline a few problems that I encountered in my research which I could not completely solve, but which should be tackled by those who come after me as research in this field advances.

The first major problem I encountered in modeling the specific rotation of solvated molecules had not to do with the chiroptical response functions, but simply the relative energies of differing conformers of the molecules being modeled. Most of the benchmarking studies on modeling chiroptical activity have been done on rigid molecules that only adopt one conformer, conveniently avoiding this problem. But if studies of molecules with more than one conformer are done, and they should be, energy calculations more accurate than the DFT/COMSO model that I used must be developed to yield more accurate Boltzmann populations of these

conformers. Particularly with molecules that could form intramolecular hydrogen bonds, the error due to improper weighting of the optical rotations of these conformers had a very significant impact on my research, and seriously detracts from the utility of TDDFT modeling of specific rotation as a means of assigning absolute configuration.

Another issue which must be dealt with is the infamous charge transfer problem of TDDFT. This has to some extent, affected all of my results. It is a problem I did not solve; I could only do my best to avoid its effects. The only sure way I could avoid the issue entirely was to not use TDDFT. The alternative method available to me, coupled cluster theory, is quite useful for verification of certain DFT based results, but is not as practical due to its less efficient scaling. For the modeling of such response properties involving electronic transitions of molecules exceeding 10 atoms in size, DFT based methods are the only practical method. Newer hybrid methods, such as CAM-B3LYP are under development by others that aim to lessen the charge transfer problem. However a perfect solution has yet to be found, and those modeling the chiroptical response of molecules in the UV/visible regions must be aware of how this problem will affect their results.

One final issue involves development of a better explicit solvation model. My point charge model is an effective but simplified way to model the polar water molecule, but it cannot model the polarization that these waters undergo. The DRF model, already developed can do this quite well, but is not yet implemented in a program that can do hybrid TDDFT response calculations. One of two things should happen for this research to advance to the next level: either hybrid TDDFT must be perfected in the ADF program for which the DRF model is implemented, or a polarizable model such as DRF should be implemented for a program for

which hybrid DFT is already available. This is really just an issue of writing the code, but I believe that the results produced will be useful enough that such code should be written.

APPENDIX A:

Quick and Dirty Instructions for TURBOMOLE

I. Before You Begin

Whenever someone gets a new account on a computer there is always a little work that must be done to allow easy access to the software that he or she will be using. Here we will assume that you have been given a new account on the Autschbach group's computing cluster called "thewho". You want to use the computational chemistry program "Turbomole"[166]. So the first thing to do is to setup your account to tell it where to find the Turbomole programs. We do this by editing your `.bashrc` file, a hidden file in your home directory that should be executed every time you log into thewho. You can use "vi", "emacs" or any text editor to edit this file. For example:

```
emacs .bashrc
```

This will open up your .bashrc file for editing. There may or may not already be information in this file regarding other programs. Regardless, the following three lines should be added to the file:

```
#TURBOMOLE
export TURBOMOLE_SYSNAME=i786-pc-linux-gnu
export TURBODIR=/util/tmole/v5.7.1/TURBOMOLE
source $TURBODIR/Config_turbo_env
```

Save these changes to the file and exit the editor. Now you can log out of your account and log back in. From now on every time you log into your account on thewho the computer will know where to find the general information for running TURBOMOLE.

To test this, type

```
which dscf
```

on the command line. If the answer is something like 'file not found' you need to issue the following command in your home directory:

```
ln -s .bashrc .profile
```

(don't forget the dots in front of the file name). After logout and login, things should work.

II. Building Your Molecule

Every TURBOMOLE job should be in its own directory. This is true because the program generates a number of files with generic names, and if you were to run two jobs out of the same directory these files would overwrite each other and calculations would crash. Create the directory using the Linux command "mkdir":

```
mkdir {mydirectoryname}
```

Then switch to it using "cd":

```
cd {mydirectoryname}
```

The most straight forward way to build a molecule is to set it up as an xyz matrix. Here is an example of a properly formatted .xyz file for a water molecule:

```
------------ file starts next line ----------------------
3
(this line can be blank or filled with a comment)
O     0.000000     0.000000     0.000000
H     0.000000     0.500000     0.950000
H     0.000000     0.500000    -0.950000
------------ file ends above previous line -----------------
```

The first line is the number of atoms. The second may contain the energy of the molecule, but for input purposes is can be left blank. All lines thereafter are in the format "Atom X Y Z". The Atom is designated by its atomic symbol and its x y and z coordinates are in angstroms. Creating a .xyz file by hand for a very small molecule such as water is easy, but for bigger

molecules it is best to use a graphical utility such as “molden” which can be executed by simply typing:

```
molden
```

As with any graphical user interface the best way to learn how to use molden is to play with it. Once you have your molecule built be sure to save it in the .xyz format. Exit molden by clicking on the skull and crossbones icon on the menu bar, don’t just try to close the molden window or this will cause the program to hang.

The final step to setting up your molecule is converting it from the .xyz format to the TURBOMOLE format. This can be accomplished using the “x2t” utility:

```
x2t {filename.xyz} > coord
```

This creates a TURBOMOLE coordinate file called “coord” from your xyz file. Of course insert whatever you named your .xyz file in the `{filename.xyz}` spot, but it is best to always call the final TURBOMOLE file “coord”.

III. Setting up your calculation using define

The primary setup utility program for TURBOMOLE is “define”. It is an interactive text based program that walks you through several questions that you must answer and parameters that you must set in order to run your calculation. This program is started by typing on the command line:

```
define
```

It asks for a title - type anything you want and hit return, then type

```
a coord
```

to link the program to your coord file.

Now to setup the symmetry:

```
sy c1
```

sets the molecule to no symmetry, or use

```
desy
```

to determine symmetry automatically. If using the COSMO solvent model or some other component that is incompatible with symmetry, you must use `sy c1`. Else it is best to use `desy`. Of course if your molecule has no symmetry i.e. it belongs to point group c1 then both sy c1 and desy will give the same result. Once you have this set type `*` to get out of the menu.

It asks if you want to use internal coordinates; type `no` to just use Cartesian.

In the next menu, we setup the basis set by typing `b`. Then type `all {name of your basis set}` to assign the basis sets for your molecule. Ask for advice on which basis set to use, this will usually be a compromise decision based on the size of your molecule. For geometry optimizations and energy calculations where COSMO solvation is being employed, basis sets for which the model has been calibrated (like SV(P) for big molecules or def-TZVP for smaller ones) seem to yield the best results[33], though others can be used as well. For TDDFT response calculations where electronic excitations are involved, Dunning's augmented, correlation consistent basis sets are quite popular[64], their diffuse functions being essential for modeling some excited states. Assuming the def-TZVP basis is chosen, a typical input would be:

```
b all def-TZVP
```

You may also in the future find a need to use different basis sets for different atoms, but for this example we'll keep it simple. Once the basis set in place type `*` for the next menu

Type `eht` to set up orbitals using extended Huckel theory. It will then ask three or four questions. In general just hit `{return}` a few times to accept the default answers. *The only*

question you may have to pay attention to here is the question about your molecule's charge. If it is neutral then the default value (0) is fine, if it is charged you must enter that value.

With the orbitals all set it's time to turn on the DFT setup which is found in the menu shown *after* you accept the Huckel occupation numbers.

Type `dft`. Type `on` to activate density functional theory. Type `func {name of your functional}`. When is doubt of which functional to use, start with `func b3-lyp`. This hybrid, consisting of 80% DFT and 20% exact exchange, is the most commonly used functional in the world[167]…because it often works well. Note that for electronic response calculations using TDDFT, functions with a greater portion of exact exchange, like `pbe0` (25%ee) or even `bhlyp` (50%ee) should be considered. Generally, the more diffuse the states, the more ee is appropriate; those who will be performing many TDDFT calculations in their research should familiarize themselves with TDDFT's charge transfer problem[140] to see why this is the case. Once your function is set, type `grid m4` to set better grid size, then `*` for the next menu. For frequency calculations where greater precision is crucial, the tighter `grid m5` should be used.

Here it is good practice to increase the "scfiterlimit" value and convergence criteria for the job. While this is not the case for many calculations, for some jobs the default value for this is too low for the SCF to converge, which results in job crashes. To increase this value, do the following: Type `scf`, followed by `conv`, then `7`, then `iter`, and finally the limit for the number of scf cycles, `100` should be sufficient to converge to within 10-7 hartree, but if not, this number can be made even higher. When finished, hit `{return}` to exit the scf module.

At this point you're finished with the basic setup and if you're not interested in calculating any response properties (optical rotation, CD spectrum) you can type `*` a couple of times to exit the define program. However if you want to calculate some response properties, such as circular dichroism or optical rotatory dispersion you must continue:

If you want to setup a response calculation for ORD:

Type `ex`, followed by `dynpol` and `*`. Type `nm` to set up using nanometers. Then type `a` to start entering your wavelengths. Enter them one by one hitting `{return}` after each. There is a limit to how many wavelengths TURBOMOLE can calculate in one run, which seems to depend on the size of one of the files created in the process. The only way to find out this limit is to exceed it ☺. I've found doing 8 wavelengths is okay, but 16 causes crashes. So to do 16 I just do two calculations of 8 wavelengths each in two different directories. When you're done type `*` then another `*` to stop appending wavelengths.

If you want to setup a response calculation for CD:

Type `ex`, followed by `rpas` and `*`. Then type the symmetry designation of the type of excitations desired followed by the number; for an asymmetrical molecule, all of these symmetries will be of the "a" type, so type `a {number of excitations}`. Then type `*` to exit this sub-menu.

If doing either type of response calculation (CD or ORD), one should increase the memory available to the escf module for this purpose. Type `rpacor {memory size}` to change the mem size to a higher number of MB, as the default value does not take advantage of the memory

available in current technology. Jochen recommends 400. Just keep it well under the 1GB memory size for thewho's double processor nodes. To get out of the program now type `*`, `{return}` to accept a default, then `*` to exit define.

IV. Incorporating solvent effects using cosmoprep

Once you're done with define you're ready to run cosmoprep to set up your solvent field: Type `cosmoprep` to start the program.

Type `78.5` for the 1st number; this is the dielectric constant for water at room temperature. If your solvent is something other than water you can usually look up the number you'll want to use in the CRC handbook (on-line at the library web pages).

Hit `{return}` several times to accept default values until the radius menu comes up.

Most atoms have a "COSMO optimized" radius available, to use this, type "`r all o` ". For the less common atoms, optimized radii are not defined, but Turbomole has radii available which are set at 1.17 times the van-der-Waals radii[34], type "`r all b`" to use these radii.

Finally type `*` to exit cosmoprep. If you go back and look at your control file, you'll notice that cosmoprep has added a few lines to it. Now you're ready to start your calculations.

V. Post define modifications to the control file

After define (and cosmoprep) has (have) completed:

The *define* program generates several files, having the names of "basis" (the basis set) "mos" (the molecular orbitals information) and "control", the file that tells the other Turbomole programs what to do. It is this control file which you may want to alter somewhat after define is

finished generating it. As with any other file, it may be edited with emacs or vi, or any other editor of choice. If doing response calculations, one line that may be of use is the following:

```
$escfiterlimit 100
```

This is so because, as with the dscfiterlimit, the escfiterlimit default value is sometimes too low, and can occasionally cause response calculations to fail.

For those calculating a cd spectrum, the a line with the keyword `$cdspectrum rcm` is useful, in that it generates a circular dichroism spectrum file (using reciprocal centimeters as the energy units) with which Jochen's "plotspecTM" script can be used to plot a graphical CD spectrum.

Keywords such as these may be placed almost anywhere inside the control program, provided they are not inside other blocks of commands (which are indented) or after the "$end" at the end of the file.

VI. Submitting your calculation with PBS

PBS stands for Portable Batch System. It is software developed in part by NASA to sort of direct traffic on super computer systems. It's not rocket science. What PBS does is to take input for computations, what resources are needed, and starts the calculation on one or more of the nodes of a parallel computer once the resources are available. It is convenient to use if several people (like in our research group) use the same cluster for many different computations. In essence you just need to tell PBS what resources you need (how much time on how many processors) and it finds an efficient way to allocate resources to your job. Here is an example of a simple PBS script that runs a geometry optimization, a single point calculation, a response calculation and a frequency calculation on one processor for 24 hours:

```
---------------------- file starts next line---------------
```

```
#!/bin/sh
#PBS -q em64t
#PBS -l nodes=1:ppn=1
#PBS -l walltime=24:00:00
#PBS -o PBSoutputfilename.out
#PBS -N myjobname
#PBS -j oe
#PBS -M youremail@buffalo.edu
#PBS -m e

#TURBOMOLE
export TURBOMOLE_SYSNAME=i786-pc-linux-gnu
export TURBODIR=/util/tmole/v5.7.1/TURBOMOLE
source $TURBODIR/Config_turbo_env

ulimit
cd $PBSTMPDIR
echo "running in directory $PBSTMPDIR"
cp $PBS_O_WORKDIR/* $PBSTMPDIR
##########################USER_INPUT############################

jobex -c 100
dscf
escf
NumForce

##########################INPUT_END#############################
mv $PBSTMPDIR/* $PBS_O_WORKDIR
cd $PBS_O_WORKDIR
rmdir $PBSTMPDIR
---------------------- file ends previous line -------------
```

The first block tells PBS how many nodes (machines) you want to use and how many processors on each node. Running on more than one processor is called "running in parallel" and it requires a more complex PBS script than this. Filling out this block are the walltime - the time limit for your job and some information on what to call the output file and whom to e-mail when the job is done. Following this is a little information specific to running TURBOMOLE plus a couple of lines setting up a temporary directory and copying your files there.

Next comes the "user input" which is all the mini-programs you want TURBOMOLE to run. `jobex` runs a geometry optimization; the `- c 100` extension tells it to run for a maximum of 100 cycles, as the default of 20 is almost never enough. `dscf` runs a single point calculation,

which pretty much just gives you the energy of one configuration of a molecule, but sometimes you need to run it to set up for other calculations. `escf` calculates the response properties: any ORD or CD calculations you may have set up. `NumForce` calculates force constants and vibrational frequencies and gives you the essential information to make thermal corrections to your calculated energies.

The last few lines move all of your files back out of the temporary PBS directory and delete that temporary directory.

VI. Extracting your results

Unless you tell them to do otherwise jobex, dscf and escf will simply dump their output to default, which in this case would be your PBS output file. If you wish you may re-direct any of their output to the file of your choosing, for example: `dscf > mydscffile.out`.

NumForce puts most of its output into a directory it creates called "numforce". Any of these output files can be opened and dissected with a text editor.

There are also some scripts in the bin directory that can be used to extract certain information from these files. For example there is a command in the bin directory that when called: `tm-extract-ord < {escfoutputfile}` will output any specific rotation found in a particular output file. Little scripts like this can be written in simple programming languages such as "awk" and can be quite useful. If you think you can make your data creating and analyzing life easier by writing your own scripts, by all means do it.

APPENDIX B:

A basic guide to running GROMACS

I am not an expert on the GROMACS molecular dynamics program. However late in my graduate career I have become proficient enough in it to get it to run and generate structures of water solvated amino acids that were reasonable enough to get good results from. The following is a brief guide describing how I set up my calculations, and extracted my results. It is by no means exhaustive, but should provide a neophyte the basic knowledge needed to execute a basic set of calculations.

All calculations must begin with a Protein Data Bank (.pdb) file, which contains the geometry of the solute. Documents describing this format can be found at http://www.wwpdb.org/docs.html. It is not a forgiving format. Unlike .xyz files which are delimited by spaces, the fields in .pdb files are determined strictly by the columns in which the information appears. An example .pdb file of the proline zwitterion is shown below; the numbers in red are not part of the file but are included in order to emphasize the columns in which that data fall.

```
          1         2         3         4         5         6         7
1234567890123456789012345678901234567890123456789012345678901234567890
ATOM      1  O    PRO     1      -3.730  12.010  12.290  1.00  0.00
ATOM      2  C    PRO     1      -4.300  11.940  11.170  1.00  0.00
ATOM      3  OXT  PRO     1      -5.000  12.820  10.640  1.00  0.00
ATOM      4  CA   PRO     1      -4.130  10.640  10.380  1.00  0.00
ATOM      5  HA   PRO     1      -4.956  10.563   9.672  1.00  0.00
ATOM      6  CB   PRO     1      -4.150   9.410  11.270  1.00  0.00
ATOM      7  HB1  PRO     1      -5.166   9.101  11.516  1.00  0.00
ATOM      8  HB2  PRO     1      -3.490   9.592  12.118  1.00  0.00
ATOM      9  CG   PRO     1      -3.470   8.430  10.340  1.00  0.00
ATOM     10  HG1  PRO     1      -4.079   8.203   9.465  1.00  0.00
ATOM     11  HG2  PRO     1      -3.138   7.573  10.926  1.00  0.00
ATOM     12  CD   PRO     1      -2.250   9.270   9.990  1.00  0.00
ATOM     13  HD1  PRO     1      -1.706   8.847   9.145  1.00  0.00
ATOM     14  HD2  PRO     1      -1.665   9.401  10.900  1.00  0.00
ATOM     15  N    PRO     1      -2.860  10.560   9.640  1.00  0.00
ATOM     16  H2   PRO     1      -3.040  10.630   8.660  1.00  0.00
ATOM     17  H1   PRO     1      -2.240  11.300   9.920  1.00  0.00
```

The three familiar numbers starting in column 33 are the x, y and z coordinates of the molecule, in angstroms. The info in columns 18-20 corresponds to the amino acid identity, and the data in columns 14-16 identify the atoms. Note that all of the atom identifiers are unique, e.g. H1, H2, for hydrogen number one, hydrogen number two, etcetera. The letters A, B, G, and D in column 15 designate the alpha, beta, gamma and delta positions, respectively. All of this information is critical for the GROMACS program to assign force field parameters to the atoms. Non-amino acids can also be calculated, however such non-standard molecules may require additional input parameters beyond what is included in a .pdb file such as this.

The only other file that must exist in order to setup an MD calculation for a solvated amino acid is an .mdp file. While the .pdb file contains all the information about the amino acid structure and atom types, the .mdp file tells the program how the molecular dynamics program is to be run. An example file is shown below:

```
title                =  cpeptide MD
cpp                  =  /usr/bin/cpp
constraints          =  all-bonds
integrator           =  md
dt                   =  0.001     ; ps !
nsteps               =  2560000   ; total 2560 ps.
nstcomm              =  1
nstxout              =  10000
nstvout              =  0
nstfout              =  0
nstlist              =  10
ns_type              =  grid
rlist                =  1.0
rcoulomb             =  1.0
rvdw                 =  1.0
; Berendsen temperature coupling is on in two groups
Tcoupl               =  berendsen
tau_t                =  0.1           0.1
tc-grps              =  protein       sol
ref_t                =  300           300
; Pressure coupling is not on
Pcoupl               =  no
tau_p                =  0.5
compressibility      =  4.5e-5
ref_p                =  1.0
```

```
; Generate velocites is on at 300 K.
gen_vel             =  yes
gen_temp            =  300.0
gen_seed            =  173529
```

Most of these parameters I simply leave alone. The ones that are critical to change are "nsteps" and "nstxout", which control the number of molecular dynamics steps to be run and the frequency with which a geometry is written to output. This example is for a rather long production run of 2650 picoseconds, with output being written every 10ps (all numbers are in fs, 1ps = 1000fs). Of course for short test jobs smaller numbers should be used. The other number that I have had reason to change is the gen_seed. This can be virtually anything, but when executing multiple production runs where different geometries are desired, using different random number seeds are a requirement.

Once these two files are available in the same folder, the setup of the MD run can begin. The first program that needs to be run is "pdb2gmx" An example of the run line is below, using the .pdb file name "eiwit.pdb", which is the default name for a .pdb input. TIP3P type waters, which are very common water solvent models are used here:

```
pdb2gmx -f eiwit.pdb -water tip3p -o pdb2gmx.gro -p pdb2gmx.top
```

The program will ask for a force field. For starters I recommend number 6, the OPLS-AA-FF. It is an all atom force field. Unless using extremely large molecules (for which subsequent ab-initio calculations would be impractical, anyway), an all atom force field, as opposed to a unified one should be used. Upon completion of the execution of pdb2gmx, two files should be generated, pdb2gmx.gro and pdb2gmx.top.

Next three programs must be run which setup the solute molecule in a periodic box of waters. The box size, the only critical parameter here, is set to 2.5 nanometers on each edge:

```
editconf -f pdb2gmx.gro -o editconf.gro -box 2.5 2.5 2.5
```

```
genbox -cp pdb2gmx.gro -cs -o genbox.gro -p pdb2gmx.top -box 2.5
2.5 2.5
grompp -f gro.mdp -c genbox.gro -p pdb2gmx.top -o topol.tpr
```

Upon completion of these programs, the molecular dynamics run is ready to begin. GROMACS is nice in that only one file is needed for the execution of the production run. Here the file is named "topol.tpr". As such, I like to take time here to clear out all the files created during the setup that are no longer needed:

```
rm editconf.gro
rm genbox.gro
rm pdb2gmx.gro
rm mdout.mdp
rm pdb2gmx.top
rm '#pdb2gmx.top.1#'
rm posre.itp
```

At this point, the molecular dynamics run is ready for execution. It can be run in the command line with by simply typing, "mdrun". For a simple test job this is okay, but time consuming production runs should be submitted through the PBS que. An example PBS input file follows:

```
---------------------- file starts next line---------------
#!/bin/sh
#PBS -q em64t
#PBS -l nodes=1:ppn=1
#PBS -l walltime=24:00:00
#PBS -o PBSoutputfilename.out
#PBS -N myjobname
#PBS -j oe
#PBS -M youremail@buffalo.edu
#PBS -m e

ulimit
cd $PBSTMPDIR
echo "running in directory $PBSTMPDIR"
cp $PBS_O_WORKDIR/* $PBSTMPDIR
##########################USER_INPUT############################

mdrun
trjconv -o movie.pdb  << eor > /dev/null
0
eor
```

```
##########################INPUT_END##########################
mv $PBSTMPDIR/* $PBS_O_WORKDIR
cd $PBS_O_WORKDIR
rmdir $PBSTMPDIR
---------------------- file ends previous line -------------
```

As with the Turbomole PBS script in the previous section, this simply copies all the files in the directory to the production node, runs the job, then copies all the files back and destroys the temporary directory created on the production node. Here I have it do two executions: first the mdrun, which runs the dynamics, of course. Next I have it execute trjconv, creating a file called, "movie.pdb" from which geometry data may be extracted from later. I have several scripts written for the purposes of this extraction. One may of course manipulate the data in whatever way he sees fit once the GROMACS program has output it.

Vitae of Matthew David Kundrat

Education:

- 2008 Ph.D. in Computational Chemistry
  State University of New York at Buffalo

- 2004 M.A. in Inorganic Chemistry
  Wayne State University

- 1997 B.S. in Chemistry with a minor in Computer Science
  Clarion University of Pennsylvania

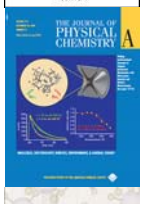
Awards:

- 1997 Clarion University Chemistry Department Competitive Award
- 1994 Advanced Placement Scholar With Distinction

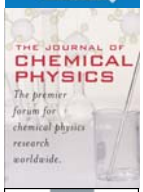
Publications:

9. Kundrat, Matthew D.; Autschbach, Jochen. Ab-initio and density functional theory modeling of the chiroptical response of glycine and alanine in solution using explicit solvation and molecular dynamics. Journal of Chemical Theory and Computation *Submitted.*

8. Kundrat, Matthew D.; Autschbach, Jochen. Computational modeling of the optical rotation of amino acids: A new look at an old rule for pH dependence of optical rotation. Journal of the American Chemical Society; **2008**; *130*(13) pp 4004-4414.

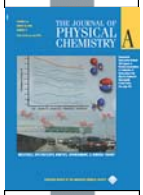


7. Kundrat, Matthew D.; Autschbach, Jochen. Time dependent density functional theory modeling of specific rotation and optical rotatory dispersion of the aromatic amino acids in solution. Journal of Physical Chemistry A; **2006**; *110*(47) pp 12908-12917.



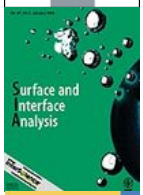
6. Krykunov, Mykhaylo; Kundrat, Matthew D.; Autschbach, Jochen. Calculation of CD spectra from optical rotatory dispersion, and vice versa, as complementary tools for theoretical studies of optical activity using time-dependent density functional theory. Journal of Chemical Physics; **2006**; *125* pp 194110-194122.

5. Balogh, Michael P.; Jones, Camille Y.; Herbst, J.F.; Hector, Jr., Louis G.; Kundrat, Matthew. Crystal structures and phase transformation of deuterated lithium imide, Li_2ND. Journal of Alloys and Compounds; **2006**; *420* pp 326-336.



4. Kundrat, Matthew D.; Autschbach, Jochen. Time dependent density functional theory modeling of chiroptical properties of small amino acids in solution. Journal of Physical Chemistry A; **2006**; *110*(11) pp 4115-4123.

3. Meisner, Gregory P.; Pinkerton, Frederick E.; Meyer, Martin S.; Balogh, Michael P.; Kundrat, Matthew D.. Study of the lithium-nitrogen-hydrogen system. Journal of Alloys and Compounds; **2005**; *404-406* pp 24-26.



2. Gaarenstroom, Stephen W.; Balogh, Michael P.; Militello, Maria C.; Waldo, Richard A.; Wong, Curtis A.; Kelly, Nelson A.; Gibson, Thomas L.; Kundrat, Matthew D.. Characterization of indium-tin-oxide films with improved corrosion resistance. Surface and Interface Analysis **2005**; *37*(4) pp 385-392.

1. Pinkerton, Frederick E.; Meisner, Gregory P.; Meyer, Martin S.; Balogh, Michael P.; Kundrat, Matthew D.. Hydrogen desorption exceeding ten weight percent from the new quaternary hydride $Li_3BN_2H_8$. Journal of Physical Chemistry B; **2005**; *109*(1) pp 6 - 8.

Conference Participation:

6. Kundrat, Matthew D.; Autschbach, Jochen. Modeling the effects of an achiral solvent on the chiroptical response of chiral amino acids. Talk presented at 40th Midwest Theoretical Chemistry Conference, Ann Arbor, MI, June 26-28, 2008.

5. Kundrat, Matthew D.; Autschbach, Jochen. Time-dependent density functional theory calculation of specific rotation using molecular dynamics and explicit solvation. Poster and talk presented at 233rd ACS National Meeting, Chicago, IL, March 25-29, 2007; PHYS 56.

4. Kundrat, Matthew D.; Autschbach, Jochen. Specific rotation of amino acids: A new look at an old rule. Poster presented at 232nd ACS National Meeting, San Francisco, CA, Sept. 10-14, 2006, PHYS 449

3. Kundrat, Matthew D.; Autschbach, Jochen. Time-dependent density functional theory modeling of the chiroptical properties of aromatic amino acids in solution. Poster presented at 38th Midwest Theoretical Chemistry Conference, Columbus, OH, June 15-17, 2006

2. Kundrat, Matthew D.; Autschbach, Jochen. Comparison of methods of density functional theory modeling of the specific rotation of amino acid solutions. Poster presented at 230th ACS National Meeting, Washington, DC, Aug. 28-Sept. 1, 2005, COMP 215.

1. Second Annual Wayne State University Graduate Research Symposium: Organizing committee member. Detroit, MI August 18, 2000.

Patent:

1. Meisner, Gregory P.; Pinkerton, Frederick E.; Meyer, Martin S.; Balogh, Michael P.; Kundrat, Matthew D.. Imide/amide hydrogen storage system. United States Patent #6967012, Nov. 22, 2005

REFERENCES:

1. Polavarapu, P.L., *Optical rotation: Recent advances in determining the absolute configuration.* Chirality, 2002. **14**(10): p. 768-781.
2. Polavarapu, P.L., *Optical rotation: Recent advances in determining the absolute configuration (vol 14, pg 768, 2002).* Chirality, 2003. **15**(3): p. 284-285.
3. Autschbach, J., et al., *Chiroptical properties from time-dependent density functional theory. II. Optical rotations of small to medium sized organic molecules.* Journal of Chemical Physics, 2002. **117**(2): p. 581-592.
4. Stephens, P.J., et al., *Determination of absolute configurations of chiral molecules using ab initio time-dependent density functional theory calculations of optical rotation: How reliable are absolute configurations obtained for molecules with small rotations?* Chirality, 2005. **17**: p. S52-S64.
5. Pecul, M. and K. Ruud, *The ab initio calculation of optical rotation and electronic circular dichroism*, in *Advances In Quantum Chemistry, Vol 50*. 2005, Elsevier Academic Press Inc: San Diego. p. 185-212.
6. Crawford, T.D., *Ab initio calculation of molecular chiroptical properties.* Theoretical Chemistry Accounts, 2006. **115**(4): p. 227-245.
7. Autschbach, J., F.E. Jorge, and T. Zlegler, *Density functional calculations on electronic circular dichroism spectra of chiral transition metal complexes.* Inorganic Chemistry, 2003. **42**(9): p. 2867-2877.
8. Mennucci, B., et al., *Polarizable continuum model (PCM) calculations of solvent effects on optical rotations of chiral molecules.* Journal of Physical Chemistry A, 2002. **106**(25): p. 6102-6113.
9. Marchesan, D., et al., *Optical rotation calculation of a highly flexible molecule: The case of paraconic acid.* Journal of Physical Chemistry A, 2005. **109**(7): p. 1449-1453.
10. Specht, K.M., et al., *Determining absolute configuration in flexible molecules: A case study.* Journal of the American Chemical Society, 2001. **123**(37): p. 8961-8966.
11. Giorgio, E., et al., *Ab initio calculation of optical rotatory dispersion (ORD) curves: A simple and reliable approach to the assignment of the molecular absolute configuration.* Journal of the American Chemical Society, 2004. **126**(40): p. 12968-12976.
12. Lattanzi, A., et al., *Time dependent density functional response theory calculation of optical rotation as a method for the assignment of absolute configuration of camphor-derived furyl hydroperoxide and alcohol.* Journal Of Physical Chemistry A, 2004. **108**(48): p. 10749-10753.
13. da Silva, C.O., B. Mennucci, and T. Vreven, *Density functional study of the optical rotation of glucose in aqueous solution.* Journal of Organic Chemistry, 2004. **69**(23): p. 8161-8164.
14. Stephens, P.J., et al., *Determination of absolute configuration using ab initio calculation of optical rotation.* Chirality, 2003. **15**: p. S57-S64.
15. Berova, N., K. Nakanishi, and R.W. Woody, *Circular Dichroism: Principles and Applications*. 2nd ed. 2000, New York: John Wiley & Sons, Inc.
16. Jorgensen, E.C., *Sector rule for relating optical rotatory dispersion with the conformation and absolute configuration of alpha-amino acids.* Tetrahedron Letters, 1971. **13**: p. 863-866.

17. Crawford, T.D., M.C. Tam, and M.L. Abrams, *The current state of ab initio calculations of optical rotation and electronic circular dichroism spectra.* Journal of Physical Chemistry A, 2007. **111**(48): p. 12057-12068.
18. Polavarapu, P.L., *Renaissance in chiroptical spectroscopic methods for molecular structure determination.* Chemical Record, 2007. **7**(2): p. 125-136.
19. Autschbach, J., *Computation of Optical Rotation using Time-Dependent Density Functional Theory.* Computing Letters, 2007. **3**(2-4): p. 131-150.
20. Kundrat, M.D. and J. Autschbach, *Time dependent density functional theory modeling of specific rotation and optical rotatory dispersion of the aromatic amino acids in solution.* Journal of Physical Chemistry A, 2006. **110**(47): p. 12908-12917.
21. Grimme, S., F. Furche, and R. Ahlrichs, *An improved method for density functional calculations of the frequency-dependent optical rotation.* Chemical Physics Letters, 2002. **361**(3-4): p. 321-328.
22. Klamt, A. and V. Jonas, *Treatment of the outlying charge in continuum solvation models.* Journal Of Chemical Physics, 1996. **105**(22): p. 9972-9981.
23. Wada, G., et al., *On The Ratio Of Zwitterion Form To Uncharged Form Of Glycine At Equilibrium In Various Aqueous-Media.* Bulletin Of The Chemical Society Of Japan, 1982. **55**(10): p. 3064-3067.
24. Jorgensen, W.L., et al., *COMPARISON OF SIMPLE POTENTIAL FUNCTIONS FOR SIMULATING LIQUID WATER.* Journal Of Chemical Physics, 1983. **79**(2): p. 926-935.
25. Tu, Y.Q. and A. Laaksonen, *The electronic properties of water molecules in water clusters and liquid water.* Chemical Physics Letters, 2000. **329**(3-4): p. 283-288.
26. Swart, M. and P.T. van Duijnen, *DRF90: a polarizable force field.* Molecular Simulation, 2006. **32**(6): p. 471-484.
27. Kondru, R.K., P. Wipf, and D.N. Beratan, *Atomic contributions to the optical rotation angle as a quantitative probe of molecular chirality.* Science, 1998. **282**(5397): p. 2247-2250.
28. Pecul, M., et al., *Conformational effects on the optical rotation of alanine and proline.* Journal of Physical Chemistry A, 2004. **108**(19): p. 4269-4276.
29. Stephens, P.J., et al., *Calculation of optical rotation using density functional theory.* Journal Of Physical Chemistry A, 2001. **105**(22): p. 5356-5371.
30. Ahlrichs, R., et al., *Electronic-Structure Calculations On Workstation Computers - The Program System Turbomole.* Chemical Physics Letters, 1989. **162**(3): p. 165-169.
31. Becke, A.D., *Density-Functional Thermochemistry .3. The Role Of Exact Exchange.* Journal Of Chemical Physics, 1993. **98**(7): p. 5648-5652.
32. PNNL, *EMSL Gaussian Basis Set Order Form*. 2005, Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 02/25/04, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multi-program laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact Karen Schuchardt for further information. .
33. Schafer, A., et al., *COSMO Implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems.* Physical Chemistry Chemical Physics, 2000. **2**(10): p. 2187-2193.

34. Bondi, A., *van der Waals Volumes and Radii.* Jounal of Physical Chemistry, 1964. **68**(3): p. 441.
35. Schaftenaar, G., *Molden*. 2005, Centre for Molecular and Biomolecular Informatics.
36. Bour, P., J. Kapitan, and V. Baumruk, *Simulation of the Raman optical activity of L-alanyl-L-alanine.* Journal Of Physical Chemistry A, 2001. **105**(26): p. 6362-6368.
37. Ruud, K. and T. Helgaker, *Optical rotation studied by density-functional and coupled-cluster methods.* Chemical Physics Letters, 2002. **352**(5-6): p. 533-539.
38. Lovy, D., *WinDIG*. 2005.
39. Jensen, J.H. and M.S. Gordon, *On The Number Of Water-Molecules Necessary To Stabilize The Glycine Zwitterion.* Journal Of The American Chemical Society, 1995. **117**(31): p. 8159-8170.
40. Tomasi, J., B. Mennucci, and R. Cammi, *Quantum mechanical continuum solvation models.* Chemical Reviews, 2005. **105**(8): p. 2999-3093.
41. Tortonda, F.R., et al., *Aminoacid zwitterions in solution: Geometric, energetic, and vibrational analysis using density functional theory continuum model calculations.* Journal Of Chemical Physics, 1998. **109**(2): p. 592-602.
42. *The Merck Index*. Twelfth ed, ed. S. Budavari. 1996, Whitehouse Station, NJ: Merck & Co., Inc.
43. Djerassi, C., *Optical Rotatory Dispersion Applications to Organic Chemistry*. McGraw-Hill Series in Advanced Chemistry. 1960, New York: McGraw-Hill.
44. Crawford, T.D., et al., *Ab initio calculation of optical rotation in (P)-(+)- 4 triangulane.* Journal of the American Chemical Society, 2005. **127**(5): p. 1368-1369.
45. Bauernschmitt, R. and R. Ahlrichs, *Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory.* Chemical Physics Letters, 1996. **256**(4-5): p. 454-464.
46. Jankowski, K., F. Soler, and M. Ellenberger, *Nmr Investigations Of Proline And Its Derivatives .5. Semiempirical Calculation Of Endo-Exo Equilibrium Of Pyrrolidine Ring And Its Ph-Dependence.* Journal Of Molecular Structure, 1978. **48**(1): p. 63-68.
47. Haasnoot, C.A.G., et al., *Relationship Between Proton-Proton Nmr Coupling-Constants And Substituent Electronegativities .3. Conformational-Analysis Of Proline Rings In Solution Using A Generalized Karplus Equation.* Biopolymers, 1981. **20**(6): p. 1211-1245.
48. Stepanian, S.G., et al., *Conformers of nonionized proline. Matrix-isolation infrared and post-Hartree-Fock ab initio study.* Journal Of Physical Chemistry A, 2001. **105**(47): p. 10664-10672.
49. Cappelli, C., S. Monti, and A. Rizzo, *Effect of the environment on vibrational infrared and circular dichroism spectra of (S)-proline.* International Journal Of Quantum Chemistry, 2005. **104**(5): p. 744-757.
50. Pogliani, L., M. Ellenberger, and J. Valat, *Nmr Investigation Of Proline And Its Derivatives .2. Conformational Implications Of H-1 Nmr-Spectrum Of L-Proline At Different Ph.* Organic Magnetic Resonance, 1975. **7**(2): p. 61-71.
51. Martin, R.B. and R. Mathur, *Analysis of Proton Magnetic Resonance Spectra of Cysteine and Histidine Deravitives. Conformational Equilibria.* Journal of the American Chemical Society, 1965. **87**(5): p. 1065-1070.

52. Casey, J.P. and R.B. Martin, *Disulfide Stereochemistry. Conformations and Chiroptical Properties of L-Cystine Derivatives.* Journal of the American Chemical Society, 1972. **94**(17): p. 6141-6151.
53. Kraszni, M., Z. Szakacs, and B. Noszal, *Determination of rotamer populations and related parameters from NMR coupling constants: a critical review.* Analytical and Bioanalytical Chemistry, 2004. **378**(6): p. 1449-1463.
54. Fujiwara, S., H. Ishizuka, and S. Fudano, *NMR Study of Amino Acids and Their Derivatives: Dissociation Constant of Each Rotational Isomer of Amino Acids.* Chemistry Letters, 1974. **11**: p. 1281-1284.
55. Noszal, B., W. Guo, and D.L. Rabenstein, *Rota-Microspeciation Of Serine, Cysteine, And Selenocysteine.* Journal Of Physical Chemistry, 1991. **95**(23): p. 9609-9614.
56. Feeney, J., *Improved Component Vicinal Coupling-Constants For Calculating Side-Chain Conformations In Amino-Acids.* Journal Of Magnetic Resonance, 1976. **21**(3): p. 473-478.
57. Noguera, M., et al., *Protonation of glycine, serine and cysteine. Conformations, proton affinities and intrinsic basicities.* Journal of Molecular Structure-Theochem, 2001. **537**: p. 307-318.
58. Kundrat, M.D. and J. Autschbach, *Time dependent density functional theory modeling of chiroptical properties of small amino acids in solution.* Journal Of Physical Chemistry A, 2006. **110**(11): p. 4115-4123.
59. Autschbach, J., et al., *Chiroptical properties from time-dependent density functional theory. I. Circular dichroism spectra of organic molecules.* Journal of Chemical Physics, 2002. **116**(16): p. 6930-6940.
60. Iizuka, E. and J.T. Yang, *Optical rotatory dispersion of L-amino acids in acid solution.* Biochemistry, 1964. **3**(10): p. 1519-1524.
61. Toome, V. and M. Weigele, *Determination of the Absolute Configuration of alpha-Amino Acids and Small Peptides by Chiroptical Means.* The Peptides, 1981. **4**: p. 85-184.
62. Rogers, D.M. and J.D. Hirst, *First-principles calculations of protein circular dichroism in the near ultraviolet.* Biochemistry, 2004. **43**(34): p. 11092-11102.
63. Tanaka, T., et al., *Electronic and vibrational circular dichroism of aromatic amino acids by density functional theory.* Chirality, 2006. **18**(8): p. 652-661.
64. Woon, D.E. and T.H. Dunning, *Gaussian-Basis Sets For Use In Correlated Molecular Calculations .4. Calculation Of Static Electrical Response Properties.* Journal Of Chemical Physics, 1994. **100**(4): p. 2975-2988.
65. Polavarapu, P.L., *Kramers-Kronig transformation for optical rotatory dispersion studies.* Journal Of Physical Chemistry A, 2005. **109**(32): p. 7013-7023.
66. Kainosho, M. and K. Ajisaka, *Conformational-Analysis Of Amino-Acids And Peptides Using Specific Isotope Substitution .2. Conformation Of Serine, Tyrosine, Phenylalanine, Aspartic-Acid, Asparagine, And Aspartic-Acid Beta-Methyl Ester In Various Ionization States.* Journal Of The American Chemical Society, 1975. **97**(19): p. 5630-5631.
67. Hansen, P.E., J. Feeney, and G.C.K. Roberts, *Long-Range C-13-H-1 Spin-Spin Coupling-Constants In Amino-Acids, Conformational Applications.* Journal Of Magnetic Resonance, 1975. **17**(2): p. 249-261.
68. Dezube, B., C.M. Dobson, and C.E. Teague, *Conformational-Analysis Of Tryptophan In Solution Using Nuclear Magnetic-Resonance Methods.* Journal Of The Chemical Society-Perkin Transactions 2, 1981(4): p. 730-735.

69. Juy, M., L.T. Hung, and S. Fermandjian, *CD and proton NMR studies on the side-chain conformation of tyrosine derivatives and tyrosine residues in di- and tripeptides.* International Journal of Peptide & Protein Research 1982. **20**(4): p. 298-307.
70. Reddy, M.C., et al., *Precise Values Of The C-13 Component Vicinal Coupling-Constants For Calculating Side-Chain Conformations In Amino-Acids.* Organic Magnetic Resonance, 1984. **22**(7): p. 464-467.
71. Merrett, J.H., et al., *The Synthesis And Rotational-Isomerism Of 1-Amino-2-Imidazol-4-Ylethylphosphonic Acid [Phosphonohistidine, His(P)] And 1-Amino-2-Imidazol-2-Ylethylphosphonic Acid [Phosphonoisohistidine, Isohis(P)].* Journal Of The Chemical Society-Perkin Transactions 1, 1988(1): p. 61-67.
72. Greenstein, J.P. and M. Winitz, *Chemistry of the Amino Acids*. 1961, New York: John Wiley & Sons.
73. ScienceLab.com. 2006 [cited 2006 July 15]; Available from: http://www.sciencelab.com/page/S/PVAR/10427/SLT3205.
74. Nishino, H., et al., *The pH dependence of the anisotropy factors of essential amino acids.* Journal Of The Chemical Society-Perkin Transactions 2, 2002(3): p. 582-590.
75. Moscowitz, A., A. Rosenberg, and A.E. Hansen, *On the 260-milimicron Cotton effect in L-phenylalanine.* Journal of the American Chemical Society, 1965. **87**(8): p. 1813-1815.
76. Neugebauer, J., et al., *Importance of vibronic effects on the circular dichroism spectrum of dimethyloxirane.* Journal of Chemical Physics, 2005. **122**(23): p. 7.
77. Autschbach, J., et al., *Time-dependent density functional calculations of optical rotatory dispersion including resonance wavelengths as a potentially useful tool for determining absolute configurations of chiral molecules.* Journal Of Physical Chemistry A, 2006. **110**(7): p. 2461-2473.
78. Giorgio, E., et al., *Determination of the absolute configuration of flexible molecules by ab initio ORD calculations: A case study with cytoxazones and isocytoxazones.* Journal Of Organic Chemistry, 2005. **70**(17): p. 6557-6563.
79. Patterson, J.W. and W.R. Brode, *A rotatory-dispersion study of alpha-amino acids.* Archives of Biochemistry, 1943. **2**: p. 247-257.
80. Hooker Jr., T.M. and C. Tanford, *Optical rotatory dispersion of L-tyrosine.* Journal of the American Chemical Society, 1964. **86**(22): p. 4989-4991.
81. Krykunov, M., M.D. Kundrat, and J. Autschbach, *Calculation of circular dichroism spectra from optical rotatory dispersion, and vice versa, as complementary tools for theoretical studies of optical activity using time-dependent density functional theory.* Journal of Chemical Physics, 2006. **125**(19): p. 13.
82. Polavarapu, P.L., *Protocols for the analysis of theoretical optical rotations.* Chirality, 2006. **18**(5): p. 348-356.
83. Bijvoet, J.M.P., A. F.; van Bommel, A. J., *Determination of the absolute configuration of optically active compounds by means of x-rays.* Nature, 1951. **168**: p. 271-272.
84. Haesler, J., et al., *Absolute configuration of chirally deuterated neopentane.* Nature, 2007. **446**(7135): p. 526-529.
85. Lightner, D.A. and T.C. Chang, *Octant Rule .3. Experimental Proof for Front Octants.* Journal of the American Chemical Society, 1974. **96**(9): p. 3015-3016.
86. Lightner, D.A., J.K. Gawronski, and T.D. Bouman, *Octant Rule .7. Deuterium as an Octant Perturber.* Journal of the American Chemical Society, 1980. **102**(6): p. 1983-1990.

87. Lightner, D.A., et al., *The Octant Rule .15. Antioctant Effects - Synthesis and Circular-Dichroism of 2-Exo-Alkylbicyclo[2.2.1]Heptan-7-Ones and 2-Endo-Alkylbicyclo[2.2.1]Heptan-7-Ones and 2-Endo-Bicyclo[3.2.1]Octan-8-Ones.* Journal of Organic Chemistry, 1985. **50**(20): p. 3867-3878.
88. Harada, N.N., K., *Exciton Coupling in Organic Stereochemistry*. 1983, Mill Valley, CA: University Science Books.
89. Clough, G.W., *Relationship between the optical rotatory powers and the relative configurations of optically active compounds. Influence of certain inorganic halides on the optical rotatory powers of a-hydroxy acids, a-amino acids, and their derivatives.* Journal of the Chemical Society, Transactions, 1918. **113**: p. 526-54.
90. Lutz, I.O. and B. Jirgensons, *A new method of assigning optically active a-amino acids to the d- or l-series.* Berichte der Deutschen Chemischen Gesellschaft [Abteilung] B: Abhandlungen, 1930. **63B**: p. 448-60.
91. Lutz, O. and B. Jirgensons, *New method for the grouping of optically active a-amino acids in the dextro- or levo-series. II. .* Berichte der Deutschen Chemischen Gesellschaft [Abteilung] B: Abhandlungen, 1931. **64B**: p. 1221-32.
92. Vleggaar, R., L.G.J. Ackerman, and P.S. Steyn, *Structure Elucidation Of Monatin, A High-Intensity Sweetener Isolated From The Plant Schlerochiton-Ilicifolius.* Journal Of The Chemical Society-Perkin Transactions 1, 1992(22): p. 3095-3098.
93. Baerends, E.J., et al., *Amsterdam Density Functional*. 2006, Scientific Computing and Modeling: Amsterdam, The Netherlands.
94. Frisch, M.J., et al., *Gaussian03*. 2004, Gaussian, Inc.: Wallingford CT.
95. *DALTON, a molecular electronic structure program, Release 2.0 (2005), see* http://www.kjemi.uio.no/software/dalton/dalton.html 2005.
96. Bylaska, E.J., et al., *NWChem, A Computational Chemistry Package for Parallel Computers*. 2006, Pacific Northwest National Laboratory: Richland, Washington 99352-0999, USA.
97. Cheeseman, J.R., et al., *Hartree-Fock and density functional theory ab initio calculation of optical rotation using GIAOs: Basis set dependence.* Journal Of Physical Chemistry A, 2000. **104**(5): p. 1039-1046.
98. Ruud, K. and R. Zanasi, *The importance of molecular vibrations: The sign change of the optical rotation of methyloxirane.* Angewandte Chemie-International Edition, 2005. **44**(23): p. 3594-3596.
99. Kongsted, J., et al., *Coupled cluster and density functional theory studies of the vibrational contribution to the optical rotation of (S)-propylene oxide.* Journal Of The American Chemical Society, 2006. **128**(3): p. 976-982.
100. Mort, B.C. and J. Autschbach, *Temperature dependence of the optical rotation in six bicyclic organic molecules calculated by vibrational averaging.* Chemphyschem, 2007. **8**(4): p. 605-616.
101. Pecul, M., et al., *Polarizable continuum model study of solvent effects on electronic circular dichroism parameters.* Journal Of Chemical Physics, 2005. **122**(2).
102. Mukhopadhyay, P., et al., *Solvent effect on optical rotation: A case study of methyloxirane in water.* Chemphyschem, 2006. **7**(12): p. 2483-2486.
103. Pecul, M., *Conformational structures and optical rotation of serine and cysteine.* Chemical Physics Letters, 2006. **418**(1-3): p. 1-10.

104. Mori, T., Y. Inoue, and S. Grimme, *Time dependent density functional theory calculations for electronic circular dichroism spectra and optical rotations of conformationally flexible chiral donor-acceptor dyad.* Journal of Organic Chemistry, 2006. **71**(26): p. 9797-9806.
105. Coriani, S., et al., *Solvent effects on the conformational distribution and optical rotation of gamma-methyl paraconic acids and esters.* Chirality, 2006. **18**(5): p. 357-369.
106. Kowalczyk, T.D., M.L. Abrams, and T.D. Crawford, *Ab initio optical rotatory dispersion and electronic circular dichroism spectra of (S)-2-chloropropionitrile.* Journal of Physical Chemistry A, 2006. **110**(24): p. 7649-7654.
107. Wiberg, K.B., et al., *Sum-over-states calculation of the specific rotations of some substituted oxiranes, chloropropionitrile, ethane, and norbornenone.* Journal of Physical Chemistry A, 2006. **110**(51): p. 13995-14002.
108. Peiponen, K.E., et al., *Kramers-Kronig relations and sum rules in nonlinear optical spectroscopy.* Applied Spectroscopy, 2004. **58**(5): p. 499-509.
109. Moffitt, W.W., R. B.; Moscowitz, A.; Klyne, W.; Djerassi, Carl, *Structure and the optical rotatory dispersion of saturated ketones.* Journal of the American Chemical Society, 1961. **83**: p. 4013-18.
110. Klyne, W., *Carboxyl and aromatic chromophores: optical rotatory dispersion and circular dichroism studies.* Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences, 1967. **297**(1448): p. 66-78.
111. Rinderspacher, B.C. and P.R. Schreiner, *Structure-property relationships of prototypical chiral compounds: Case studies.* Journal Of Physical Chemistry A, 2004. **108**(15): p. 2867-2870.
112. Barron, L.D., *Molecular Light Scattering and Optical Activity*. Second ed. 2004, Cambridge, UK: Cambridge University Press. 443.
113. Sahu, P.K. and S.L. Lee, *Effect of microsolvation on zwitterionic glycine: an ab initio and density functional theory study.* Journal of Molecular Modeling, 2008. **14**(5): p. 385-392.
114. Aikens, C.M. and M.S. Gordon, *Incremental solvation of nonionized and zwitterionic glycine.* Journal of the American Chemical Society, 2006. **128**(39): p. 12835-12850.
115. Takahashi, H., et al., *A novel quantum mechanical/molecular mechanical approach to the free energy calculation for isomerization of glycine in aqueous solution.* Journal Of Chemical Physics, 2005. **123**(12).
116. Ramaekers, R., et al., *Neutral and zwitterionic glycine.H2O complexes: A theoretical and matrix-isolation Fourier transform infrared study.* Journal of Chemical Physics, 2004. **120**(9): p. 4182-4193.
117. Balta, B. and V. Aviyente, *Solvent effects on glycine II. Water-assisted tautomerization.* Journal Of Computational Chemistry, 2004. **25**(5): p. 690-703.
118. Balta, B. and V. Aviyente, *Solvent effects on glycine. I. A supermolecule modeling of tautomerization via intramolecular proton transfer.* Journal of Computational Chemistry, 2003. **24**(14): p. 1789-1802.
119. Cui, Q., *Combining implicit solvation models with hybrid quantum mechanical/molecular mechanical methods: A critical test with glycine.* Journal Of Chemical Physics, 2002. **117**(10): p. 4720-4728.

120. Gontrani, L., B. Mennucci, and J. Tomasi, *Glycine and alanine: a theoretical study of solvent effects upon energetics and molecular response properties.* Journal of Molecular Structure-Theochem, 2000. **500**: p. 113-127.
121. Kassab, E., et al., *Theoretical study of solvent effect on intramolecular proton transfer of glycine.* Journal Of Molecular Structure-Theochem, 2000. **531**: p. 267-282.
122. Tortonda, F.R., et al., *Why is glycine a zwitterion in aqueous solution? A theoretical study of solvent stabilising factors.* Chemical Physics Letters, 1996. **260**(1-2): p. 21-26.
123. Kikuchi, O., et al., *Ab initio MO and Monte Carlo simulation study on the conformation of L-alanine zwitterion in aqueous solution.* Journal Of Physical Organic Chemistry, 1997. **10**(3): p. 145-151.
124. D'Abramo, M., et al., *Calculation of the optical rotatory dispersion of solvated alanine by means of the perturbed matrix method.* Chemical Physics Letters, 2005. **402**(4-6): p. 559-563.
125. Sagarik, K. and S. Dokmaisrijan, *A theoretical study on hydration of alanine zwitterions.* Journal of Molecular Structure-Theochem, 2005. **718**(1-3): p. 31-47.
126. Kundrat, M.D. and J. Autschbach, *Computational modeling of the optical rotation of amino acids: A new look at an old rule for pH dependence of optical rotation.* Journal of the American Chemical Society, 2008. **130**(13): p. 4404-4414.
127. Pecul, M., et al., *Solvent effects on Raman optical activity spectra calculated using the polarizable continuum model.* Journal Of Physical Chemistry A, 2006. **110**(8): p. 2807-2815.
128. Mukhopadhyay, P., et al., *Contiribution of a solute's chiral solvent imprint to optical rotation.* Angewandte Chemie-International Edition, 2007. **46**(34): p. 6450-6452.
129. Jensen, L., et al., *Circular dichroism spectrum of [Co(en)(3)](3+) in water: A discrete solvent reaction field study.* International Journal of Quantum Chemistry, 2006. **106**(12): p. 2479-2488.
130. Christiansen, O., H. Koch, and P. Jorgensen, *THE 2ND-ORDER APPROXIMATE COUPLED-CLUSTER SINGLES AND DOUBLES MODEL CC2.* Chemical Physics Letters, 1995. **243**(5-6): p. 409-418.
131. Pedersen, T.B., et al., *Origin invariant calculation of optical rotation without recourse to London orbitals.* Chemical Physics Letters, 2004. **393**(4-6): p. 319-326.
132. Lindahl, E., B. Hess, and D. van der Spoel, *GROMACS 3.0: a package for molecular simulation and trajectory analysis.* Journal of Molecular Modeling, 2001. **7**(8): p. 306-317.
133. Jorgensen, W.L., D.S. Maxwell, and J. TiradoRives, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.* Journal of the American Chemical Society, 1996. **118**(45): p. 11225-11236.
134. Isborn, C., K. Claborn, and B. Kahr, *The optical rotatory power of water.* Journal of Physical Chemistry A, 2007. **111**(32): p. 7800-7804.
135. Legrand, M. and R. Viennet, *Dichroisme circulaire optique. XV. - Etude de quelques acids amines.* Bulletin a la Societe Chimque, 1965: p. 679-681.
136. Neugebauer, J., et al., *The merits of the frozen-density embedding scheme to model solvatochromic shifts.* Journal Of Chemical Physics, 2005. **122**(9): p. 13.
137. Yanai, T., D.P. Tew, and N.C. Handy, *A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP).* Chemical Physics Letters, 2004. **393**(1-3): p. 51-57.

138. Neugebauer, J., O. Gritsenko, and E.J. Baerends, *Assessment of a simple correction for the long-range charge-transfer problem in time-dependent density-functional theory.* Journal Of Chemical Physics, 2006. **124**(21): p. 11.
139. Lange, A. and J.M. Herbert, *Simple methods to reduce charge-transfer contamination in time-dependent density-functional calculations of clusters and liquids.* Journal of Chemical Theory and Computation, 2007. **3**(5): p. 1680-1690.
140. Bernasconi, L., M. Sprik, and J. Hutter, *Hartree-Fock exchange in time dependent density functional theory: application to charge transfer excitations in solvated molecular systems.* Chemical Physics Letters, 2004. **394**(1-3): p. 141-146.
141. Osted, A., et al., *Linear response properties of liquid water calculated using CC2 and CCSD within different molecular mechanics methods.* Journal Of Physical Chemistry A, 2004. **108**(41): p. 8646-8658.
142. Coutinho, K., M.J. De Oliveira, and S. Canuto, *Sampling configurations in Monte Carlo simulations for quantum mechanical studies of solvent effects.* International Journal Of Quantum Chemistry, 1998. **66**(3): p. 249-253.
143. Kwit, M., et al., *Determination of absolute configuration of conformationally flexible cis-dihydrodiol metabolites: Effect of diene substitution pattern on the circular dichroism spectra and optical rotations.* Chirality, 2008. **20**(5): p. 609-620.
144. Mort, B.C. and J. Autschbach, *A pragmatic recipe for the treatment of hindered rotations in the vibrational averaging of molecular properties.* Chemphyschem, 2008. **9**(1): p. 159-170.
145. Neugebauer, J., *Induced chirality in achiral media - How theory unravels mysterious solvent effects.* Angewandte Chemie-International Edition, 2007. **46**(41): p. 7738-7740.
146. Mori, T., S. Grimme, and Y. Inoue, *A combined experimental and theoretical study on the conformation of multiarmed chiral aryl ethers.* Journal of Organic Chemistry, 2007. **72**(18): p. 6998-7010.
147. Rossi, S., et al., *Specific anion effects on the optical rotation of alpha-amino acids.* Journal of Physical Chemistry B, 2007. **111**(35): p. 10510-10519.
148. Mori, T., Y. Inoue, and S. Grimme, *Experimental and theoretical study of the CD spectra and conformational properties of axially chiral 2,2 '-, 3,3 '-, and 4,4 '-biphenol ethers.* Journal of Physical Chemistry A, 2007. **111**(20): p. 4222-4234.
149. Jansik, B., A. Rizzo, and H. Agren, *Ab initio study of the two-photon circular dichroism in chiral natural amino acids.* Journal Of Physical Chemistry B, 2007. **111**(2): p. 446-460.
150. Mort, B.C. and J. Autschbach, *Temperature dependence of the optical rotation of fenchone calculated by vibrational averaging.* Journal of Physical Chemistry A, 2006. **110**(40): p. 11381-11383.
151. Osted, A., et al., *The electronic spectrum of the micro-solvated alanine zwitterion calculated using the combined coupled cluster/molecular mechanics method.* Chemical Physics Letters, 2006. **429**(4-6): p. 430-435.
152. Gaenko, A.V., et al., *Spectral and density functional studies on the absorbance and fluorescence spectra of 2-R-5-phenyl-1,3,4-oxadiazoles and their conjugate acids.* Journal Of Physical Chemistry A, 2006. **110**(28): p. 8750-8757.
153. Russ, N.J. and T.D. Crawford, *Local correlation domains for coupled cluster theory: optical rotation and magnetic-field perturbations.* Physical Chemistry Chemical Physics, 2008. **10**(23): p. 3345-3352.

154. Polavarapu, P.L., *Why is it important to simultaneously use more than one chiroptical spectroscopic method for determining the structures of chiral molecules?* Chirality, 2008. **20**(5): p. 664-672.
155. Krykunov, M., et al., *Calculation of the magnetic circular dichroism B term from the imaginary part of the Verdet constant using damped time-dependent density functional theory.* Journal Of Chemical Physics, 2007. **127**(24): p. 16.
156. Crawford, T.D., M.C. Tam, and M.L. Abrams, *The problematic case of (S)-methylthiirane: electronic circular dichroism spectra and optical rotatory dispersion.* Molecular Physics, 2007. **105**(19-22): p. 2607-2617.
157. Ferrighi, L., L. Frediani, and K. Ruud, *Degenerate four-wave mixing in solution by cubic response theory and the polarizable continuum model.* Journal of Physical Chemistry B, 2007. **111**(30): p. 8965-8973.
158. Autschbach, J., *Density functional theory applied to calculating optical and spectroscopic properties of metal complexes: NMR and optical activity.* Coordination Chemistry Reviews, 2007. **251**(13-14): p. 1796-1821.
159. Jiemchooroj, A. and P. Norman, *Electronic circular dichroism spectra from the complex polarization propagator.* Journal Of Chemical Physics, 2007. **126**(13): p. 7.
160. Krykunov, M. and J. Autschbach, *Calculation of static and dynamic linear magnetic response in approximate time-dependent density functional theory.* Journal Of Chemical Physics, 2007. **126**(2): p. 12.
161. Mori, T., Y. Inoue, and S. Grimme, *Quantum chemical study on the circular dichroism spectra and specific rotation of donor-acceptor cyclophanes.* Journal of Physical Chemistry A, 2007. **111**(32): p. 7995-8006.
162. Pinkerton, F.E., et al., *Hydrogen Desorption exceeding ten weight percent from the new quaternary hydride Li3BN2H8.* Journal Of Physical Chemistry B, 2005. **109**(1): p. 6-8.
163. Gaarenstroom, S.W., et al., *Characterization of indium-tin-oxide films with improved corrosion resistance.* Surface and Interface Analysis, 2005. **37**(4): p. 385-392.
164. Meisner, G.P., et al., *Study of the lithium-nitrogen-hydrogen system.* Journal of Alloys and Compounds, 2005. **404**: p. 24-26.
165. Balogh, M.P., et al., *Crystal structures and phase transformation of deuterated lithium imide, Li2ND.* Journal of Alloys and Compounds, 2006. **420**(1-2): p. 326-336.
166. Ahlrichs, R., et al., *Turbomole*. 2004, COSMOlogic GmbH & Co.: Leverkusen, Germany.
167. Sousa, S.F., P.A. Fernandes, and M.J. Ramos, *General performance of density functionals.* Journal of Physical Chemistry A, 2007. **111**(42): p. 10439-10452.