Solvent effects on the electronic spectra of indoles : theoretical methods and laser induced fluorescence excitation in supersonic jet
by Pedro Luis Muino Montero

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry
Montana State University

Abstract:
The effect of solvent on the first two π<—π excited states of indole and 3-methylindole, (1La and 1Lb), has been studied using a combination of hybrid theoretical methods and experimental techniques. A procedure that couples molecular dynamics with semiempirical molecular orbital methods has yielded information about the mechanism and time scales involved in bulk solvent reorganization after excitation of the indole molecule to the 1La, (or 1Lb), state. The fluorescence red shifts predicted in several solvents, (water, butanol, methanol and dimethylether), agree reasonably well with experimental values. These time resolved calculations also indicate that the solvent relaxation has two components: The first one is inertial in character, with a Gaussian shape having a half-width at half-maximum of ~15 fs for water, and of 100-300 fs for the other, (larger), solvents. The second component shows an exponential decay behavior and seems related to the longitudinal relaxation time of the solvent. The correlation times for this component are -170'fs for water, and a few picoseconds for the larger solvents. By coupling molecular mechanics with semiempirical molecular orbital methods, information was obtained about the relative populations and spectroscopic properties of the different conformers of tryptophan at 300 K. We also present very detailed one-photon excitation spectra of indole complexed with H2O, D2O, and methanol, along with an analysis of the intermolecular vibrational modes of the two types of complexes present in each spectrum. The two-photon excitation spectra of selected peaks helped in determining the 1Lb or 1La character of a few transitions. The origin of the first of the complexes has been assigned as 1Lb and the putative origin of the 1La manifold has a polarization ratio compatible with 1La character. Analysis of the ring vibrational modes associated with the other complex suggests that its origin also has 1Lb character.

SOLVENT EFFECTS ON THE ELECTRONIC SPECTRA OF INDOLES: THEORETICAL METHODS AND LASER INDUCED FLUORESCENCE EXCITATION IN SUPERSONIC JET

by

Pedro Luis Muiño Montero

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

November, 1993

SOLVENT EFFECTS ON THE ELECTRONIC SPECTRA OF INDOLES: THEORETICAL METHODS AND LASER INDUCED FLUORESCENCE EXCITATION IN SUPERSONIC JET

Pedro Luis Muiño Montero

Advisor: Dr. Patrik R. Callis

Montana State University
1993

Abstract

The effect of solvent on the first two $\pi^* \leftarrow \pi$ excited states of indole and 3-methylindole, (^1L_a and ^1L_b), has been studied using a combination of hybrid theoretical methods and experimental techniques. A procedure that couples molecular dynamics with semiempirical molecular orbital methods has yielded information about the mechanism and time scales involved in bulk solvent reorganization after excitation of the indole molecule to the ^1L_a, (or ^1L_b), state. The fluorescence red shifts predicted in several solvents, (water, butanol, methanol and dimethylether), agree reasonably well with experimental values. These time resolved calculations also indicate that the solvent relaxation has two components: The first one is inertial in character, with a Gaussian shape having a half-width at half-maximum of ~15 fs for water, and of 100-300 fs for the other, (larger), solvents. The second component shows an exponential decay behavior and seems related to the longitudinal relaxation time of the solvent. The correlation times for this component are ~170 fs for water, and a few picoseconds for the larger solvents. By

coupling molecular mechanics with semiempirical molecular orbital methods, information was obtained about the relative populations and spectroscopic properties of the different conformers of tryptophan at 300 K. We also present very detailed one-photon excitation spectra of indole complexed with H_2O, D_2O, and methanol, along with an analysis of the intermolecular vibrational modes of the two types of complexes present in each spectrum. The two-photon excitation spectra of selected peaks helped in determining the ^1L_b or ^1L_a character of a few transitions. The origin of the first of the complexes has been assigned as ^1L_b and the putative origin of the ^1L_a manifold has a polarization ratio compatible with ^1L_a character. Analysis of the ring vibrational modes associated with the other complex suggests that its origin also has ^1L_b character.

APPROVAL

of a thesis submitted by

Pedro Luis Muiño Montero

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

16 NOV 93
Date

Patrick R Callis
Chairperson, Graduate Committee

Approved for the Major Department

Nov. 16, 1993
Date

David M Dooley
Head, Major Department

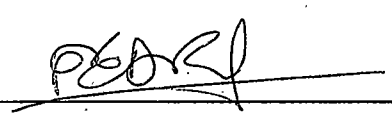
Approved for the College of Graduate Studies

11/21/93
Date

R L Brown
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation for sale in and from microform or electronic format, along with the right to reproduce and distribute my abstract in any format in whole or in part."

Signature 

Date NOV 16/93

To those closest to me: mamá, papá, Toni, madriña, padriño, and Penny.

ACKNOWLEDGMENTS

I wish to extend my gratitude to certain individuals who helped me along the way towards completion of my graduate work.

I want to thank Dr. Patrik Callis for his unselfish supervising of my education and for his concern about my well-being as a person. Dr. Lee Spangler has provided me with his expertise in many aspects of my work with laser and supersonic jet.

The rest of the members of my graduate committee, Dr. Reed Howald, Dr. Edward Dratz, Dr. Arnold Craig, Dr. Jarvis Brown, and Dr. Ronald Johnson, have dedicated their time and efforts to this project.

Mr. Gonzalo Fajardo, Fr. Antonio Iglesias, and Dr. Xosé Luis Armesto must be credited for shaping and enhancing my interest for science along my years in school.

I would like to acknowledge Honeywell Europe and Honeywell España, and, specifically, Mr. Angel Hermosa, for their help and partial financial support during my stay at Montana State University.

Finally, I want to thank my fellow graduate students, my friends, and my wife, Penny, for their help, support, and friendship. Because of them, these years in a foreign land have been full of warmth.

TABLE OF CONTENTS

TABLE OF CONTENTS-Continued

TABLE OF CONTENTS-Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES-Continued

LIST OF FIGURES-Continued

LIST OF FIGURES-Continued

ABSTRACT

The effect of solvent on the first two $\pi^* \leftarrow \pi$ excited states of indole and 3-methylindole, (^1L_a and ^1L_b), has been studied using a combination of hybrid theoretical methods and experimental techniques. A procedure that couples molecular dynamics with semiempirical molecular orbital methods has yielded information about the mechanism and time scales involved in bulk solvent reorganization after excitation of the indole molecule to the ^1L_a, (or ^1L_b), state. The fluorescence red shifts predicted in several solvents, (water, butanol, methanol and dimethylether), agree reasonably well with experimental values. These time resolved calculations also indicate that the solvent relaxation has two components: The first one is inertial in character, with a Gaussian shape having a half-width at half-maximum of ~15 fs for water, and of 100-300 fs for the other, (larger), solvents. The second component shows an exponential decay behavior and seems related to the longitudinal relaxation time of the solvent. The correlation times for this component are ~170 fs for water, and a few picoseconds for the larger solvents. By coupling molecular mechanics with semiempirical molecular orbital methods, information was obtained about the relative populations and spectroscopic properties of the different conformers of tryptophan at 300 K. We also present very detailed one-photon excitation spectra of indole complexed with H_2O, D_2O, and methanol, along with an analysis of the intermolecular vibrational modes of the two types of complexes present in each spectrum. The two-photon excitation spectra of selected peaks helped in determining the ^1L_b or ^1L_a character of a few transitions. The origin of the first of the complexes has been assigned as ^1L_b and the putative origin of the ^1L_a manifold has a polarization ratio compatible with ^1L_a character. Analysis of the ring vibrational modes associated with the other complex suggests that its origin also has ^1L_b character.

INTRODUCTION

The indole molecule has been studied[1-7] in detail in order to get information about the amino acid tryptophan, (see Figure 1), which is the main contributor to the UV absorption and fluorescence of proteins[4,8-10]. The spectrum of indole and derivatives, however, is complicated by the presence of two near degenerate $\pi^* \leftarrow \pi$ transitions[11] labelled ^1L_a and ^1L_b[12].

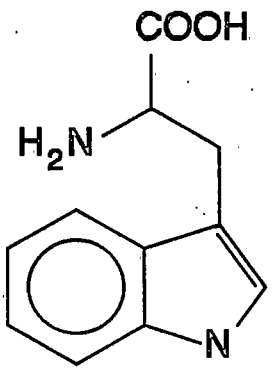


Figure 1. Tryptophan

The work of Valeur and Weber[6], in 1977, was successful in applying fluorescence polarization techniques to separate the ^1L_a and ^1L_b contributions to the excitation spectrum of indole in propylene glycol at 215 K. As a result of this work, more information about the nature of these two states was obtained: the ^1L_a state presents a broad, featureless spectrum, possibly because of a Franck-Condon progression stemming from a geometry displaced from that of the ground state; while the ^1L_b state results in a

sharp band as a result of a geometry similar to that of the ground state.

Early studies by Strickland *et al.*[4] suggested that the nature of the solvent, and substitution at the 3 position, (see Figure 2 for numbering of the ring), affected the shape of the spectra. More thorough studies indicated that indeed substitution[4,8,9,13], solvent effects[4,8,9,13-15], and charge effects[16] affect the ^1L_a and ^1L_b states differently, thus leading to changes in the spectra.

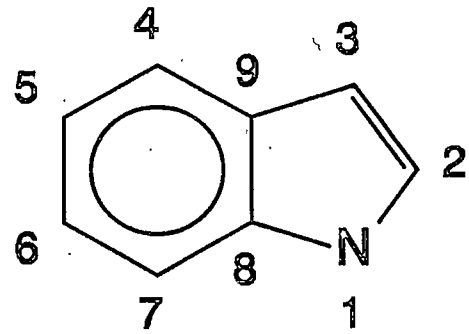


Figure 2. Numbering of Indole

Therefore, it was understood that differences in the electronic spectra of proteins could be analyzed in order to get information on the environment surrounding tryptophan inside a protein[14,17,18,19,20], yielding information about its tertiary structure. A better understanding of the indole chromophore is necessary to accomplish this task. In this respect, Lautié *et al.*[7] presented the first infrared and Raman spectra of indole, which resulted in a complete study of its ground state vibrational frequencies. Lami and Glasser[13] performed a very detailed analysis of the effect of substitution and solvent on the ^1L_a and ^1L_b absorption and ^1L_a fluorescence shifts.

When the technique of laser excitation in supersonic jet was first applied to

indole[21] and to indole+solvent[22,23] it was clear that it would prove useful to get information about the vibrational frequencies of the excited[24-27] and ground states[27,28] for different indole derivatives. However, through this period, a very important question, that of the location of the ^1L_a origin of indole, remained unanswered. Wallace *et al.*[27,29] and Sulkes *et al.*[30] suggested that the ^1L_a transitions could not be seen in supersonic jet because this state is dissociative, leading to a breaking (probably) of the N-H bond. It is possible, however, that the ^1L_a origin was mistakenly identified as a ^1L_b vibrational mode, (or combination of), because the fluorescence lifetime would be similar to those typical of the ^1L_b state. The fact that indoles present normally single state emission[31] can be responsible for this behavior, since the collected fluorescence would come from S_1, independently of which state was excited.

On a parallel course, Callis *et al.* had been successful in applying the technique of two-photon spectroscopy to resolve the ^1L_a and ^1L_b bands of indole and benzimidazole[32], methylindoles[33], and other aromatic systems[34]. In 1990, this technique was used in combination with a supersonic expansion for indole[35], and succeeded in yielding the first positive identification of ^1L_a transitions in jet. Callis *et al.* reported that the transitions at 455 and 480 cm^{-1} to the blue of the ^1L_b 0-0 have ^1L_a character. As the spacing between them, ~25 cm^{-1}, was too small to correspond to a vibrational mode of the ^1L_a state, it was suggested that the ^1L_a origin is split, possibly as the result of a coupling with a nearby dark ^1L_b state.

The effect of solvent dynamics has also been subject to a detailed study during the past years. Specifically, it is now well understood how the solvent affects electron transfer reactions[36-38]. Fleming *et al.*[39-41] and Barbara *et al.*[42,43] have pioneered work

on subpicosecond spectroscopy which has resulted in a much better understanding of molecular phenomena and solvent reorganization at time scales as short as 50 fs[44] for coumarin dyes in water. These experiments have been coupled to numerous calculations on the behavior of different molecules in polar solvents[45-51]. This work has helped in establishing that the temporal response of the solvent is dominated by solvent properties, but the magnitude of the shift clearly depends on the difference of the permanent dipoles of ground and excited states[51].

It should also be pointed out that self consistent reaction field methods are appearing as of lately[52-55]. In these methods, the solvent is modelled as a dielectric continuum. They are useful for predicting, in an easy way, equilibrium solvation energies.. However, they cannot predict good spectroscopic information for different solvents because the dielectric continuum formulation is not sensitive to changes in dielectric constant when $\varepsilon>10$, (see RESULTS AND DISCUSSION). Lami and Glasser[13] have measured different shifts for indoles in five different solvents with dielectric constant larger than 10, (ranging from 17 to 80). They found that each solvent induced a different shift, contrary to what the self consistent reaction field methods would predict.

The question of which mechanism is responsible for the red shift induced by solvation has often been addressed in the literature. A very interesting theory explains the fluorescence red shift by means of *exciplex* formation. Exciplexes are excited-state solute-solvent complexes. Lumry *et al.*[2,5] suggested that exciplex formation was responsible for a large red shift and a loss of vibrational structure in the fluorescence spectra of indole in polar solvents. They reported a solute-solvent stoichiometry of 1:2 and 1:1 for associating and non-associating solvents and dismissed the possibility of

hydrogen bonding between the N-H group and the solvent. One important characteristic of this work is that the spectra were taken in indole solutions in non-polar solvents containing small amounts of polar solvents. Thus, it is conceivable that the formation of the exciplex was induced by the presence of the non-polar environment which would tend to enhance the degree of association of the polar components.

The idea of exciplex formation has been used by Tubergen and Levy[26] and Sulkes and Arnold[30] to explain the broad and red shifted emission of the second of the complexes that indole forms with water in a supersonic jet. They have stated that the first complex is hydrogen bonded to the N-H group, while the second is located over the π cloud of the ring, producing an effect similar to the one seen in solution[2]. The lifetime results of the fluorescence taken by Sulkes and Arnold are compatible with the idea that a second state is emitting. This would indicate that, upon complexation, the ^1L_a state origin becomes nearly degenerate with the ^1L_b origin and emission from both states follows. Presumably, the broad, red shifted emission would be caused by the ^1L_a state.

Statement of the Problem

Our goal in this research is to shed some light on the problem of whether the assigned ^1L_a origin is truly ^1L_a or is a pair of ^1L_b transitions borrowing ^1L_a character through Herzberg-Teller coupling. We will also investigate the effect of the solvent on the ground and electronically excited states. To this end, we will be using two approaches which are going to yield information about two different aspects of the phenomenon.

Molecular dynamics coupled with semiempirical molecular orbital methods can yield good results about the behavior of indole in a bulk solution. The time scales involved in molecular dynamics, (femtoseconds to picoseconds), allow one to follow the dynamics of the indole molecule in the instants following excitation, when the solvent molecules are responding to the changes in the electric field generated by the solute. This will allow us to predict fluorescence shifts in solution. Comparison of results in different solvents will also provide information about the effect of the polarity of the environment and the size of the solvent molecules. Direct inspection of the solvent reorientation in response to the changing electric field generated by the solute will be very useful in helping to determine the mechanism responsible for the red shift.

On the other hand, we will apply the techniques of one- and two-photon laser induced fluorescence in supersonic jet to obtain information on how a single solvent molecule, forming a van der Waals complex with indole, affects the electronic states. We can use one-photon spectroscopy to get a very detailed picture of the vibrational frequencies, (both intra- and intermolecular), of the complexes using different solvents, and two-photon spectroscopy to obtain information on the nature of some of the transitions in the electronic spectra. This information can be coupled to that obtained by analysis of the differential shifts on the ^1L_a and ^1L_b peaks induced by complexation.

Comparison of these two techniques, which represent the many molecules as well as the single molecule limit, will provide information that will help in answering the questions of how, in which way, how fast, and how much the solvent molecules affect the indole chromophore.

BACKGROUND

Theory

This section contains a quick overview of the methods we are going to use in our calculations. It is our intention to explain their purpose, advantages and drawbacks, without giving an in-depth explanation of the theory behind them.

Molecular Orbital Methods

In general, quantum mechanics methods try to solve the Schrödinger equation, (see Equation 1), in order to obtain the associated eigenvalues and eigenfunctions.

$$\hat{H}\Psi_n = E_n\Psi_n \qquad (1)$$

However, when three or more particles, (including nuclei and electrons), are present, Equation 1 does not have an analytical solution, and approximate methods are necessary. The reason for this impossibility to solve the equation stems from the presence of electron repulsion terms, (of the form of $1/r_{12}$), in the Hamiltonian of the system.

One way to write a one-electron Hamiltonian which includes the repulsion between that electron and the others consists of explicitly using the one-electron wavefunctions when building $\hat{H}$. However, this introduces the problem of having to know $\hat{H}$ to get Ψ, and having to know Ψ in order to have $\hat{H}$. To get around this situation, the best way to proceed is to start with a *guess* wavefunction, $\psi^{(0)}$, and use it to construct an approximate

Hamiltonian, $\hat{H}^{(0)}$. Now, we can solve the Schrödinger equation and obtain a better wavefunction, $\psi^{(1)}$, which is used to obtain a better Hamiltonian, $\hat{H}^{(1)}$. This process is repeated until self-consistency, that is, until the wavefunction obtained from the Schrödinger equation is the same that was used to build the Hamiltonian, (that is, $\psi^{(n)}=\psi^{(n+1)}$). Obviously, the better $\psi^{(0)}$, the faster the method will converge.

The wavefunctions obtained and used in the Self Consistent Field (SCF) method are multielectron wavefunctions and, in their simplest form, can be thought of as products of one-electron spin-orbitals[56]. However, one of the postulates of Quantum Mechanics states that an acceptable wavefunction must be antisymmetric with respect to the exchange of electrons. To this effect, Slater[57] showed that one can assure antisymmetry with respect to the exchange of electrons by writing the wavefunction as a determinant whose elements are one-electron spin-orbitals. For instance, the multielectron wavefunction for the helium atom in its ground state can be expressed as

$$\Psi = \frac{1}{\sqrt{2}}\begin{vmatrix} 1s(1)\alpha(1) & 1s(2)\alpha(2) \\ 1s(1)\beta(1) & 1s(2)\beta(2) \end{vmatrix} \qquad (2)$$

The SCF method cannot reach the true energy of the system because the electronic motion is not correlated. Each electron is seeing the other electrons as a time-averaged charge cloud, but there is no penalty built in the wavefunction preventing all the electrons from simultaneously occupying the same point in space. The minimum energy that can be obtained without including electron correlation is called the Hartree-Fock limit. Thus, a Slater determinant is an approximate solution of the Schrödinger equation, but it cannot

be used, by itself, to reach the true energy of the system.

To go beyond this limit, one has to include electron correlation. There are several ways to do so. The method most widely used is called Configuration Interaction, (CI). In this method, the wavefunctions are built as linear combinations of Slater determinants corresponding to different configuration of the system, (see Equation 3),

$$\Psi = \sum_i c_i D_i \qquad (3)$$

where the D_is are Slater determinants.

The configurations previously mentioned correspond to the ground state of the molecule, (1 configuration), and as many excited state configurations as one desires. The latter ones correspond to promotion of one electron to an unoccupied orbital, (singly excited configuration), of two electrons to two unoccupied orbitals, (doubly excited configuration), et cetera.

With the inclusion of CI, it is possible to reach the exact energy limit. The difference between this and the Hartree-Fock limit is called the correlation energy, and it is rather large. For instance, for helium, it is on the order of 25 kcal/mol[58]. In our calculations, we are going to include CI, because it helps in better reproducing the spacing between the ^1L_a and ^1L_b states of indole, as shown by Callis[59].

Approximate Methods: INDO/S

With the use of computers, every integral stemming from the Schrödinger equation can be computed, even for fairly large systems. This is the basis for the *ab initio*

methods. These, however, are computationally intensive, because the number of integrals to be evaluated grows as the fourth power of the number of the basis function used.

If the molecule is large, the number of integrals extended over four nuclei is very large. Clementi and Mehl[60] have shown that many of these four center integrals are almost negligible. It would be interesting to devise a method that allows one to selectively neglect these integrals without having to spend time in calculating them. There are several considerations to take care of. For instance, if all the integrals neglected were very small, but all of them added together had a fairly large value, it would not be very wise to neglect them. Another problem to be addressed is that of relating the process of neglecting integrals to the basis function in a simple way. If one is not careful, different energies can be obtained for a molecule if its orientation changes.

Pople and coworkers[61-65] have developed a series of methods that take care of these and other problems and yield good results when calculating different molecular properties. In the first paper of the series[61], Pople *et al.* introduced 5 approximations in the process of calculating the integrals. This method, called CNDO, (Complete Neglect of Differential Overlap), is the simplest of a group of methods based on the concept of Zero Differential Overlap, (see reference 56). Normally, these methods are also parametrized, indicating that scaling factors or parameters are included in the program so it yields good results for certain molecular properties.

The method we are going to use in our calculations is called INDO/S, (Intermediate Neglect of Differential Overlap, calibrated for Spectroscopy), due to Zerner *et al.*[66-69]. Ridley and Zerner[66] preferred the INDO over the CNDO formulation because

the latter does not include the one-center exchange integrals which are required to separate different terms within a configuration, (like singlets and triplets in $\pi^* \leftarrow n$ transitions). At any rate, the program allows one to choose between the CNDO and the INDO formulations, (and also between different *versions* of them, corresponding to different parametrizations). The method has been widely tested, yielding good results for aromatic molecules[70,71], aromatic molecules in a crystal field[72], and, more specifically, for indole[59].

Zerner's program allows the selection of different types of electron repulsion parameters, like Mataga-Nishimoto[73], Ohno-Klopman[74], Pariser[75], and theoretical parameters. A general formula for the electron repulsion between two atoms A and B, at a distance R_{AB} is given by Callis *et al.*[59,71]

$$\gamma_{AB} = \left(R_{AB}^{n} + \left[0.5(\gamma_A + \gamma_B)\right]^{-n}\right)^{-1/n} \qquad (4)$$

In the Mataga-Nishimoto scheme, n=1 and R_{AB} is divided by 1.2[66]. Meanwhile, n=2 in the Ohno-Klopman formulation. As pointed out by Callis[59], when n increases, the splitting between ^1L_a and ^1L_b decreases because γ_{AB} decreases more gradually with R_{AB}. Zerner's program also permits one to choose scaling factors for the interaction between molecular orbitals, (see METHODS section).

As stated in the previous section, CI can be used to include the electron correlation. The program allows for the inclusion of up to 210 configurations. Zerner *et al.*[66,76,77] have shown that the use of singly excited configurations with Mataga-Nishimoto electron repulsion yields results in good agreement with experiments

for benzene and related molecules. This procedure was previously tested by Jaffé *et al.*[78] for CNDO. Callis[59], working on a suggestion by Zerner, has shown that this procedure is not the only one to effectively reproduce experimental results. He obtained good agreement by using doubly excited configurations along with Ohno-Klopman repulsion, because, although the doubly excited configurations bring the ^1L_b state farther down than the ^1L_a, the use of Ohno-Klopman repulsion, (and different interaction scaling factors), reduces the gap between the two states and compensates for the previous effect.

In our work, we chose to follow the first scheme, since it is easier to implement, especially when working with thousands of different geometries, as in our case. To that effect, (see METHODS section), we used a total of 14×14=196 excited configurations; that is, those formed from promotion of one electron from any of the 14 highest occupied to any of the 14 lowest unoccupied molecular orbitals.

Molecular Mechanics

Molecular Mechanics differs from molecular orbital methods in that no wavefunctions are sought. The idea behind molecular mechanics is that an appropriate potential energy function for the system, (where each atom is basically treated as a point charge), can be calculated and the geometry optimized towards the minimum of such potential energy.

Its expression is a rather complicated one, and can depend on a lot of parameters. For instance, a very simple expression for the potential energy for CO_2 could be as the one shown in Equation 5:

$$V = \frac{1}{2}k_{CO}(r_1 - r_{CO})^2 + \frac{1}{2}k_{CO}(r_2 - r_{CO})^2 + \frac{1}{2}k_{OCO}(\theta - \theta_{OCO})^2 \qquad (5)$$

where k_{CO} and k_{OCO} are force constants, and r_{CO} and θ_{OCO} are values for the standard length and angle of such bond types. These magnitudes must be input into the program, (or the program has to store them). It must be pointed out that the expression for V in Equation 5 depends on r_i, (intermolecular distances), so it can be minimized.

This expression for V, however, is extremely simplified. As an example, the Central Valence ForceField, (CVFF)[79], includes up to eleven different terms, (Equation 5 includes only one term for bond distances and one for bond angles). The first four of those eleven are what is normally called *diagonal terms*. They include the energy of deformation of bond lengths, angles, torsional angles, and out of plane interactions. The other seven terms correspond to the *off-diagonal terms* and represent couplings of the diagonal deformations.

It is clear from Equation 5 that this method requires heavy parametrization, because each term needs at least two parameters for each type of bond included. CVFF has parameters for 50 different types of atoms and ions, thus, the number of bonds, angles, and dihedrals is going to be rather large.

Another widely used forcefield is the AMBER forcefield, due to Kollman *et al.*[80,81], which has six terms. The first three handle the deformations of bonds, angles, and dihedral angles, the fourth and fifth terms correspond to the van der Waals and electrostatic interactions, and the sixth corrects the electrostatic interaction for a hydrogen bond.

One important feature of molecular mechanics relies on the algorithm used in the minimization. There is a wide variety of algorithms, but we will mention here only the two that will be used in our calculations.

The *steepest descents* algorithm is the simplest of the methods. It just calculates the gradient along each coordinate and follows that path. After each step, the previous direction is substituted by the newly calculated, and the process is repeated until the convergence criterium has been met. The disadvantage of this method is that it is rather slow, especially when the system approaches a minimum. However, the advantage this method has is that it does converge even when far from a minimum. This is crucial in calculations involving many atoms, where the starting geometry is normally not very good. Other methods do not converge to a minimum in this case, but *steepest descents* does, even if the system does not behave harmonically.

The other method we used is the *Quasi-Newton-Raphson* algorithm, also known as the *variable metric method*, developed by Fletcher and Powell[82]. It is based on the *Newton-Raphson* method, which, besides using gradient information, uses the Hessian matrix calculated from second derivatives, (curvature), to predict where the function will change directions, (which indicates that it will pass through a minimum). The *Quasi-Newton-Raphson* method calculates the Hessian matrix numerically from the first derivatives, rather than analytically from second derivatives, and, therefore, it is much faster. This method is not as forgiving of bad initial structures as *steepest descents* is, thus, it is most suitable to be used in small systems where the initial geometry is not that far from a local minimum.

Molecular Dynamics

Molecular Dynamics can be considered a generalization of molecular mechanics in the sense that the system not only has a potential energy term but is also given kinetic energy and is allowed to change in time. The same potential energy term that was mentioned in the previous section is used, but the system does not evolve along the line that minimizes the gradients. Rather, the gradients of the potential energy along each coordinate correspond to the force applied to the particles according to Newton's law,

$$F_i = -\frac{\partial V}{\partial r_i} = m_i \frac{\partial^2 r_i}{\partial t_i^2} = m_i a_i \qquad (6)$$

In trying to solve this equation, we face the same problem that was faced when trying to solve the Schrödinger equation, namely that there is no closed form solution for a system of three or more particles. Thus, a numerical solution has to be found, and the new position of each particle has to be approximated by a Taylor expansion,

$$r(t+\Delta t) = r(t) + \frac{\partial r}{\partial t}\Delta t + \frac{\partial^2 r}{\partial t^2}\frac{\Delta t^2}{2} + \frac{\partial^3 r}{\partial t^3}\frac{\Delta t^3}{6} + \ldots \qquad (7)$$

To calculate the new position, it is necessary to know the original position, velocity, and acceleration, and it is also required to know the value for higher order terms. As a result of the approximation made here, there arises a limitation on the time step used. The best results for systems like ours are obtained when $\Delta t \leq 1$ fs, because the integration algorithms require that velocities and accelerations are virtually constant over

the time step.

Another important issue here is that of temperature control. The temperature is initialized by assigning velocities to the ensemble of molecules, according to the Maxwell-Boltzmann equation,

$$f(v)dv = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) 4\pi v^2 dv \qquad (8)$$

which expresses the probability for a molecule of mass m to have a velocity v at a temperature T, (k is the Boltzmann constant). When dynamics is started, (initialization period), the velocities are scaled at every step to instantaneously achieve the desired temperature. However, in later stages of the calculation, the scaling has to be much smoother, in order to not introduce artifacts that can affect the results. The best method to achieve the scaling has been developed by Berendsen and coworkers[83]. Each calculated velocity is scaled by a factor λ, given by Equation 9.

$$\lambda = 1 + \frac{\Delta t}{2\tau}\left(\frac{T_0}{T} - 1\right) \qquad (9)$$

where T_0 is the desired temperature, T is the temperature corresponding to the distribution of velocities, and τ is the characteristic relaxation time, which can be adjusted to obtain better results.

A big drawback of both molecular mechanics and dynamics consists on the fact

that all the calculations are purely classical. Thus, there are some systems that cannot be accurately described, like electron or proton transfer, or bond breaking or formation, since they have to be treated by means of quantum mechanics methods.

Periodic Boundary Conditions

In trying to simulate a liquid solution, a common problem is that of finding an adequate geometry for the system. The most obvious solution would consist on surrounding the solute with a drop of solvent. However, this is not an accurate description, because it is actually equivalent to a *tiny* sphere of a few angstroms in diameter surrounded by vacuum. This is the reason why periodic boundary conditions, (PBC), are used. In this arrangement, the solute is surrounded, not by a sphere, but by a cube of solvent. This cube is reproduced by the program, and an "infinite" number of these reproductions, (*ghosts*), are stacked alongside the original in all directions of space. Therefore, the molecules in the original surface do not see a vacuum surrounding them. They actually see other solvent molecules, as they would in nature. This brings the problem of reducing the number of molecular interactions to a manageable size, which is taken care of by setting a cut-off distance where the interactions are no longer felt. Should a ghost molecule return inside the cut-off distance, its presence would be felt by the real molecule once again. When cut-off distances are chosen properly, MD runs even slightly faster using PBC than using a drop of solvent, (because some of the solvent-solvent interactions across the box are neglected while all of them are computed using a drop).

Another advantage of PBC is that the number of solvent molecules in the vicinity

of the solute is always the same. When using a drop, if the original geometry is strained, (especially for large systems), some of the solvent molecules can leave the system, acting as if they were evaporating.

Solvents

The Onsager-Mataga-Lippert Model

The Onsager reaction field[84] is the first useful formulation of the effect of a solute molecule in a medium of dielectric constant ε. The mathematical expression for the dielectric constant is attributed to Debye[85], who stated that

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{4\pi}{3} \sum N\left(\alpha + \frac{\mu^2}{3kT}\right) \qquad (10)$$

where N is the concentration, (in molecules/volume), α and μ are the polarizability and permanent dipole of the molecule, k is the Boltzmann constant, T is the temperature, and the summation refers to the different components of the solvent mixture.

Onsager's model refers to a spherical solute of radius a inside a continuous medium of dielectric constant ε, (see Figure 3). In this model, the solvent polarizability can be expressed in terms of the solute size and the refractive index, (n), of the solvent

$$\alpha = \frac{n^2 - 1}{n^2 + 2} a^3 \qquad (11)$$

The question of which molecular size to use is not a trivial one. Onsager himself

acknowledged that it is not reasonable to expect that the solute fills all the available space and that, therefore, a *void*, where ε=1, should be added to the formulation. In our analyses, whenever *a* is needed in order to calculate some variable, we are going to calculate the molecular volume from density values, (*D*), in crystalline phase, assuming that the shape of the molecule is spherical, and using the molecular mass of the molecule, (*M*). From this size, a *radius* can be calculated, (see Equation 12). Alternatively, using the formulation described below, one can calculate values from spectroscopic results.

$$a = \sqrt[3]{\frac{3M}{4\pi D}} \qquad (12)$$

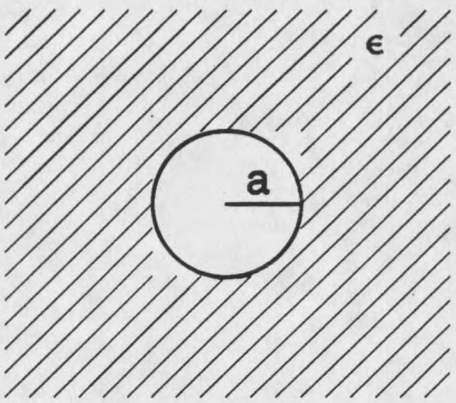


Figure 3. Onsager model: Spherical solute of radius *a* in a continuous solvent of dielectric constant ε

An additional problem in this model is the continuous nature of the solvent. As noted by Frisch *et al.*[86], this model, is not realistic when hydrogen bonds are present. Our calculations include hydrogen bonding between solvent molecules and, at least, partial hydrogen bond between solvent and solute. And, as it will be seen later, we can develop

a good quantitative analysis.

Further work by Ooshika[87], Mataga[88,89], Lippert[90,91], and McRae[92,93] has improved Onsager's model and resulted in useful formulations for the spectral shifts reported in these previous references and in the work of Muiño and Callis[51].

When an indole molecule goes from vacuum to solution in a polar solvent, both the ground and the ^1L_a state are stabilized, although the magnitude of the stabilization is larger for the excited state as a result of the larger dipole moment of this state. The amount of stabilization induced on the ground state is given by Equation 13.

$$hc\Delta\bar{\nu}_{ground} = -\mu_g^2\frac{2}{a^3}\left[\frac{\epsilon-1}{2\epsilon+1}-\frac{n^2-1}{2n^2+1}\right]-\mu_g^2\frac{2}{a^3}\left[\frac{n^2-1}{2n^2+1}\right] \qquad (13)$$

where μ_g is the permanent dipole of the ground state of the solute, *h* is Planck's constant, and *c* is the speed of light in vacuum.

Although the expressions containing f(n) in Equation 13 cancel out, they have been included to explicitly state the contribution of this term to the orientational polarizability of the solvent, (in the first bracket), and to its electronic polarizability, (second bracket).

Both components of the polarizability of the solvent play different roles in an excitation process. When the solute is promoted to the excited state, the solvent does not have time to reorient in the time elapsed in the excitation, (Born-Oppenheimer approximation). Thus, the orientational component of the polarizability remains constant during excitation. On the other hand, the electronic component remains in equilibrium with the solute dipole during the excitation process.