ANALYSIS OF BINDING AFFINITY IN DRUG DESIGN BASED ON AN AB-INITIO APPROACH

A Thesis

by

PABLO FELIX SALAZAR ZARZOSA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2009

Major Subject: Chemical Engineering

ANALYSIS OF BINDING AFFINITY IN DRUG DESIGN BASED ON AN AB-INITIO APPROACH

A Thesis

by

PABLO FELIX SALAZAR ZARZOSA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Jorge Seminario
Committee Members,	Perla Balbuena
	Debjyoti Banerjee
Head of Department,	Michael Pishko

May 2009

Major Subject: Chemical Engineering

ABSTRACT

Analysis of Binding Affinity in Drug Design Based on an Ab-initio Approach.

(May 2009)

Pablo Felix Salazar Zarzosa, B.S., University of Piura, Peru

Chair of Advisory Committee: Dr. Jorge Seminario

Computational methods are a convenient resource to solve drawbacks of drug research such as high cost, time-consumption, and high risk of failure. In order to get an optimum search of new drugs we need to design a rational approach to analyze the molecular forces that govern the interactions between the drugs and their target molecules. The objective of this project is to get an understanding of the interactions between drugs and proteins at the molecular level. The interaction energy, when protein and drugs react, has two components: non-covalent and covalent. The former accounts for the ionic interactions, the later accounts for electron transfer between the reactants. We study each energy component using the most popular analysis tools in computational chemistry such as docking scoring, molecular dynamics fluctuations, electron density change, molecular electrostatic potential (MEP), density of states projections, and the transmission function.

We propose the probability of transfer of electrons (transmission function) between reactants in protein-drug complexes as an alternative tool for molecular recognition and as a direct correlator to the binding affinity. The quadratic correlation that exists between the electron transfer rate and the electronic coupling strength of the

reactants allow a clear distinguishability between ligands. Thus, in order to analyze the binding affinity between the reactants, a calculation of the electronic coupling between them is more suitable than an overall energetic analysis such as free reaction energy.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Seminario, and my committee members, Dr. Balbuena and Dr. Banerjee, for their guidance and support throughout the course of this research.

I also want to extend my gratitude to the Defense Threat Reduction Agency (DTRA) and the Army Research Office (ARO) for the support of this project.

TABLE OF CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

LIST OF TABLES x

1 INTRODUCTION 1

1.1 Reality of drug research 1
1.2 Computer aided drug design 2
1.3 Problem statement 4
1.4 Approach to solve the problem 7

2 METHODS 9

2.1 Binding affinity and the HSAB principle 9
2.2 Electron transfer theory 12
2.3 Docking 14
2.4 Molecular dynamics 15
2.5 Ab-initio: Electron density change 16
2.6 Ab-initio: Molecular electrostatic potential 18
2.7 Ab-initio: Density of states projections 19
2.8 Ab-initio: Transmission function 19

3 TEST CASES 22

3.1 Oxidation and reduction affinities on catalytic surfaces 22
3.2 PIM-1 kinase: importance, known drugs, active site 26

4 RESULTS 28

4.1 Docking calculation 28
4.2 Molecular dynamics simulation 30

Page

4.3 Ab-initio: Electronic binding energy and the HSAB principle results 33
4.4 Ab-initio: Electron density change 35
4.5 Ab-initio: Molecular electrostatic potential 37
4.6 Ab-initio: Density of states projections 39
4.7 Ab-initio: Transmission function 42

5 CONCLUSIONS 43

REFERENCES 45

VITA 50

LIST OF FIGURES

Page

Figure 1 Estimated cost of drugs showing the slower increase rate of preclinical than clinical tests because of the use of computerized methods..... 1

Figure 2 Drug discovery pipeline vs Computer aided drug design. 2

Figure 3 Potential energy surface of reactants and products..... 13

Figure 4 Picture of the test junction Pseudo protein (green) - Platinum (blue) – O2 (red) - Platinum-Pseudo protein. 22

Figure 5 I-V curves for A: G-Pt-HH-Pt-G with H_2 dissociated (blue) and H_2 bonded (pink) showing higher current for H_2 bonded, B: G-Pt-OO-Pt-G with O_2 dissociated (green) and O_2 bounded (red) showing higher current for O_2 dissociated, C: G-Pd-OO-Pd-G (green) and G-Pt-OO-Pt-G (maroon) showing higher current for palladium. 23

Figure 6 Differential heats of oxygen adsorption on 0.5 and 1 wt% Pt/TiO_2 measured at 30 degrees Celsius. 25

Figure 7 Structures of known drugs for the PIM-I kinase 26

Figure 8 Imaging of the characterization of the active site of PIM-I kinase showing a dominant number of negative amino acids on one side of the active site and positive and hydrophobic on the other. 27

Page

Figure 9 Imaging of A) PIM-I Kinase (magenta), an oncogene-encoded serine/threonine kinase primarily expressed in hematopoietic and germ cell lines showing the active site with its ligand (blue). B) Schematic picture of the modeling of the active site by the sphere centers method. 29

Figure 10 RMSD fluctuations of the heat atoms of the neutral set of drugs in the active site along the equilibration step. 31

Figure 11 RMSD fluctuations of the heat atoms of the protonated set of drugs in the active site showing the sensitivity of the system to a change of charges. 32

Figure 12 Electron density change (blue means accumulation, and read means depletion of electrons). 35

Figure 13 Molecular electrostatic potentials (MEPs) 37

Figure 14 Density of States projections of the protein PIM-1 Kinase before (shaded curves) and after the reaction (light curves) with the neutral drugs 40

Figure 15 Density of States projections of the protein PIM-1 Kinase before (shaded curves) and after the reaction (light curves) with the protonated drugs. 41

Figure 16 Transmission functions for the neutral set (left) and for the protonated set of drugs (right). 42

LIST OF TABLES

Page

Table 1 Enthalpy of formation of known drugs obtained from the DOCK program, and experimental values. 29

Table 2 Electronic binding energy and driving force energy 33

Table 3 Total charge transfer from the protein to the drugs (neutral/protonated) 35

1. INTRODUCTION

1.1 REALITY OF DRUG RESEARCH

The discovery of new drugs is one of the most challenging and exciting tasks in science. For example, the average cost out-of-pocket per new drug (linking both the cost of unsuccessful drugs and the ones that obtained marketing approval) in 2000 was $401 million. Capitalizing over the time expended in the project of about 12 years (average time to get a drug to the market), at a rate of 11%, the amount rises to $800 millions [1]. Comparing the costs over the last 20 years, as shown in figure 1, there is an increase in the total cost at an annual rate of 7.4% above the general price inflation. There are several hypotheses for the increase in the clinical test cost: the focus on treatments for chronic and degenerative diseases that are generally more expensive; and the increase of investigated drugs turning in larger and longer recruits of patients at higher costs. Notice that, the slow increasing rate cost in the preclinical tests is mainly because of new computerized discovery technologies such as high throughput screening (HTS) and combinatorial chemistry.

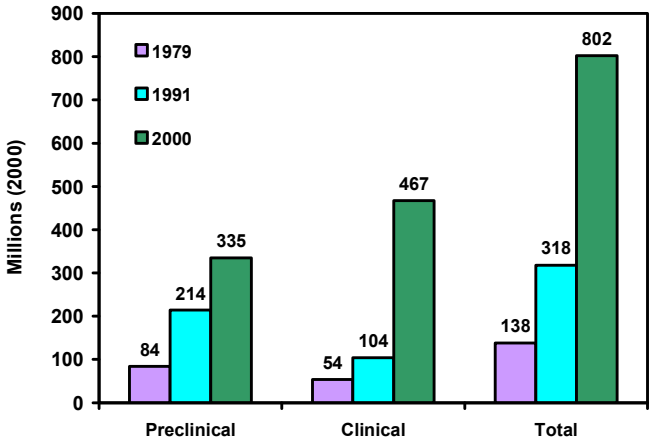


Figure 1: Estimated cost of drugs (Ref [1]) showing the slower increase rate of preclinical than clinical tests because of the use of computerized methods.

This thesis follows the style of *Journal of Physical Chemistry A.*

1.2 COMPUTER AIDED DRUG DESIGN

Efforts, mainly from pharmaceutical companies, for reducing the costs and increasing efficiency of new drugs have made of computational and molecular modeling methods a key step in the drug discovery pipeline as shown in figure 2.

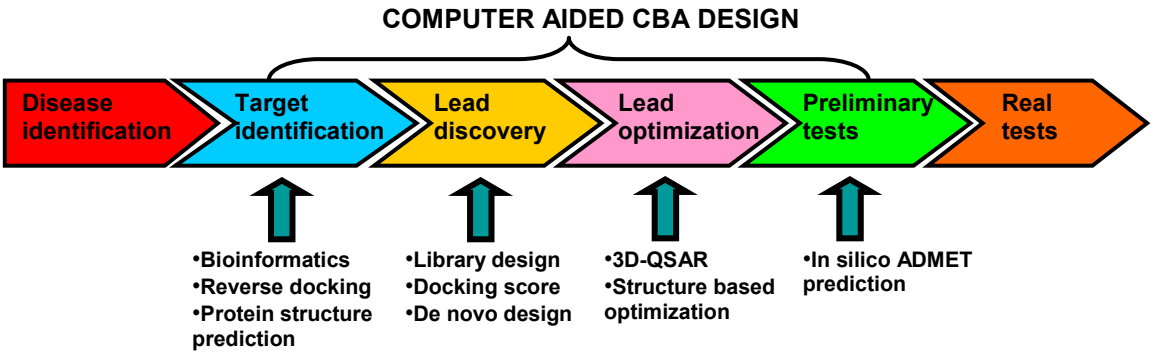


Figure 2: Drug discovery pipeline vs Computer aided drug design[2]. Applications of computational methods in the drug design process.

Computational chemistry and molecular modeling play an important role in the drug discovery process allowing a rational approach for the development of new potential drugs [3]. The first step in the drug discovery process is the target identification of an enzyme, protein or DNA sequence related to the dysfunction. The conventional approach for target identification consists of the comparison of functional genomics from normal and abnormal human's tissues. Using computational bioinformatics tools, we identify the protein that is transcript in the abnormal gene code. Another approach is the reverse docking: an existing drug with known activity is docked it against a protein database; the protein hits go through a homologous analysis to find similar proteins. Finally, the relation of the heat proteins with the disease is validated using experimental procedures.

The next stage is the lead discovery and optimization. Here the first step is the use of combinatorial chemistry and high-throughput screening (HTS) to screen many thousands of biologically active compounds in a short time. The improvement of library strategies, design principles, scaffold design, solid-phase methods and a new approach called dynamic combinatorial chemistry make of combinatorial chemistry a powerful tool for researchers. But, combinatorial chemistry also gives rise to a large number of compounds that are not suitable for synthesis as drug molecules. Therefore, the analysis of large dataset of drugs needs an efficient method for the discrimination of the ligands, resulting in lower number of potential leads. The method also has to be able to recognize, optimize and find new drug compounds by predicting, rationalizing and estimating their properties.

In the next stage of the drug discovery process, we distinguish two cases when the three-dimensional structure of the target protein is not available and when is known. In the first class known as ligand-based drug design, one tries to relate electronic and structural properties with the known activities of the training test of drugs. As an example we have the comparative molecular field analysis (CoMFA) [4], in which the activity of known set of molecules is correlated with a steric and electrostatic field generated along the molecule. In order to generate the field, we have to align the molecule in a 3-dimensional space following a suitable criteria such as maximizing the steric overlap following pharmacophore theories [5,6], or basing on automated field fit methods [7]. Even though the results are promising, the strong dependence of CoMFA on the quality of the alignment rule makes the results unstable. In the line of ligand-based drug design, several quantum mechanics studies have been also reported [8,9].

When the three-dimensional structure of the target protein is known, the analysis of the receptor-ligand interactions is directly done through parametric relations of electrostatic, van der Waals, hydrogen bond, aromatic and hydrophobic interactions.

Computational techniques to predict drug-likeness properties such as absorption, distribution, metabolism, excretion and toxicity (ADMET) have been developed as well. This technique, made of regression equations and/or neural networks trained on experimental data, reduces the costs of late stage preclinical and clinical trials.

1.3 PROBLEM STATEMENT

Computer aid drug design (CADD) is a combined procedure from theoretical and experimental aspects of established sciences as biochemistry, biophysics, chemistry, and computer science. In the first stages of CADD of the drug discovery process, combinatorial chemistry has made available large libraries of potential ligands for a high throughput screening (HTS), but yielding a large number of active compounds unsuitable for chemical synthesis[10]. To reduce the number of active compounds and increase the effectiveness of the drugs we need a rational protocol procedure based on first-principles that maximizes the chances of finding, discrimination and optimization of new drugs. In structure based drug design, inexpensive and fast docking algorithms based on scoring functions derived from force fields are used for a static search of the conformational orientation of the ligands inside of the active site of the receptor molecule. The search goes until an energetic minimum system is found. Having the total energy as a parameter, the set of ligands are ranked.

The many degrees of freedom of protein-drugs systems make the search of a global minimum in the potential energy surface of the receptor-ligand complex a big task. And it turns out to be more complicated since the nuclei fluctuations of the reactants, due to thermal and vibrational energies, help to overcome several barriers of their complex potential energy surface. Furthermore, thermal and vibrational energies strongly influence thermodynamic properties such as free energy, which is conventionally related to the binding affinity of the ligand-receptor. To overcome the

nuclei fluctuations problem, molecular dynamics and Monte Carlo methodologies [11,12], computationally more expensive than docking scoring, and free energy perturbation and thermodynamic integration techniques [13] are used in the determination of the free energy of the reactants. But the accuracy of the free energy calculations (based on force fields parameters) is insufficient to estimate slight energy differences of the ranked drugs. Therefore, the correlation of the free energy with the binding affinity lacks of reliability. Instead of calculating the free energy, one could also analyze the binding affinity based on qualitative properties such as fluctuation distances between the protein and the ligand, or root mean square deviations of the ligands over the trajectory.

Any improvement on the Docking scoring or the MD approaches comes from ab-initio calculations and experimental information. Thus, we need a model for the analysis of binding affinity based on ab-initio methods that help us to explain and quantify the structural and interaction features in a distinguishable and accurate manner. First principles methods solve not only all the nuclear motions: thermal, vibrational and orientational fluctuations; they also consider the electronic structure in an explicit manner. At least in the reactive site, we will have a chemical accuracy of the multiple interactions concerned, such as electrostatic forces, hydrogen bonding, Van der Waals interactions and bonded interactions.

Among all quantum chemistry techniques, density functional theory (DFT) provides the best compromise between accuracy and cost. The applications of DFT in the study of protein-ligand interactions are wide [14-17]. Common analysis at this level of theory is based on the study of molecular orbitals, binding energies, molecular electrostatic potentials, density of states projections and structural conformations of biomolecular complexes. But the quadratic relationship between the electronic coupling strength and the rate of electron transfer motivate us to emphasize the study of the electron transfer as a correlator to the binding affinity.

The breakpoint of the study of electron transfer rates starts in the earlier 80's with the Nobel Prize winner Rudolph A. Marcus and his document "Electron transfer in

chemistry and biology". Since then, further study of electron transfer in biological context has been done [18] along with computational approaches [19].

The nature of our variable is based on the electron transfer in biological processes. Even though the functions of electron transfer in protein-ligand complexes are not well understood yet, almost all the biological processes go along a series of transfer of electrons until a determined function is accomplished. In some way, nature must discriminate between ligands according to the transfer rate of electrons in order to make it specific and efficient.

The physical basis of the interaction energy relies on fundamental variational principles that govern the chemical reactivity. The interaction energy is decomposed in two complementary components: the electron transfer and the electrostatic contribution. The first is associated with the covalent bonding and is described by the variational principle with respect to the electron density: for the two compounds (receptor and ligand) interact, a rearrangement of the electron density, mainly in the interface, must occur until the energy of the complex is minimized. The electrostatic contribution is associated with uncovalent bonding and is explained by the variational principle with respect to the external potential: for two reactants come together, a rearrangement of the nuclei and charges must occur until the energy of the complex is minimized.

We propose the following protocol: solving the nuclear motion problem without considering the electronic part using Docking and Molecular Dynamics methods. Then we solve the electronic wavefunction using ab-initio approaches and calculate the protein-drug electronic coupling.

1.4 APPROACH TO SOLVE THE PROBLEM

Relying on variational principles implies a calculation based on first principles theories, but the simulation would result impractical for large protein-ligand systems in terms of computational costs.

In order to facilitate the search of an energy minimum configuration, we use the Born-Oppenheimer approximation: as electrons move much faster than the nuclei, the electrons are considered moving in a potential created by fixed nuclei. Thus, we separate the configuration problem in two, one for the nuclei and the other for the electrons. We use force field based methods to approach the solution of the nuclei without considering the electrons explicitly. Then, we fix the new configuration of the atoms, and try to solve the electronic part by using a self-consistent calculation.

The protocol proposed to solve the nuclei motion starts with a docking calculation. Knowing the 3D configuration of the target molecule and of the different drugs, we try to find, by scoring functions, a minimum orientational mode of the ligand in the active site of the receptor. We add explicitly solvent molecules (water) to the best ranked configuration from the DOCK results. After that, we start the molecular dynamics simulation. We make a minimization by steps: first we minimize the solvent molecules with the ligand and the receptor fixed; then we unfix all the hydrogen from the receptor and the ligand, keeping also free the solvent molecules; and finally we let free all atoms of the system. After the minimization, we heat the system quasi-adiabatically until 298K and then, we equilibrate the system at constant pressure of 1 atm. and 298K. By this procedure we try to not denaturalize the conformation previously obtained from the DOCK results and reach a conformation under living conditions.

For the electronic part, we approximate the analysis to be local. Therefore, for an ab-initio studio we restrict the calculation to a certain region of the active site, and we make a self-consistent field calculation using the density functional theory (DFT) approach. Instead of solving the wavefunction, as in standard ab-initio methods, DFT

yields the electron density energy of a polyatomic system from its atomic structure, based in the existence of a one to one relation between the electron density and its external potential. The Hamiltonian and overlap matrices are solved and used to calculate the electron transfer rate between the reactants.

Finally, we create large libraries with the electron transfer behaviors between several combinations of common functional fragments of drugs and all the known amino acids, and use them as new tools for molecular recognition. The libraries with electron transfer behaviors can be used at every step of the drug design. When the structure of the target molecule is not known (ligand-based drug design), we make use of the known activity of the drugs and correlate it with any structural or electronic property of them, this procedure is known as Quantitative Structure Activity Relationship (QSAR). The created libraries will serve as a resource for new relationships. In Structure Based Drug Design (SBDD), when the protein is known in a 3D configuration, we could include the libraries to create new scoring functions for the search of conformational orientations of the drug inside of the active site. And obviously in De Novo design (the optimization and creation of new drugs), the electron transfer characteristics would be a new variable to be included in the evolutionary algorithms general used in the growing fragment by fragment of the drug.

2. METHODS

2.1 BINDING AFFINITY AND THE HSAB PRINCIPLE

Binding affinity is a constant that estimates the tendency of two systems (generally one in much larger concentrations than the other) to dissociate, usually defined as the inverse of the dissociation constant:

$$K_d = \frac{[P][L]}{[C]} \quad \text{... eq. 1}$$

where [P], [L], [C] mean concentrations of the protein, ligand and the complex respectively. Clearly, as the concentration of the complex (protein and ligand bonded) increases, the dissociation constant goes down and the binding affinity high.

We should notice that a major property in drug research is the activity of the drug, defined as the concentration needed to activate or inactivate a specific target molecule. How do we estimate the binding affinity from computational chemistry?

Many statements to predict the tendency of the reactions has been proposed over the last 30 years. After definition of electronegativity, by Parr [20], as the desire of a molecule to accept electrons, or strictly speaking, the energy needed/obtained to remove/add an amount of electron density at space *r***;** the principle of electronegative equalization came up, in which the energy is minimized only if the electronegativity (the negative of the chemical potential) is equalized. Thus, the same tendency in many similar ways such as in acid-base reactions, donor-acceptors, oxidation-reduction were proposed [21,22]. The electronegativity, χ, is then the variation of the energy with respect to the electron density, as in eq. 2, and it is also the negative of the chemical potential, μ.

$$\chi(r) \equiv -\left(\frac{\delta E(\rho)}{\delta \rho(r)}\right) = -\mu \quad \text{... eq. 2}$$

One approximation of the electronic chemical potential, μ, is to relate it to the electron affinity, *A*, and ionization potential, *I*

$$\mu \approx -\frac{(I+A)}{2} \quad \text{... eq. 3}$$

As the first derivative of the energy with respect to the electron density is important, the second derivative should be as well, as proposed by Parr [20]. Defined as chemical hardness, the second derivative of the energy with respect to the number of electrons can be approximated as well in terms of electron affinity, *A*, and ionization potential, *I*:

$$\eta \equiv \left(\frac{\delta^2 E}{\delta N^2}\right) \approx I - A \approx \quad \text{... eq. 4}$$

The Hard Soft Acid Base principle states that hard acids tend to associate with hard bases and soft bases tend to associate with soft acids. Its basis relies in the two components of the interaction energy and their relationships to the chemical potential and hardness.

The first component is due to the electron transfer when the two reactants come together; they go through changes in the electron density between the reactants lowering the energy until the chemical potential is equalized,

$$\Delta E_{et} = -\frac{(\mu_B - \mu_A)^2}{2(\eta_A + \eta_B)} \quad \text{... eq. 5}$$

here μ_B, μ_A, η_A and η_B denote the chemical potential and the hardness of the acid and the base involved in the reaction; and μ_o is the chemical potential after the reaction. At similar strengths of chemical potential, the electron transfer is proportional to the softness (inverse of the hardness), explaining the tendency of soft molecules to react each other.

The second component corresponds to the non-bonded interactions. Once the chemical potential is equalized the system goes through changes in the external potential at a constant chemical potential as expressed by:

$$\Delta E_{qq} = \frac{q_A q_B}{(r_A + r_B)} \quad \text{... eq. 6}$$

expressing the energy that stabilizes the system of two opposite charges when they approach to each other. For hard reagents, which generally are highly charged, electron transfer becomes unimportant the electrostatic interaction dominates in the system.

Another change in the reaction is the polarization of the pseudo acids and bases, which lowers the energy by:

$$\Delta E_{pol} = -\frac{\alpha_B (q_A)^2}{2(r_A + r_B)^4} - \frac{\alpha_A (q_B)^2}{2(r_A + r_B)^4} \quad \text{... eq. 7}$$

α_A and α_B mean the polarizability of the reactants. But induced charges are only present when highly charged molecules react, that is, when electrostatic interactions are present as well. Quantitativily the former one is dominant, so generally the polarizability term in neglected.

It is also possible quantify, in the same terminology, the non-bonded interactions of neutral or small charged components. Those kinds of interactions are called the London dispersion interaction:

$$\Delta E_{London} = -\frac{3}{2} \frac{\varepsilon_A \varepsilon_B}{\varepsilon_A + \varepsilon_B} \frac{\alpha_A \alpha_B}{(r_A + r_B)^6} \quad \text{... eq. 8}$$

ε_i represents the average excitation energy of the molecule. From the previous approximation ($\eta \approx (\text{LUMO - HOMO})$), the chemical hardness is approximately equal to the first excitation energy:

$$\Delta E_{London} = -\frac{3}{2} \frac{\eta_A \eta_B}{\eta_A + \eta_B} \frac{\alpha_A \alpha_B}{(r_A + r_B)^6} \quad \text{... eq. 9}$$

It can be proved (not shown here) that the hardness is inversely proportional to the radius of the active site of interaction and to the polarizability coefficient, but in different degrees. From that follows that:

$$\Delta E_{London} = -\frac{3}{2} \frac{c_\alpha^{\,2}}{c_r^{\,6}} \frac{(\eta_A \eta_B)^4}{(\eta_A + \eta_B)^7} \quad \text{... eq. 10}$$

2.2 ELECTRON TRANSFER THEORY

Classical theory of long distance electron transfer reactions between reactants, in which the interaction of the relevant electronic orbitals is weak, is described by the Fermi's golden rule (equation 11). According to the Fermi's rule the first order rate constant is proportional to the square of the coupling, H_{AB}, of the wavefunctions of the reactants:

$$k = \frac{2\pi}{\hbar} H_{AB}^2 * FC \qquad \text{... eq. 11}$$

The second term FC is the "Franck-Condon" factor, which is a sum of products of overlap integrals of the vibrational and solvational wavefunctions of the reactants with those of the products suitably weighted by density of nuclear states. $\hbar$ is the reduced Planck constant equal to $6.58211899 \times 10^{-16}$ eV·s.

The separation of the two components (electrons and nuclei) is allowed by the Franck-Condon principle: since the electrons transferred are light particles, during the actual electron transfer the nuclei do not have time to change their positions, and then the transfer occurs only at near nuclear configurations for which the total potential energy of the reactants is equal to the products.

A simplified example of a one-dimensional potential energy profile is given in figure 3. The electron transfer only occurs at the intersection of those two curves. In order for the reactants to reach it, thermal, vibrational and orientational coordinates are needed. The probability of going from the reactants (pink curve) to the products (blue curve) surface depends on a number of factors such as the extent of coupling of the electronic orbitals of the two reactants and their electronic motion.

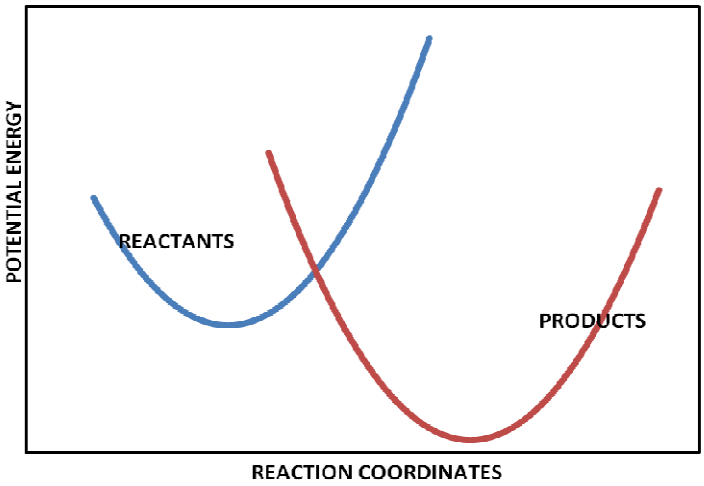


Figure 3: Potential energy surface of reactants and products. Starting from the minimum of the reactants surface, thermal fluctuations take the system to the intersection where electron transfer takes place and the complex is formed.

2.3 DOCKING

Molecular docking is the prediction and ranking of the conformational complex structure as a result of the chemical reaction of a given ligand and a receptor of known structure.

There are two main issues in protein-ligand docking. The first is the functions that are going to score or rank every conformation. The second is the search algorithm between all positions and orientations, thus yielding huge number degrees of freedom. If the algorithm searches the best conformation going through all the degrees of freedom (with some conformation exceptions like the structure of a benzene ring for example), we call it a systematic algorithm. Sometimes it also uses a likelihood function, generated either by Monte Carlo methods or genetic algorithms in order to converge faster to a minimal conformation.

Different approaches in the score functions have been developed such as force fields that parameterize the movement of the atomic centers by potential functions [23,24]; empirical scoring functions that calibrate the receptor-ligand interactions based on complexes of known affinity [25-27]; or statistical potentials based on the distribution of distances between atoms [28-30].

Furthermore two characteristics of proteins are also important. One is the flexibility: x-ray structure shows that the ligands are buried or covered by the protein between 70 to 100 percent of their surface area. And the other is the solvation: it has been proved that the entropic effects of the water medium are crucial in the conformation and affinity of the complexes. The flexibility is solved by increasing the degrees of freedom of the atoms in the active site; on the other hand, the solvation is compensated implicitly in the score functions.

2.4 MOLECULAR DYNAMICS

Molecular dynamics can serve as a search algorithm to explore the conformational space, and then use random snapshots of the complexes of the MD simulation as initial conformations for the docking scoring. Also, the rigid and flexible part of the receptor can be explored by MD and then, restrict the docking search algorithm to the most flexible section in order to reduce the degrees of freedom.

After the docking process, MD can be used to relax the entire system; include the solvent molecules in an explicit manner; and compute, although it is highly expensive, the binding free energy. The computation of the free energy can be done either by free energy perturbation methods which are highly accurate but its constrained to similar ligands for a consensus comparison; or by linear integration energy methods where more empirical data are used, thus loosing accuracy.

Since at this stage the objective is still a discrimination of the ligands, a qualitative analysis is still relevant. One type of analysis is looking at the atomic position fluctuations of the ligand inside of the active site, or the bond distance changes along the simulation. The basic correlation here is: the less flexible the more stable. The number and distribution of the H-bonding and ionic bonding between the reactants is also an indicator of the stability. The accessible surface area of the ligand in the active site is correlated with the flexibility as well.

2.5 AB-INITIO: ELECTRON DENSITY CHANGE

One method to solve the Schrödinger equation:

$$\hat{H}\Psi = (\hat{T} + \hat{V}_{ext} + \hat{V}_{ee})\Psi = E\Psi \quad \text{... eq. 12}$$

where $\hat{H}$ is the Hamiltonian operator for a system of nuclei and electrons. The Hamiltonian operator is composed of three contributions; $\hat{T}$ is the operator for the electron kinetic energy contribution, $\hat{v}_{ext}$ is the operator for the potential energy of the electrons in the external field generated by the nuclei and, $\hat{v}_{ee}$ is the operator for the energy of all the electron-electron interactions. In Density Functional Theory (DFT), the wavefunction, Ψ, is dependent of the three spatial coordinates and the spin of every electron, thus the wavefunction is a function of 4N variables (N is the number of electrons). The Schrödinger equation is solved using DFT, which is based on the Hohenberg-Kohn-Sham theorems [31,32]; using the electron density ρ as a variable, which dependent on only three variables (three coordinates in 3D space).

The total energy is

$$E(\rho) = T(\rho) + \int V_{ext}(\vec{r})\rho(\vec{r})dr + \frac{e^2}{2}\int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r}-\vec{r}'|} d\vec{r}\, d\vec{r}' + V_{xc} \quad \text{... eq. 13}$$

where the electron-electron interactions have been decomposed in two components (the last two terms of equation 13). The first of those (the integral term) accounts for the by the columbic repulsion energy of the electrons. This integrand (also called Hartree term) depends only on the local electron density through the entire space and it represents the potential energy induced by the repulsion of the electrons at distance $|\vec{r}-\vec{r}'|$ over the entire space. The second term is the so called exchange and correlation energy that accounts for the rest of the electron-electron interactions; the exchange is a non-local contribution that arises from the antisymmetric nature of the wavefunction. The

correlation term, instead, accounts for the effect of Coulomb correlation upon many-electron wavefunction. It is important to notice that the exchange and correlation term includes also the contribution of the kinetic energy arising from the electron-electron interactions; in other words, the kinetic energy difference from the total system and a non-interacting system.

The electron density for a closed shell system is written as a sum of squared one-electron wave-functions,

$$\rho(r) = 2\sum_{i}^{N} |\psi_i(\vec{r})|^2 \quad \text{... eq. 14}$$

The one-electron wave-function solutions are expanded as a linear combination of atomic-like orbitals ϕ_k :

$$\psi_i(\vec{r}) = \sum_k c_{ik}\phi_k(\vec{r}) \quad \text{... eq. 15}$$

where the c_{ik} are eigenvectors coefficients corresponding to the linear expansion of the one-electron wave function, ψ_i, in atomic orbitals.

The first Hohenberg-Kohn theorem states that external potential V_{ext} is uniquely determined by the electron density. Then, since the Hartree term and the exchange and correlation terms depend on the electron density, ρ, and in turn the electron density corresponds to a unique N-electron Hamiltonian; the solution is obtained iteratively. The solution of the Kohn-Sham equation starts with a guess of the electron density, which is built using semiempirical approaches, and then the Kohn-Shan matrix and a first eigenfunction solution are calculated. The new eigenfunction yields a new electron density and so on. The process is repeated until the total energy is minimized within certain threshold.

The electron density allows us to estimate any property of the nuclei-electrons system. Qualitatively, we can estimate the strength of the bond using the electron density. In any electron transfer reaction, even in outer-sphere or weak coupled, a change of the electron density from the most negative electrostatic molecule to the most

positive occurs forming a covalent bond, even though the bond is still weak. Then the total amount of charge density transferred is proportional to the strength of the bond.

2.6 AB-INITIO: MOLECULAR ELECTROSTATIC POTENTIAL (MEP)

The electrostatic potential $V(\vec{r})$ due to the nuclei and electrons of a molecule at a point *r* in the space is defined by

$$V(\vec{r}) = \sum_{A} \frac{Z_A}{|R_A - \vec{r}|} - \int \frac{\rho(\vec{r}')d\vec{r}'}{|\vec{r}' - \vec{r}|} \quad \text{... eq. 16}$$

where Z_A is the charge on nucleus *A*, located at R_A, and $\rho(\vec{r}')$ is the electron density at any point $\vec{r}'$. The first component is a positive contribution from the nuclei, Z_A and the second is a negative contribution from all the electrons (integral covers the whole space), thus the net potential is positive, zero or negative, depending on which contribution is the more dominant[33]. Negative regions are more likely to be attacked by an electrophile (positively charged) species. The preference depends, on the value of the negative potential. Some negative regions are due to the π-orbitals of unsaturated molecules such as those from benzene rings, heteroatoms fragment of an amino acid with carbonyl lone pairs.

2.7 AB-INITIO: DENSITY OF STATES PROJECTIONS

The density of states (DOS) spectra represents the occupancy distribution of electrons along the energy states. According to the tight binding theory, a narrow band of the DOS is an indicator of the small coupling between the valance states (the states of the reagents responsible for the reaction), meaning a more reactive system[34]. The explanation is that electrons can exist only at quantized energy levels, and the accumulation of electrons at a specific energy is only possible at a non-interactive system. For example, the DOS of the surface atoms of a cluster shows a narrow and sharp peak instead of the broad band DOS of the bulk atoms of the same cluster. A high coordination number of the atoms in the bulk cluster broads the distribution of the energy levels. It follows that a change from a narrow to a broad peak after the reaction indicates a strong chemisorption between the reactants (protein and drug).

Also, a downshift in energy of the valance states after the interaction means a strong chemisorption, since the electrons are trying to go to more stable states.

2.8 AB-INITIO: TRANSMISSION FUNCTION

Starting from the Marcus theory, a model of the current-voltage behavior adapted for a donor-bridge-acceptor system is described by the Landauer formula [35] (the full mathematical derivation can be found elsewhere [36]). The current of electrons with energy between E and $E + dE$ is determined by,

$$I(V) = \frac{2e}{h} \int_{-\infty}^{\infty} T(E,V)[f_1(E,V_1) - f_2(E,V_2)]dE \qquad \text{... eq. 17}$$

where $T(E,V)$ is the transmission coefficient (probability per transport channel), $f_i(E)$ is the Fermi-Dirac distribution and V_i is the bias voltage.

In molecular junctions, the transmission function T is calculated using the Green's function approach:

$$T(E) = Trace(\Gamma\ G_M \Gamma\ G_M^+) \quad \text{... eq. 18}$$

Γ represents the coupling protein-drug:

$$\Gamma\ = i[\tau^+ g\tau - \tau g\tau^+] \quad \text{... eq. 19}$$

where $\tau^+ g\tau$ is a Hermitian conjugate self-energy matrix to describe the coupling between the reactants. G_M describes the retarded Green function of the drug molecule affected by the protein.

From quantum chemical calculations we can approximate these functions into matrix forms. We use a partitioned Hamiltonian in atomic basis to describe the retarded Green function:

$$\mathrm{G}(E) = \begin{bmatrix} g^{-1} & -\tau & 0 \\ -\tau^+ & E - H_c & -\tau^+ \\ 0 & -\tau & g^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} G_1 & G_{1M} & G_{12} \\ G_{M1} & G_M & G_{2M} \\ G_{21} & G_{M2} & G_2 \end{bmatrix} \quad \text{... eq. 20}$$

g represents the protein, as it was a substrate, and again τ describe the protein-drug coupling. H_c is the Hamiltonian of the isolated drug and E is the electron energy.

Solving the inverse we find that G_M is equal to:

$$G_M\ = [E - H_c - 2\tau^+ g\tau]^{-1} \quad \text{... eq. 21}$$

We assign the components of the inverse of G_M to the submatrices of the partitioned KS Hamiltonian, which in turn depends of a linear combination of atomic like orbitals:

$$H_{KS} = \begin{bmatrix} H_P & H_{PD} & H_P \\ H_{DP} & H_{DD} & H_{DP} \\ H_P & H_{PD} & H_P \end{bmatrix} \quad \text{... eq. 22}$$

where the subscripts P and D refer to the protein and the drug, respectively.

Then H_{DD} is assigned to H_c, H_{PD} to τ and H_{DP} to τ^*. Furthermore, it has to be noticed that the real KS Hamiltonian matrix solved is of the form:

$$H_{KS}\Psi = eS\Psi \quad \text{... eq. 23}$$

instead of:

$$H\Psi = E\Psi \quad \text{... eq. 24}$$

where S is the overlap matrix and e are the eigenvalues, then the corresponding transformation has to be done in order to these matrices to be Hermitians.

Full detailed of the methodology to calculate the transmission function, but in a sense of molecular conduction in nano-junctions, can be found in ref. [37-39] and it has been implemented in the GENIP program [40-43].

3. TEST CASES

3.1 OXIDATION AND REDUCTION AFFINITIES ON CATALYTIC SURFACES

As preliminary results, we test the use of the electron transfer as a correlator for binding affinity. We build junctions based on known electro-catalytic interactions such as hydrogen oxidation and oxygen reduction in platinum surfaces. The results of this work were also published in our previous paper[44]. The systems, as shown in figure 4, are made of a bulk of glycine amino acids (green surface) that will resemble the biological environment. The bulks end up in platinum or palladium terminals. Between the terminals we have the test ligands, oxygen and hydrogen dimmer. Even though these systems are not so similar to real cases, both share a common characteristic: the electron transfer in the active site. To calculate the electron transfer characteristics, showing it in an I-V behavior, the procedure differs slightly from the proposed in section 1.4.

The GENIP program estimates the current-voltage behavior using the bulk of glycine as source of electrons. The terminals are either the platinum or palladium atoms. The flux of electrons goes from terminal to terminal through the bridge molecules that are either the oxygen or hydrogen dimmer. As stated before, a higher current correlates with a stronger reactivity between the compounds.

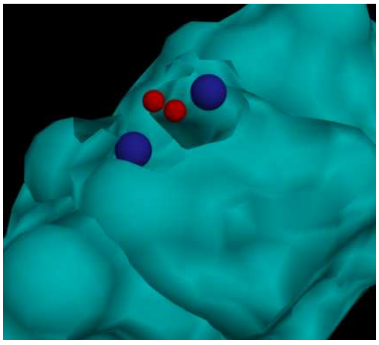
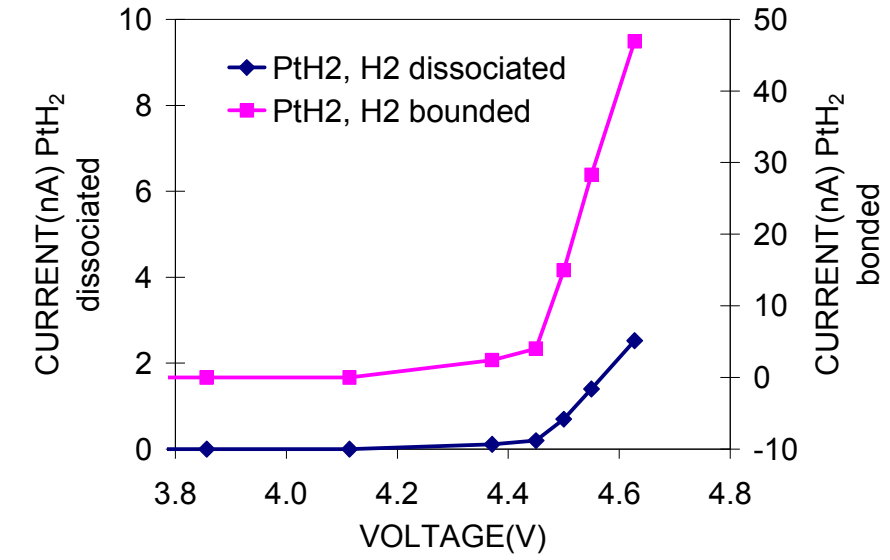
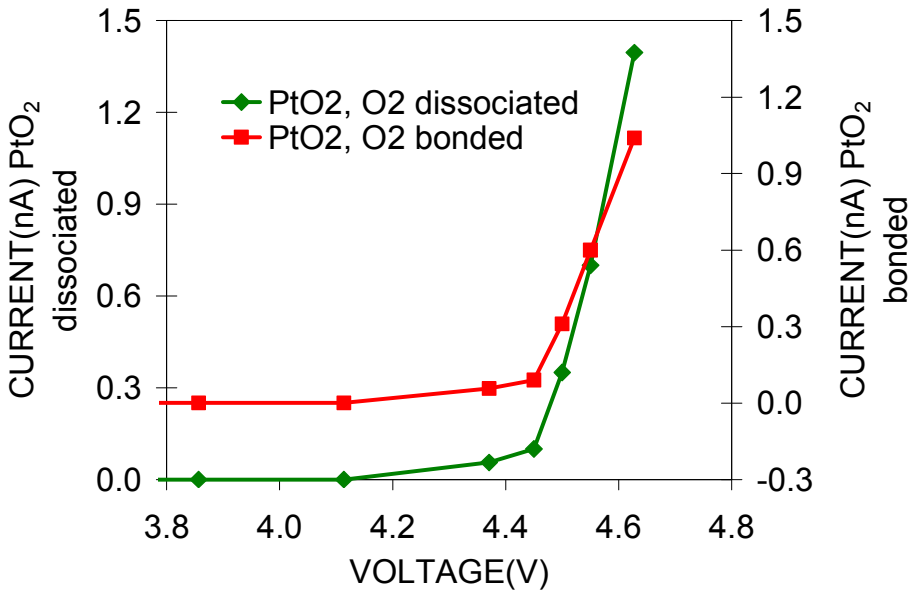


Figure 4: Picture of the test junction Pseudo protein (green) - Platinum (blue) – O_2 (red) - Platinum-Pseudo protein. Similar tests were constructed by replacing the O_2 with H_2 and platinum with palladium

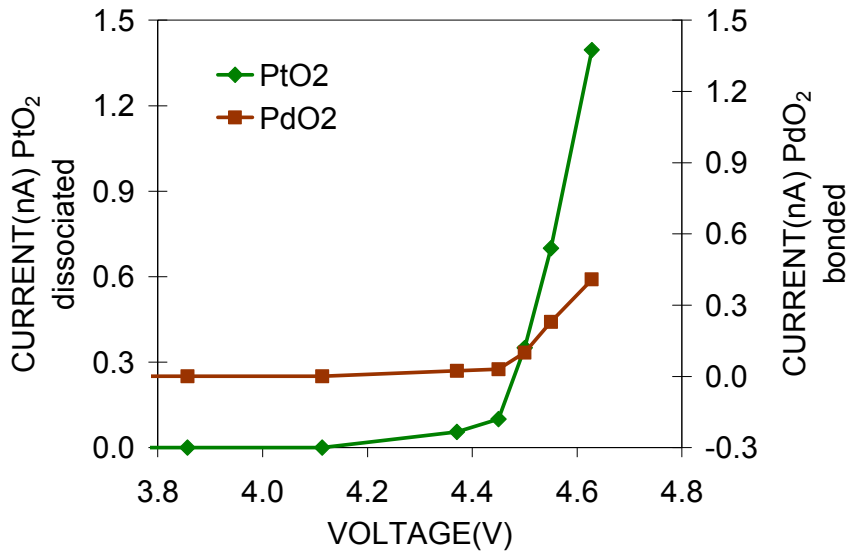


A



B

Figure 5: I-V curves for A: G-Pt-HH-Pt-G with H_2 dissociated (blue) and H_2 bonded (pink) showing higher current for H_2 bonded, B: G-Pt-OO-Pt-G with O_2 dissociated (green) and O_2 bounded (red) showing higher current for O_2 dissociated, C: G-Pd-OO-Pd-G (green) and G-Pt-OO-Pt-G (maroon) showing higher current for palladium.



C

Figure 5 Continued

The results in figure 5 are correlated with experimental tendencies. The current-voltage curve for the platinum-hydrogen interactions (figure 5A) shows a higher a current when the hydrogen dimmer is bonded rather than when it is dissociated. On the other hand, the platinum-oxygen curve (figure 5B) shows the opposite trend: a higher current when the oxygen is dissociated than when is bonded. The behavior corresponds with experimental formation of hydroxides on the platinum surfaces, due to the high affinity for oxygen dissociated.

Differential experimental heats of oxygen absorption on platinum surfaces (figure 6) show a down peak at oxygen saturation: thermal desorption spectra has shown that the energy of adsorption of oxygen dissociated on platinum surface is about 340 kJ/mol, and about 40 kJ/mol for oxygen molecular. According to figure 6, the oxygen is presented in a dissociative manner on the platinum surface at the beginning of the adsorption. Only when the metal surface is saturated of oxygen, the molecular oxygen is adsorbed.

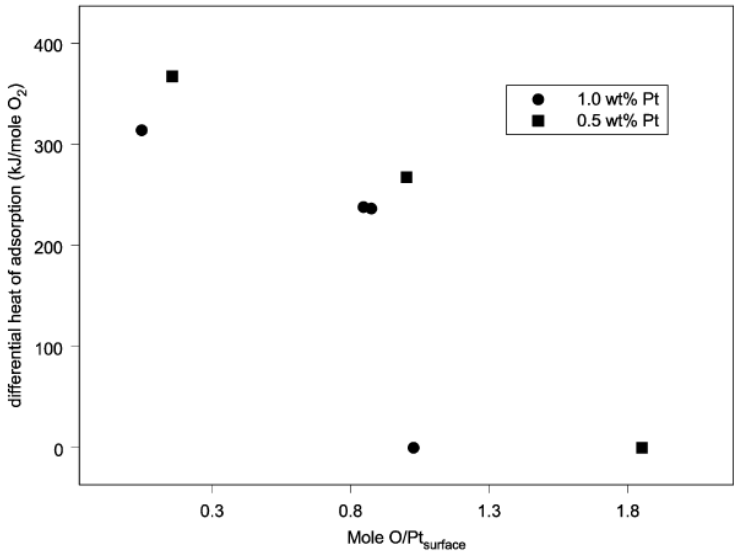


Figure 6: Differential heats of oxygen adsorption on 0.5 and 1 wt% Pt/TiO_2 measured at 30 degrees Celsius[45]. This shows that the adsorption of oxygen in the platinum surface is in a dissociative manner until the surface is saturated

Another comparison between different catalytic surfaces is done by replacing the platinum atoms with palladium (figure 5C). The I-V results show what is known: a higher affinity or reactivity for platinum surfaces than for palladium.

3.2 PIM-1 KINASE: IMPORTANCE, KNOWN DRUGS, ACTIVE SITE

We test our protocol with the PIM-1 kinase. PIM kinases are emerging as important mediators of cytokine signaling pathways in hematopoietic cells, and they contribute to the progression of certain leukemias and solid tumors. We test three of well-known drugs active to the PIM-1 kinase reported in BindingDB webpage (http://www.bindingdb.org). These drugs (figure 7) are a family of bisindolylmaleimides (BIM)[46] based on the nonspecific kinase inhibitor "Staurosporine", identified as very potent inhibitors of PIM-1 kinase. Furthermore, the effect of the charges in the specificity of the protein-ligand complexes motivates us to create another test set of drugs by protonating each one of the original set. For example, it has been shown that going from the zwitterion to the protonated state of a histidine amino acid located in the active site can totally change the activity of a certain protein. So for every drug on the test we have a similar one protonated in its most positively spot.

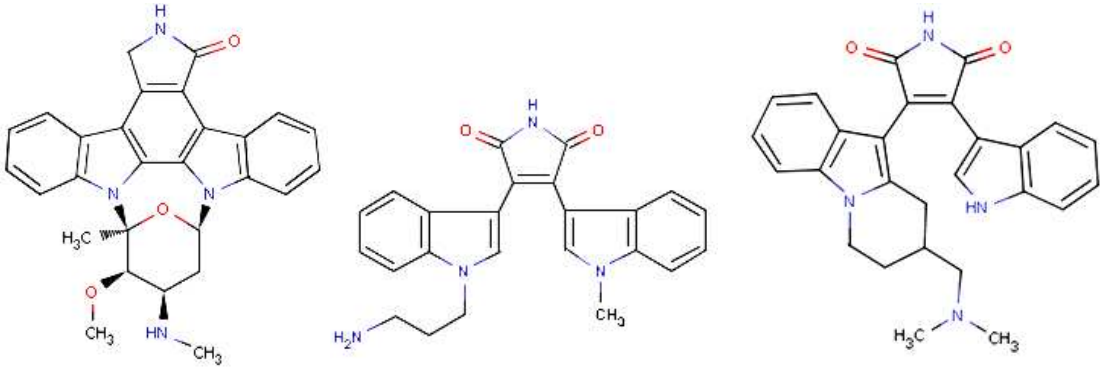


Figure 7: Structures of known drugs for the PIM-I kinase. From left to right: Staurosporine ($C_{28}H_{26}N_4O_3$), BIM-8 ($C_{24}H_{22}N_4O_2$), BIM-11 ($C_{27}H_{26}N_4O_2$)

We also do a characterization of the active site that helps us to recognize the different interactions protein-drug responsible for the binding affinity (see figure 8). The active site shows a negative charge in the bottom due the presence of Aspartic acid (Asp) and Glutamic acid (Glu) making ionic interactions with the drug. This force is balanced

with positive amino acids on the top, such as Lysine (Lys) and Arganine (Arg); and also with hydrophobic interactions between the rings of the drugs and the hydrophobic amino acids. We approximate the reaction between the protein and the drugs to be local, so for a further ab-initio analysis, we restrict the system to consider the atoms within a radius of 5.0 A from the drugs.

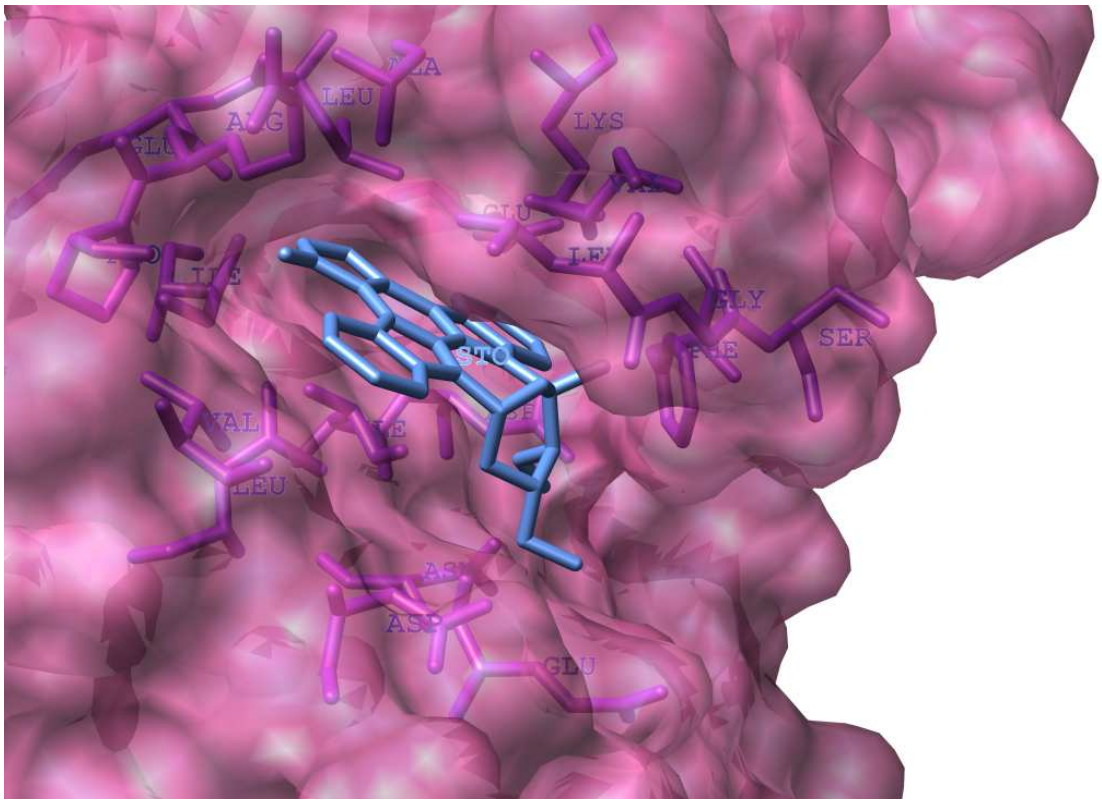


Figure 8: Imaging [47]of the characterization of the active site of PIM-I kinase showing a dominant number of negative amino acids on one side of the active site and positive and hydrophobic on the other.

4. RESULTS

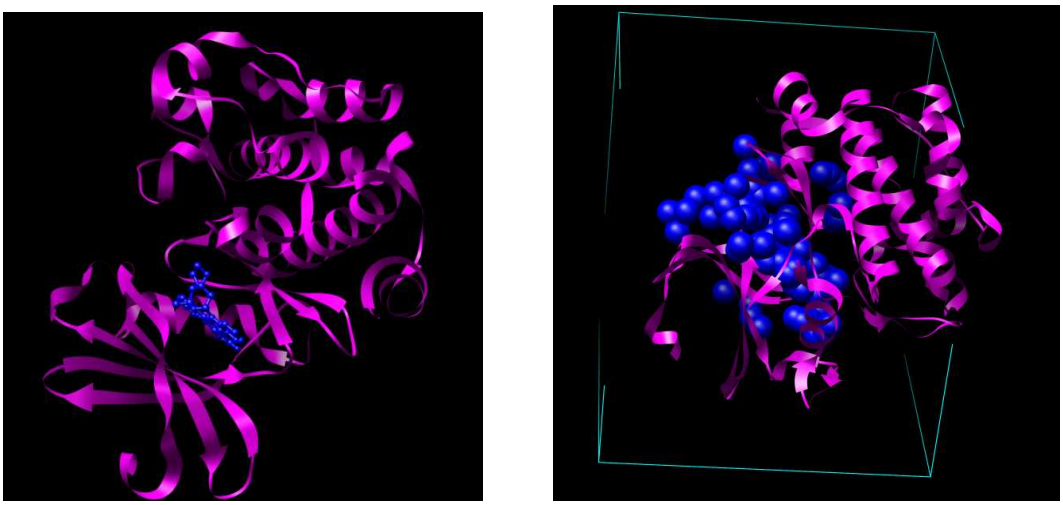
4.1 DOCKING CALCULATION

Receptor preparation: PIM-1 kinase crystal structures were obtained from the Protein Data Bank (http://www.rcsb.org). Further modification of the structure was not needed. Hydrogen atoms were then added to the entire receptor. Gasteiger charges were computed by ANTECHAMBER [48] and added using the USCF Chimera program[47].

Ligand preparation: Ligands were prepared from crystal structures reported in BindingDB webpage (http://www.bindingdb.org). Bond and atom types were checked, and hydrogens were added. Gasteiger charges were calculated for all ligands.

Molecular docking: Docking of the three complexes (figure 9A) was performed using DOCK 6.0.0 [49] and its accessory programs. DMS (dot molecular surface) tool [50] was used to create a molecular surface of the receptor. The Sphgen program[49] was then used to create a set of spheres orthogonal to the molecular surface of the protein. The resulting file was edited to include only spheres within a cutoff distance of 10.0 Å from the ligand (figure 9B). This distance was sufficient to include the entire active site. Scoring and bump grids were then generated using GRID [51]. Contact and energy scores were calculated using the AMBER99 parameter force field with an energy cutoff distance of 9A and a van der Waals repulsive exponent of 12. All other parameters were left as their defaults.

The conformational search goes through the spheres center method. The center of the spheres represents the possible positions of the center of the atoms of the ligand. Depending of the search algorithm the ligand goes through several combinations of positions until finds a conformation with the lowest energy.



A B

Figure 9: Imaging [47] of A) PIM-I Kinase (magenta), an oncogene-encoded serine/threonine kinase primarily expressed in hematopoietic and germ cell lines showing the active site with its ligand (blue). B) Schematic picture of the modeling of the active site by the sphere centers method.

Table 1: Enthalpy of formation of known drugs obtained from the DOCK program, and experimental values. Also a test of the results has been done for the Protonated states of the drugs showing indeed higher contributions of the electrostatic interaction, and similar values for the VDW contribution.

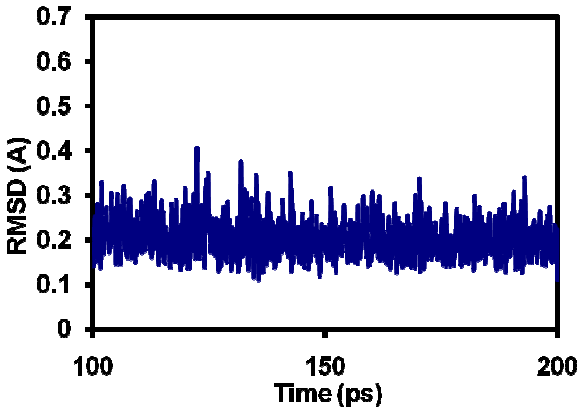
Contribution (kJ/mol)	Staurosporine		BIM-8		BIM-11	
	Neutral	Protonated	Neutral	Protonated	Neutral	Protonated
VDW contribution	-52.6	-40.4	-29.5	-28.9	-34.5	-37.6
Elect. contribution	-3.5	-9.8	-5.1	-21.4	-0.5	-11.1
TOTAL	-54.1	-50.2	-34.6	-50.2	-35.0	-48.7
Exp. enthalpy of formation	-33.8		-41.3		-13.7	
Exp. free energy	-48.0		-43.0		-35.5	

The results of Table 1 show that at least the trend of the experimental free energy is barely followed by the docking calculations.

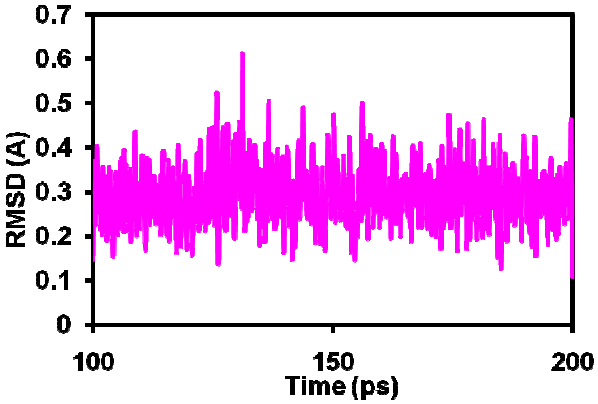
4.2 MOLECULAR DYNAMICS SIMULATION

After the docking scoring calculation, we take the best scored conformation of the receptor-ligand complex (the one with the lowest energy). We use the Perl script to generate the force field parameter file of the unknown atoms of the drugs based on the GAFF force field. At the same time, the Perl script adds water atoms into the system. Then we start minimizing the water molecules and keeping all the others atoms of the receptor-ligand fixed. At this step, the minimization is of 1000 steps and we consider the Van der Waals and electrostatic interactions of the atoms separated by three bonds with a cutoff of 8 A. A second consequent minimization is done keeping unfixed all the hydrogens atoms of the system along with the water molecules for 10000 steps as well, and considering the VDW and columbic interactions the atoms separated by two bonds. Finally, we minimize the system for 25000 steps let them free all the atoms. Then, we heat the system very slowly with an increment of 0.001 K for every time step (1 fs.) until 298K. After that we equilibrate the system for 50 ps with 1 fs time step, at 1 atm. and 298K with a rescale of the temperature every 10 fs.

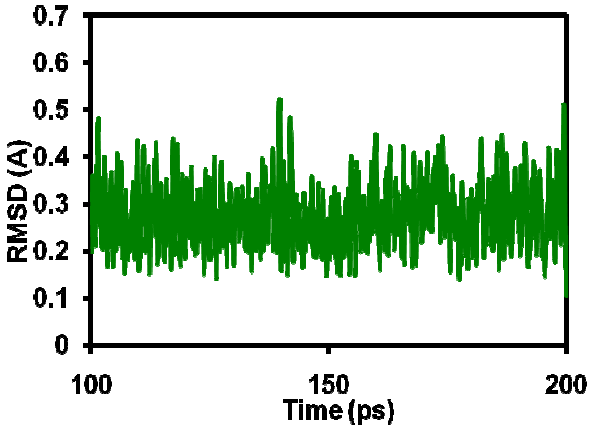
We do a qualitative analysis of the fluctuations of the drugs in the active site (figure 10 and 11), basing in the statement: “the less flexible, the stronger the binding”. We calculate the RMSD fluctuation of the heat atoms (C, O and N) of the drugs along a short range of time after the equilibration stage. The fluctuations shown in figure 10 determine that staurosporine is the most stable drug. In specific, the order of stability is staurosporine > BIM-11 > BIM-8. Instead, the protonated set of drugs shows a slight different trend (figure 11): staurosporine > BIM-8 > BIM-11



A. Staurosporine (Mean: 0.21)

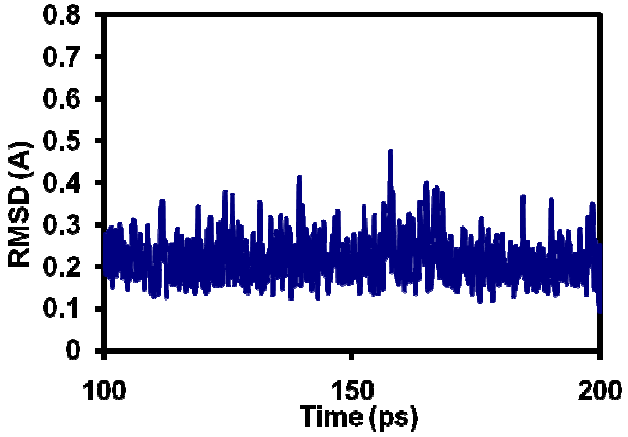


B. BIM-8 (Mean: 0.30)

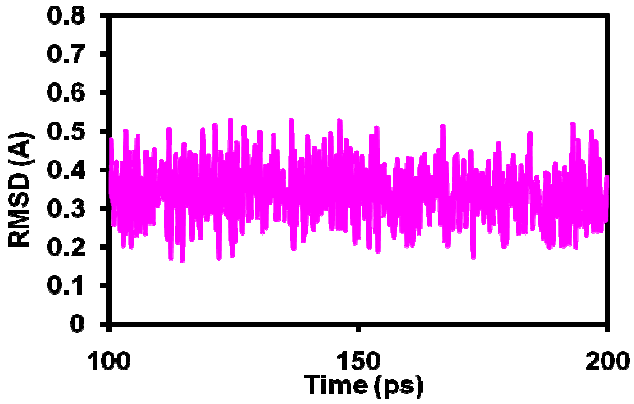


C. BIM-11(Mean: 0.28)

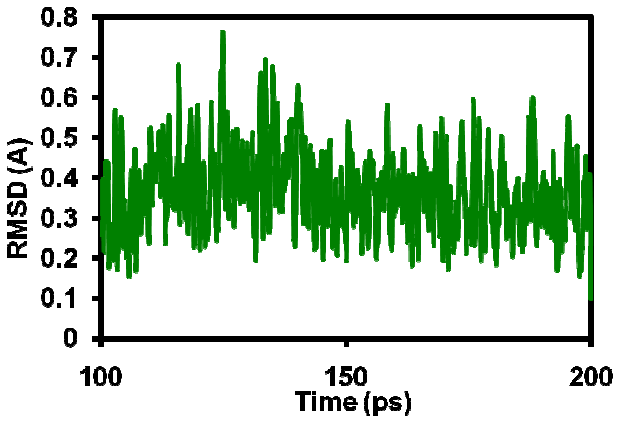
Figure 10: RMSD fluctuations of the heat atoms of the neutral set of drugs in the active site along the equilibration step. Their fluctuations are correlated with the free energy of the reaction.



A. Staurosporine (Mean: 0.22)



B. BIM-8 (Mean: 0.34)



C. BIM-11(Mean: 0.37)

Figure 11: RMSD fluctuations of the heat atoms of the protonated set of drugs in the active site showing the sensitivity of the system to a change of charges.

4.3 AB-INITIO: ELECTRONIC BINDING ENERGY AND THE HSAB PRINCIPLE RESULTS

After the molecular dynamics simulation, we take the "average conformation" of the system and restrict it to consider only the atoms in the active site (a radius of 5 A with respect to the drug). "The average conformation" is made up of the average cartesian coordinates of the atoms during a short range of time after the equilibration stage. The conformation at the equilibration stage accounts for the changes in the position of the nuclei due to the thermal energy. Reducing the system to consider only the active site allows to use more expensive method/basis set in the ab-initio approach. In order to account for the non-bonded interactions that are usually overestimated by conventional DFT methods like B3PW91, we use a novel DFT-like functional M05-2X that uses a different proportion of the Hartree-Fock exchange potential. All the ab-initio calculations in this project has been done using Gaussian 03[52].

Table 2: Electronic binding energy and driving force energy.

Energy (eV)	Staurosporine		BIM-8		BIM-11	
	Neutral	Protonated	Neutral	Protonated	Neutral	Protonated
Reorganization		9.94		6.35		9.88
energy	3.35		0.91		1.72	
Driving force	0.57	4.18	0.04	1.18	1.37	4.39

We define as the electronic binding energy to the energy difference associated to the redistribution of the electrons in the reaction of the drug and protein. The redistribution of the electrons is determined by the self-consistent calculation of the Schrödinger equation (eq. 12). The electron density, ρ, experience iterative changes until the total energy of the system (eq. 13) reaches a minimum.

$$\hat{H}\Psi = (\hat{T} + \hat{V}_{ext} + \hat{V}_{ee})\Psi = E\Psi \qquad \text{... eq. 12}$$

$$E(\rho) = T(\rho) + \int V_{ext}(\vec{r})\rho(\vec{r})dr + \frac{e^2}{2}\int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}\, d\vec{r}' + V_{xc} \qquad \text{... eq. 13}$$

In order to calculate the electronic binding energy, we calculate the energy of the complex with an ab-initio approach and compare it to the energy of the reactants alone, but maintaining the geometry as they are in the complex. Table 2 shows the electronic binding energy for all cases. At the redistribution of the electrons during the reaction, we have to recognize the two interaction energy components. One associated to the equalization of the chemical potential forming bonds, and the other involved in the electrostatic interaction due to the polarization of the charges.

The driving force energy is a renamed variable of the electron transfer contribution described before (Section 2.1). Having two components at different chemical potentials, a movement of the electron density from the high to low potential must occur in order to equalize the chemical potential. The transfer of charges represents a decrease of the electrostatic interaction. In my opinion, it is impossible from eq. 5 to predict the energy related to the change of electron density since it does not consider the properties of the final product. Instead, if the trend determined by the energy change of eq. 5 is followed, this energy change estimates the ionic character of the interaction before the reaction instead, that is, a sort of potential to the reaction due to the columbic charges.

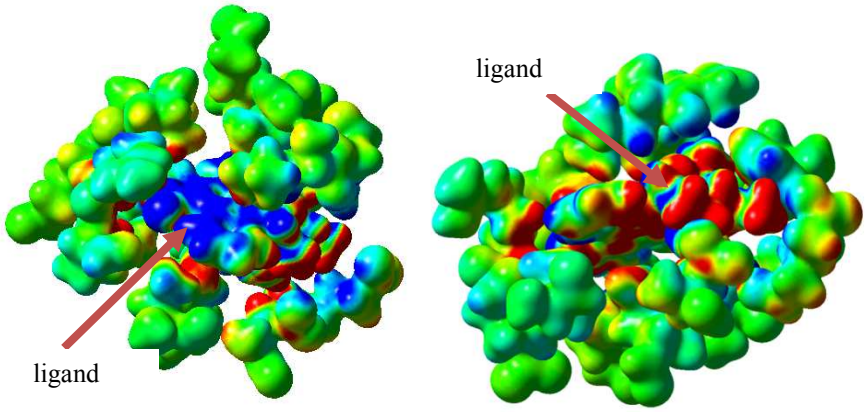
4.4 AB-INITIO: ELECTRON DENSITY CHANGE

As stated before (section 2.4), a higher charge transfer from the most negative to the most positive electrostatic molecule translates in a stronger bond. The calculated total charge transfer using the Mulliken population analysis is shown in Table 3.

Table 3: Total charge transfer from the protein to the drugs (neutral/protonated).

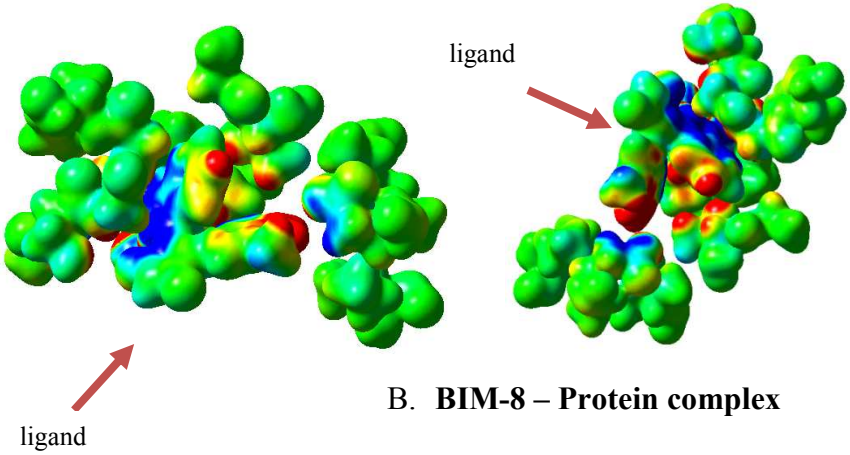
Staurosporine	BIM-8	BIM-11
-0.5/-0.9	-1.1/-0.4	-0.1/-0.3

In order to observe the local changes at every point of the active site, we have projected the electron density change in a 3-D feature. As shown in figure 13, the electron density change in the drug has two distinguishable regions, one will accumulate electrons (blue surfaces) coming from negative amino acids (and other regions of the drug), and the other will show depletion of electrons (red surfaces) that goes either to other regions of the drugs or to electrostatic positive regions of the protein. The overall of blue surfaces minus red surfaces will mean the total electron density transfer.

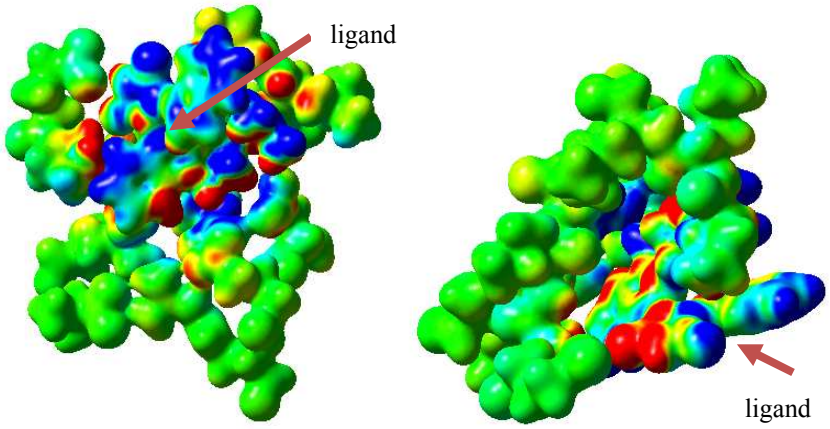


A. Staurosporine-Protein complex

Figure 12: Electron density change (blue means accumulation, and read means depletion of electrons). The electrons go from less negative chemical potential functional groups to more negative chemical potential groups



B. **BIM-8 – Protein complex**

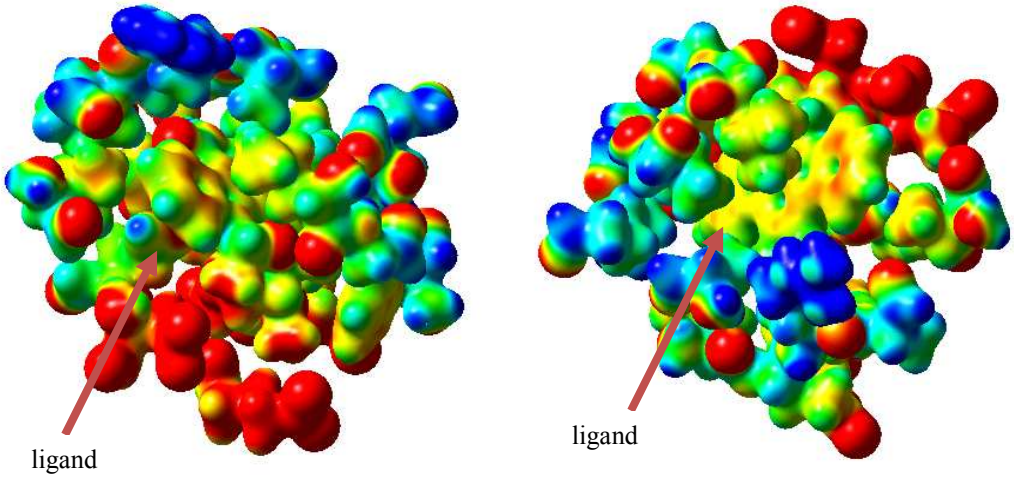


C. BIM-11 – Protein complex

Figure 12 Continued

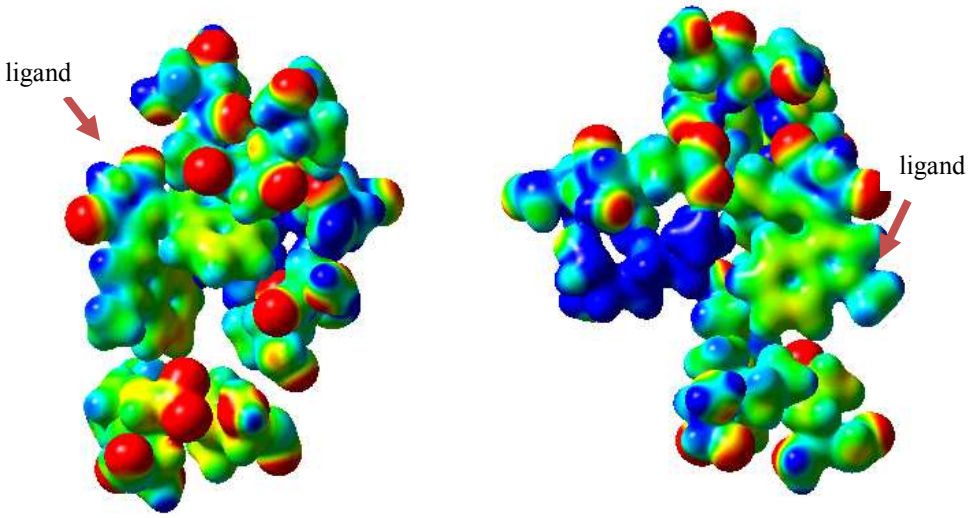
4.5 AB-INITIO: MOLECULAR ELECTROSTATIC POTENTIAL (MEP)

Analogously we present the 3-D feature of the molecular electrostatic potential of the final complexes (Figure 13). The MEP helps us to identify spots where the electrostatic interaction is strong in the entire space of the active site. Notice that in BIM-11, we had to go to higher isodensity surfaces to ensure a MEP with similar characteristics as in Staurosporine and BIM-8. The reason is that Staurosporine and BIM-8 has, from the electron density change results, a stronger electrostatic interaction than BIM-11(or weaker covalent bonding). Then, their electrons are more condensed in their nuclei. Thus going to higher isodensity surfaces means calculate the MEP at closer distances from the radius, compensating its higher condensation of electrons.

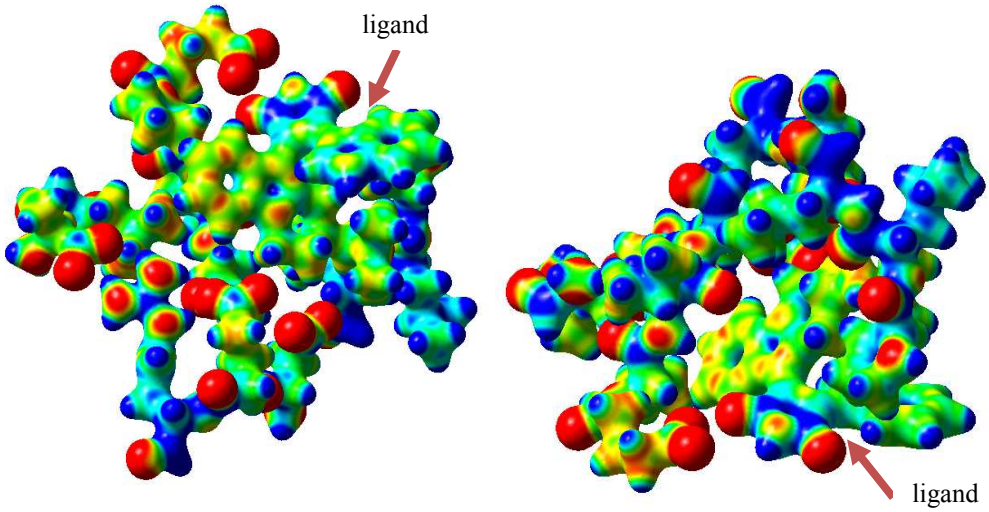


A. Staurosporine-Protein complex

Figure 13: Molecular electrostatic potentials (MEPs). Blue and red mean positive and negative electrostatic potential. Localized combination of these spots means a dominant ionic interaction



B. BIM-8 -Protein complex



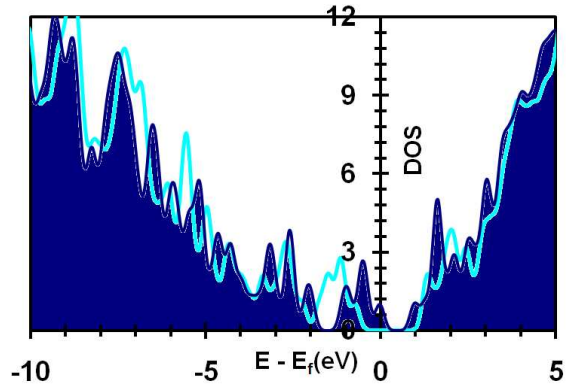
C. BIM-11 - Protein complex

Figure 13 Continued

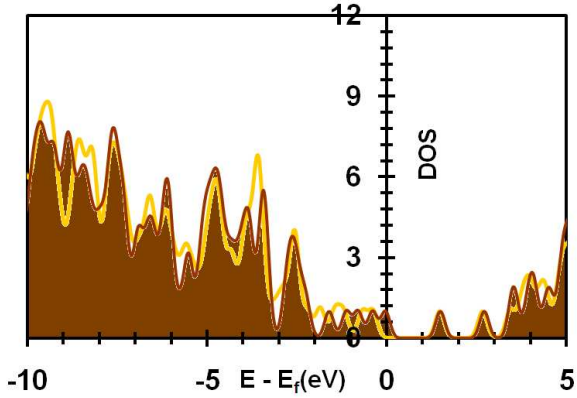
4.6 AB-INITIO: DENSITY OF STATES (DOS) PROJECTIONS

There are two characteristics to look up in the DOS projections that dictate the reactivity of the protein and their drugs. In short, the change from narrow to broad bands close to the Fermi level is a consequence of the strong chemisorption of the drug in the protein due to the stronger coupling between the valance states. Furthermore, a shift of the bands to more stable energy levels (shift to the left) is proportional to the strength of the chemisorption.

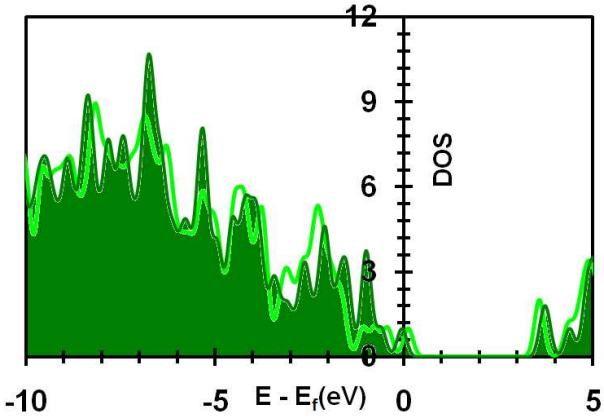
The DOS projections for each set of drugs, neutral and protonated, are shown in figure 14 and 15, respectively (shaded curves are the projection of the protein before the reaction and light curves are after the reaction). The combination of the two features of the DOS ranks the reactivity of every drug, following a similar trend determined by the electronic binding energy. staurosporine and BIM-8 does not show a change in their sharpness. Instead, they show downshifts in the energy of the bands. A greater shift in the case of staurosporine, compared to BIM-8, accounts for its greater electronic binding energy. Instead, BIM-11 shows a change into more wide bands but the increase in reactivity is diminished by an upshift of the energy of the bands. In the protonated set, all the reactions have almost the same downshift to more stable energy levels. But Staurosporine and BIM-11 have changes to more broad bands, accounting for higher electronic binding energies.



A. Protein affected by Staurosporine (neutral)

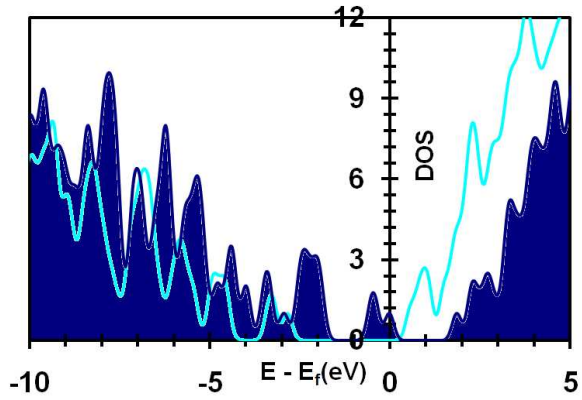


B. Protein affected by BIM-8 (neutral)

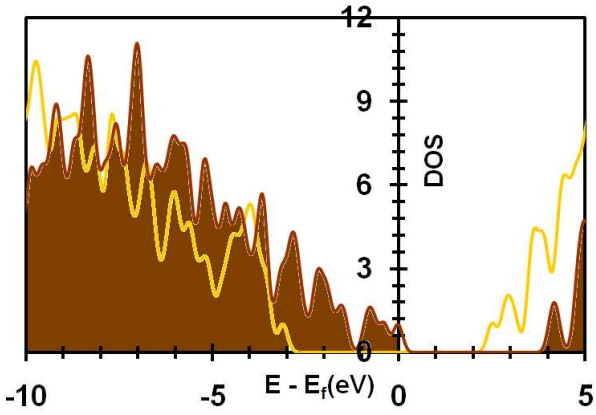


C. Protein affected by BIM-11 (neutral)

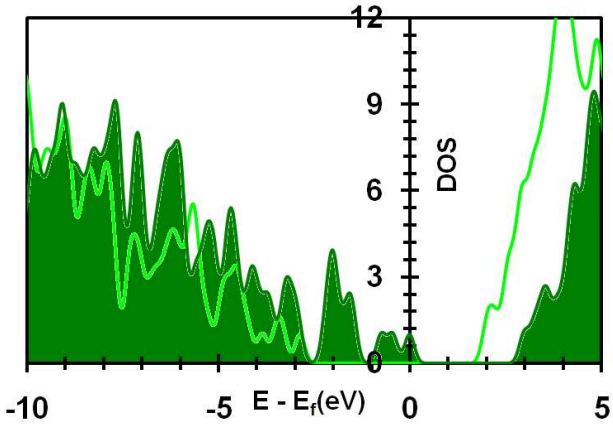
Figure 14: Density of States projections of the protein PIM-1 Kinase before (shaded curves) and after the reaction (light curves) with the neutral drugs. Change to broad curves and downshifts in energy mean the strength of the chemisorptions



A. Protein affected by Staurosporine (protonated)



B. Protein affected by BIM-8 (protonated)



C. Protein affected by BIM-11 (protonated)

Figure 15: Density of States projections of the protein PIM-1 Kinase before (shaded curves) and after the reaction (light curves) with the protonated drugs. Change to broad curves and downshifts in energy mean the strength of the chemisorptions

4.7 AB-INITIO: TRANSMISSION FUNCTION

The transmission functions dictate the probability of electron transport between donor and acceptors and, at the same time, is an estimation of the electronic coupling between the reactants. Actually the transmission is proportional to the square of the electronic coupling, making of the transmission function a variable sensitive to any change in the interactions of the reagents. There are several approaches to calculate the electronic coupling, but so far, no one has applied the Landauer approach to biological systems like we did. And the reason may be because the Landauer approach is based on molecular junctions. But looking at their derivation, the Landauer approach started with the same Marcus definition of the transfer rate of electrons. Figure 16 shows the results for both sets (a different scaling factor for each set has been applied in the calculation of the transmission, then comparison between the two graphs is not reliable), the high peak in the transmission of Staurosporine is a consequence of its strong covalent interaction; and also similar peaks of BIM-11 and Staurosporine in the protonated cases account for their similar strength of coupling.

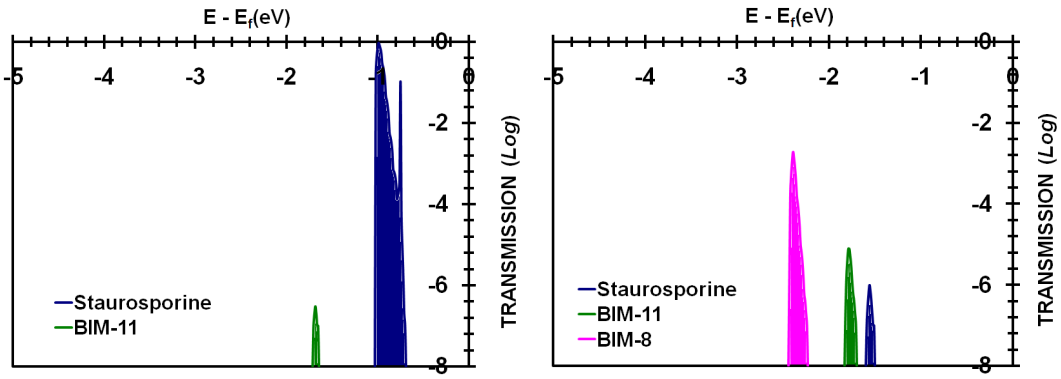


Figure 16: Transmission functions for the neutral set (left) and for the protonated set of drugs (right). Higher peaks mean higher rate transfer of electrons.

5. CONCLUSIONS

We have tested the most relevant computational chemistry techniques towards a fully understanding of the molecular interactions in biological systems like protein-ligand interactions.

In order to find a best correlation of the binding affinity with some electronic, thermodynamic or conformational property; we have analyzed the components of the interaction energy involved in the reaction, mainly, through an ab-initio approach.

We have proposed a protocol to analyze protein-drug systems by starting from docking calculation and molecular dynamics in order to take into account the nuclei fluctuations at natural conditions. Then we did a local study using first principles to account for the interactions of electrons, keeping fixed the nuclei (based on the Frank-Condon principle and/or the Born-Oppenheimer approximation).

We use the PIM-1 Kinase, responsible for diseases such as leukemia, and their known active drugs in order to test the analysis tools. The small difference in free energy makes of the discrimination of the drugs hard to do. Further molecular dynamics simulations are needed, but at this point where no electrons are treated explicitly, only qualitative analysis like RMSD fluctuations is done. The results show a better agreement with their respective experimental free energy

Then, we restrict the ab-initio studio to the active site of the protein since it is believed that any chemical reaction would have to be local. We calculated their electronic binding energies and we used the electron density change and molecular electrostatic potential to analyze the covalent and ionic interactions, respectively. A higher electron density change between the PIM-1 Kinase and their drugs corresponded to a stronger covalent bond. On the other hand, localized spots of electrostatic potential (positive and negative) are representation of strong ionic interactions.

We analyzed the chemisorption of the drugs by looking up the density of states projections of the protein before and after the interaction. We extrapolated the

conclusions from the tight binding theory, based on metal surfaces, and apply them to biological systems. The results are in agreement with the electronic binding energies. For example, the protonated set of drugs showed a higher effect in the downshift of the energy levels of the protein (shift towards more stable energies). Even more, Staurosporine and BIM-11 of the protonated set showed a change to broad bands due to a stronger chemisorption, accounting for higher electronic binding energies with respect to BIM-8.

Finally, we proposed a methodology based on the Landauer approach to calculate the electronic coupling which in turn will represent the transfer rate of electrons. The results showed that Staurosporine (neutral) accounts for the highest rate validating their lowest experimental free energy.

The combined analysis of the computational chemistry techniques used in this project gives us a better insight of the different molecular interactions in protein-ligand complexes, and furthermore it let us be able to distinguish their different interaction components.

REFERENCES

(1) DiMasi, J. A.; Hansen, R. W.; Grabowski, H. G. *Journal of Health Economics* **2003**, *22*, 151.

(2) Tang, Y.; Zhu, W.; Chen, K.; Jiang, H. *Drug Discovery Today: Technologies* **2006**, *3*, 307.

(3) Jorgensen, W. L. *Science* **2004**, *303*, 1813.

(4) Calder, J. A.; Wyatt, J. A.; Frenkel, D. A.; Casida, J. E. *Journal of Computer-Aided Molecular Design* **1993**, *7*, 45.

(5) Greco, G.; Novellino, E.; Silipo, C.; Vittoria, A. *Quantitative Structure-Activity Relationships* **1991**, *10*, 289.

(6) Prendergast, K.; Adams, K.; Greenlee, W. J.; Nachbar, R. B.; Patchett, A. A.; Underwood, D. J. *Journal of Computer-Aided Molecular Design* **1994**, *8*, 491.

(7) Klebe, G. Comparative Molecular Similarity Indices Analysis: CoMSIA. In *3D QSAR in Drug Design*, 2002; pp 87.

(8) Friesner, R. A. *PNAS* **2005**, *102*, 6648.

(9) Besalu, E.; Girones, X.; Amat, L.; Carbo-Dorca, R. *Acc. Chem. Res.* **2002**, *35*, 289.

(10) Wan, H.; Bergstram, F. *Journal of Liquid Chromatography & Related Technologies* **2007**, *30*, 681

(11) Alonso, H.; Bliznyuk, A.; Gready, J. *ChemInform* **2006**, *37*.

(12) Rognan, D. Molecular dynamics simulations: a tool for drug design. In *3D QSAR in Drug Design*, 2002; pp 181.

(13) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.

(14) Patel, M. A.; Deretey, E.; Csizmadia, I. G. *Journal of Molecular Structure: THEOCHEM* **1999**, *492*, 1.

(15) Silva, C. P. *International Journal of Quantum Chemistry* **2005**, *102*, 1131.

(16) Spiro, T. G.; Jarzecki, A. A. *Current Opinion in Chemical Biology* **2001**, *5*, 715.

(17) Wu, E.; Mei, Y.; Han, K.; Zhang, J. *Biophys. J.* **2007**, *92*, 4244.

(18) Dreyer, J. L. *Cellular and Molecular Life Sciences (CMLS)* **1984**, *40*, 653.

(19) Roberto Improta, V. B., Marshall D. Newton,. *ChemPhysChem* **2006**, *7*, 1211.

(20) Robert, G. P.; Robert, A. D.; Mel, L.; William, E. P. *The Journal of Chemical Physics* **1978**, *68*, 3801.

(21) Sanderson, R. T. *Science* **1951**, *114*, 670.

(22) Sanderson, R. T. Chemical Bonds and Bond Energy; Academic Press: New York, 1976.

(23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(24) Foloppe, N.; A. D. MacKerell, J. *Journal of Computational Chemistry* **2000**, *21*, 86.

(25) Böhm, H.-J. *Journal of Computer-Aided Molecular Design* **1992**, *6*, 61.

(26) Rarey, M.; Wefing, S.; Lengauer, T. *Journal of Computer-Aided Molecular Design* **1996**, *10*, 41.

(27) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. *Journal of Molecular Graphics and Modelling* **2003**, *21*, 289.

(28) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.

(29) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. R.; Fogel, L. J.; Freer, S. T. *Chemistry & Biology* **1995**, *2*, 317.

(30) Muegge, I.; Martin, Y. C. *J. Med. Chem.* **1999**, *42*, 791.

(31) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864.

(32) Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133.

(33) Yan, L.; Seminario, J. M. *International Journal of Quantum Chemistry* **2006**, *106*, 1964.

(34) Hammer, B.; Nørskov, J. K.; Bruce, C. G.; Helmut, K. Theoretical surface science and catalysis--calculations and concepts. In *Advances in Catalysis*; Academic Press, 2000; Vol. 45; pp 71.

(35) Landauer, R. *IBM J. Res. Dev.* **1988**, *32*, 306.

(36) Nitzan, A. *Annual Review of Physical Chemistry* **2001**, *52*, 681.

(37) Bautista, E. J.; Yan, L.; Seminario, J. M. *The Journal of Physical Chemistry C* **2007**, *111*, 14552.

(38) Derosa, P. A.; Seminario, J. M. *J. Phys. Chem. B* **2001**, *105*, 471.

(39) Yan, L.; Seminario, J. M. *International Journal of Quantum Chemistry* **2007**, *107*, 440.

(40) Seminario, J. M.; Cordova, L. E.; Derosa, P. A. *Proceedings of the IEEE* **2003**, *91*, 1958.

(41) Seminario, J. M.; De La Cruz, C. E.; Derosa, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 5616.

(42) Seminario, J. M.; Yan, L. M. *International Journal of Quantum Chemistry* **2005**, *102*, 711.

(43) Sotelo, J. A.; Yan, L.; Wang, M.; Seminario, J. M. *Phys. Rev. A* **2007**, *75*, 022511 (13 pages).

(44) Salazar, P. F.; Seminario, J. M. *The Journal of Physical Chemistry B* **2008**, *112*, 1290.

(45) Uner, D.; Tapan, N. A.; Özen, I.; Üner, M. *Applied Catalysis A: General* **2003**, *251*, 225.

(46) Bullock, A. N.; Debreczeni, J. E.; Fedorov, O. Y.; Nelson, A.; Marsden, B. D.; Knapp, S. *J. Med. Chem.* **2005**, *48*, 7604.

(47) Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. *Journal of Computational Chemistry* **2004**, *25*, 1605.

(48) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J Mol Graph Model* **2006**.

(49) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *Journal of Molecular Biology* **1982**, *161*, 269.

(50) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *Journal of Molecular Graphics* **1988**, *6*, 13.

(51) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. *Journal of Computational Chemistry* **1992**, *13*, 505.

(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; J. A. Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyenga, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Stefanov, J. C. B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

VITA

Pablo Felix Salazar Zarzosa received his Bachelor of Science degree in Mechanical and Electrical Engineering from The University of Piura, Peru in 2006. He entered the Graduate program at Texas A&M University in September 2003 and received his Master of Science degree in May 2009. His research interests include biomolecular simulations and catalysis. So far, he has five peer reviewed papers and four more to be published in the next months.

Name:	Pablo Felix Salazar Zarzosa
Address:	3122 TAMU, College Station, Texas
Email Address:	pablo1723@neo.tamu.edu
Education:	B.S., Mechanical and Electrical Engineering, University of Piura, Peru, 2006
	M.S., Chemical Engineering, Texas A&M University, 2009