AUTONOMOUS UNIVERSITY OF BARCELONA

Cationic Dummy Atom (CaDA) Software

by

Daniel Soler Viladrich

A thesis submitted in partial fulfilment for the
Bachelor of Physics
Bachelor of Chemistry

in the

Science Faculty
Physics Department
Chemistry Department

June 19, 2017

"Protons give an atom its identity, electrons its personality."

Bill Bryson

AUTONOMOUS UNIVERSITY OF BARCELONA

Abstract

Science Faculty
Physics Department
Chemistry Department

Physics and Chemistry Student

by Daniel Soler Viladrich

A long-standing challenge in modelling metalloproteins has been to rapidly define the binding properties of metallic centres. This task becomes more complex when there is a lack of empirical structural data. The project reports a dummy-atom-based software with a time-saving algorithm to determine qualitatively the stability of organometallic systems. The implementation of this algorithm automatises the parametrization process for transition metals. The software is tested on a modified carbonic anhydrase and an artificial metalloprotein. Results of the first test are in agreement with the experimental data. No binding affinity is found for the second system suggesting the proposed geometry is implausible. The results show potential applicability of the software in metalloprotein design.

Acknowledgements

I would like to use this section to thanks all people who have been involved in this project.

My warmest thanks to my project tutor and supervisor Prof. Jean-Didier Maréchal to give me the opportunity and guidance along the way.

Also, my special thanks of gratitude to Jaime Rodriguez, who gave me the golden opportunity to learn from him, coming to know about so many new things.

Nevertheless, I am also grateful to my tutor Dr. Carles Navau for their feedback, cooperation and support.

I would like to express my deepest appreciation to all Insilichem members, specially Lur, Josepe, Sergi and David Teixé whose contribution in stimulating suggestions and encouragement, helped me to finalise my project.

Last but not least, I would like to thank my girlfriend, parents and friends for accepting nothing less than excellence from me.

Contents

Chapter 1

Introduction

The primary goal of this dissertation is to present a new molecular modelling software to simplify the parametrization of metal centres in metalloproteins. Metal cations are of relevance in nature as they assist in processes like protein folding or binding to these folded proteins to tune their properties. Metalloproteins constitute almost half of all proteins in biology. Understanding the binding affinity and coordination geometry of metal cofactors would be of great value since it would elucidate mechanisms such as catalytic processes and signal transduction. This knowledge would also allow for the design of novel metalloproteins for biotechnological and pharmaceutical applications. Metalloprotein design is much more challenging than its non metallic counterpart since metals are much larger and have variable geometries. Metal-binding sights display high chromaticity, hence they exhibit magnetic properties which in the recent decade have been characterised through spectroscopic techniques. The implementation of this data in computation aids advances in representation of metalloprotein structures. This specificity, although yielding accurate results, is not functional due to time expenditure in cases where qualitative results are desired. In order to abet the process, this project aims to elaborate a software based on the cationic dummy atom approach. The software will automate the parametrization of metallic centres,therefore accelerating the representation of the desired system while determining the stability and geometrical properties at a reasonable accuracy. The improvement of this software could finally lead to faster but still accurate predictions for the comprehension of these systems.

1.1 Molecular Modeling

With the aim to achieve this goal, Molecular Modeling is needed, which acts as a bridge between microscopic world and macroscopic properties by approximating the interaction between atoms, usually derived from a potential energy, to obtain bulk properties predictions.

Molecular modeling has advanced tremendously from past years. It is no longer confined to drawn representations of chemical structures. Today, detailed molecular modeling can be accomplished using desktop computer, mimicking the behaviour of systems.

1.1.1 Historical Overview

Everything started between 1858 and 1861, when the three chemists Archibald Scott Couper, Friedrich August Kekulé von Stradonitz, and Aleksandr Mikhailovich Butlerov independently introduced the general rules of valence for organic chemistry, and the first written structures involving chains of carbons with lines drawn as "bonds" to substituent atoms and groups [1]. Also in 1861, Johann Josef Loschmidt produced the first "correct" structure of benzene. And in 1874 Jacobus Henricus van't Hoff and Joseph Achille LeBel determined that carbon compounds have tetrahedral geometry [2].

In the earliest twentieth century the concept of forcefield appeared, from vibrational spectroscopy, which considered the forces acting between every pair of atoms in the molecule, or in a lattice in the case of ionic crystals. However, this concept was not embraced by the physics and chemistry community until more quantitative approaches arose.

Despite the inspired work described before this new computational concept did not yet widely influence the thinking of the chemical community. This changed abruptly in 1950 with the publication of Barton's short note on how the conformations of steroids affect their chemistry [3], which laid the foundation of conformational analysis. From then onward, an appreciation of the three-dimensional aspect became crucial to understanding structure, stability, conformation and reactivity.

In the same year, a group of scientists from Los Alamos published their studies of, "Equation of State Calculations by Fast Computing Machines [4]". This work, carried out a computed-based Monte Carlo methods, established the Metropolis algorithm (from the first-named author) for simulated annealing, and was the ancestor of molecular dynamics calculations.

Following the earlier successes of Monte Carlo simulations, in 1957, Alder and Wainwright used an IBM 704 computer to simulate perfectly elastic collisions between hard spheres. Later in 1964, Rahman published landmark simulations of liquid argon that used a Lennard-Jones potential where calculations matched well with experimental data. Since then, molecular dynamics (MD) is frequently used to refine three-dimensional structures of proteins and other macromolecules based on experimental constraints from X-ray crystallography or NMR spectroscopy.

In the last few years, molecular dynamics have been greatly improved allowing the influence of chemical effects, electronegativity, and hyperconjugation on molecular structure and properties. Thus, being able to simulate more successfully metallic systems, catalytic and surface reaction as Yang Jiang and Tianwei Tan did in 2015 with a refined dummy atom model of ion magnesium as shown in figure 1.1.

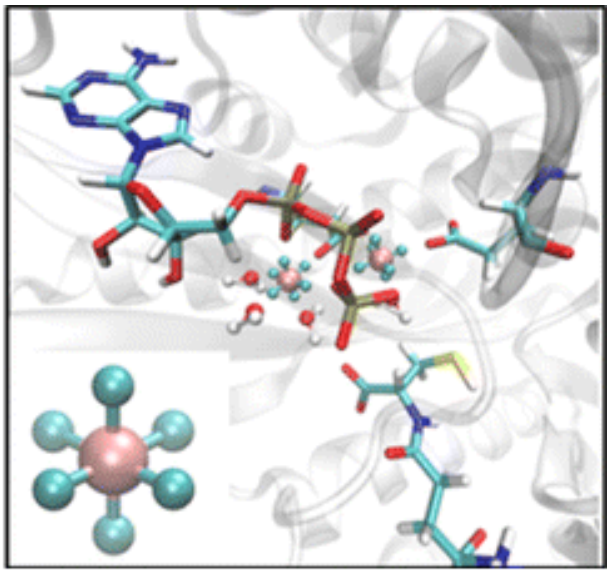

FIGURE 1.1: Refined Dummy Atom Model of Mg2+ by Simple Parameter Screening Strategy with Revised Experimental Solvation Free Energy

Having said that, there is still much to do in understanding transition metal behaviour. In order to delve into the comprehension of this subject it would be appropriate to have a general overview on the theory behind the steps that have built up to the development of this work.

1.1.2 Atomic description level

Concerning the treatment of atomic properties, the field of molecular modelling is divided in several methods. Their applicability depends on the type and desired atomic description of the system to study.

For quantitative calculations and accuracy, quantum mechanics is the best method up to date. It models electronic rearrangements during simulations which are essential for those parts of the system involved in the process. For a less accurate but faster result Molecular Mechanics is enough. The system is Newtonianly treated and no electron distribution, in other words, polarization is allowed. For something in between, one could use a hybrid method where most part of the system would be studied classically, but certain atoms from the active site would be held under the quantum mechanic laws.

1.1.2.1 Quantum Mechanics

Quantum Mechanics stem from the Schrodinger equation 1.1 first brought to light in the 1920's.

$$ih\frac{\partial}{\partial t}\psi(r,t) = [\frac{-h^2}{2\mu} \bigtriangledown^2 + V(r,t)]\psi(r,t) \tag{1.1}$$

This method treats the system as a collection of nuclei and electrons without any input concerning chemical bonds. Calculations are made by solving the system's Schrodinger equation. However, this equation cannot actually be solved for any but a one-electron system, and approximations need to be made. Depending on these approximations a wide scoop of chemical models and methods can be found in terms of their capability and reliability and their "cost".

There are methods directly derived from theoretical principles, as Hartree Fock and density functional theory (DFT) typically used to solve the time-independent Schrödinger equation for a multi-electron atom or molecule.

These methods make four major simplifications.

1. The Born–Oppenheimer approximation where the motion of atomic nuclei and electrons in a molecule can be separated as a result of their big difference in mass.
2. The momentum operator is assumed to be non-relativistic.
3. The solution is a linear combination of a finite set of basis. The computational cost will be directly related with the type and number of basis chosen.
4. The mean field approximation is applied, instead of evaluating the electronic density of each electron over a nuclei, an average density function for all electrons is calculated.

With all the approximations above, the equation 1.2 is finally resolved obtaining the potential energy surface (from which we can get the equilibrium geometry and vibrational frequencies) and the electronic wave function ψ(r; R), which contains lots of useful information about molecular properties such as dipole moments and polarizability.

$$[T_e(r) + V_{eN}(r;R) + V_{NN}(R) + V_{ee}(r)]\varphi(r;R) = E_{el}\varphi(r;R) \tag{1.2}$$

1.1.2.2 Molecular Mechanics

Systems treated within molecular mechanics are made up of atoms and bonds (as opposed to nuclei and electrons), and studied through forcefields, or functional forms used to determine the potential energy of the system.

Forcefields are split into two general terms, bonded and non-bonded interactions.

$$E_T = E_{bonded} + E_{non-bonded} \tag{1.3}$$

The latter is traditionally studied through one, two or even three body interaction terms. The former, is calculated as a sum of individual terms, as seen in equation 1.4.

$$\begin{aligned} E_{bonded} &= E_{streching} + E_{bending} + E_{torsions} \\ E_{non-bonded} &= E_{vanderwals} + E_{electrostatic} + E_{crossterms} \end{aligned} \tag{1.4}$$

The final aim is to calculate the molecular mechanics energy equation given below.

$$E_T = E_{streching} + E_{bending} + E_{torsions} + E_{vanderwals} + E_{electrostatic} + E_{crossterms} \tag{1.5}$$

Where $E_{streching}$ is the energy involved in the deformation of a bond, either by stretching or compression and $E_{bending}$ is the energy involved in angle bending.

To determine them, bonds are treated as springs and described by equation 1.6 based on the Hook's law.

$$E_{streching} = \sum_{bonds} k_b(r - r_o)^2 + \sum_{angles} k_\theta(\theta - \theta_o)^2 \tag{1.6}$$

$E_{torsions}$ is the torsional angle energy, which represents the amount of energy that must be added or subtracted to the other terms to make the total energy agree with the experimental value. This one is described through the periodical function 1.7.

$$E_{torsions} = A[1 + cos(n\tau - \phi)] \tag{1.7}$$

$E_{vanderWaals}$ is the energy related with the distance dependent interactions between atoms. On the one hand, van der Waals attraction occurs at close distance, and rapidly dies off as the interacting atoms separate a few Angstroms. On the other hand, repulsion occurs when the distance between interacting atoms becomes even slightly less than the sum of their contact radius. That is summed up on the equation 1.8, where the repulsive term diverges at short ranges and the A and B parameters control the depth and the position of the potential energy well. However, this is not the only equation that can describe this type of behaviour.

$$E_{vanderWaals} = -\frac{A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \tag{1.8}$$

$E_{electrostatic}$ is the energy involved in interactions between atoms that are not directly bonded. This can be modeled using a Coulombic potential. Where, electrostatic energy depends on the charge of the non-bonded atoms, their interatomic distance, and a molecular dielectric expression that accounts for the attenuation of electrostatic interactions by the environment.

One application were MM is found is molecular dynamics simulations (MD) based on the ergodic hypothesis, instead of calculating the partition function to obtain ensemble averages, such as energy, values are replaced by time averages over the simulation.

All of this is obviously much simpler than solving the Schrödinger equation for electron motions, but requires an explicit description of “chemical bonding”, as well as a large amount of information about the structures of molecules.

Pharmaceutical companies, research groups and many other members of the scientific community have been using MD to perform ligand dockings on the protein active site, for instance in drug design, the mechanism of inhibiton of protein kinases[5] (targets for treatment of a number of diseases) has been illuminated by the use of several of this modelling methods.

1.2 Molecular Mechanics. Metal Systems Modelling

As said in the introduction, a big number of processes in biological systems rely on the powerful reactive properties of transition metal ions. However, the theoretical modelling of transition metal centres is a notoriously difficult task. The metal-ligand interactions and the combination of suitable analytic interaction potentials can be more complex than expected. The situation can get more difficult by the fact that at finite temperature, the system might switch dynamically between different bonding situations which are characterized by different coordination numbers and geometries.

To represent all these behaviours several quantum mechanics methods shed light to this field in the late 1980s as new and better computational approaches were created. The first of them was Hartree Fock, a variational method based on the Eckart theorem which uses a linear combination of functions to described orbitals and calculate an upper bond to the true ground state energy of a given molecule.

A few years later DFT appeared. This revolutionary method uses the Kohn–Sham equation, 1.9, defined by a fictitious external potential denoted as Vs(r) and a N-particle density function to determine the energy of a system 1.10.

$$[\frac{-h^2}{2m} \bigtriangledown^2 + Vs(r)]\theta(r) = E_i\theta(r) \tag{1.9}$$

$$E_{tot(\rho)} = E_{kin(\rho)} + E_{ext(\rho)} + E_{H(\rho)} + E_{xc(\rho)} \tag{1.10}$$

Where $E_{kin(\rho)}$ is the Kohn–Sham kinetic energy , $E_{ext(\rho)}$ the energy due to external potential (usually ions), $E_{H(\rho)}$ the classic Hartree (or Coloumb) repulsion (e- - e-) and $E_{xc(\rho)}$ the exchange-correlation energy.

Even though Kohn-Sham equation is exact, the exchange-correlation energy functional is not known explicitly. That is the main reason the method is not absolutely accurate and it is also the reason for the derivation of various exchange correlation functionals.

All in all, the two approaches above and some hybrids of both are currently the most used on metal modelling. Even though, due to their high computational cost they are not feasible for large scale systems, and quantum mechanical/molecular mechanical (QM/MM) method needs to be used.

The hybrid QM/MM is a multiscale technique, which responds to two necessary requirements for the modeling of metal binding enzymes. The first is the complexity of the

system and their computational cost and the second one is the realism of the model system to be as accurate as possible to reality.

Unfortunately, most MM force fields do not take into account the electronic degrees of freedom, experiencing large difficulties in describing the metal environment accurately. In addition, force fields do not account for bond breaking and forming events. To address these issues, is it necessary to move to the parameter-free quantum mechanical (QM) approaches on certain areas of the system and keep the rest under the MM laws.

Even though when aiming to describe systems qualitatively the process explained previously is severely time consuming and its computational cost is too high. In order to pose a solution to this problem MM could be used as a more efficient alternative.

1.2.1 Qualitative Thermodynamic Approaches

Forcefield parameters are available for the 20 natural aminoacids and the five base pairs of nucleic acids. Nonetheless, in case of metal centres these parameters are related to the nature of the metal centre, its geometry and coordination. Therefore, they must be calculated and optimise for each system.

In spite of the intrinsic hardships of metal system parametrization, there have been some attempts of MD simulations applied to metalloproteins with the purpose of understanding the dynamic properties and analysis of the internal motions of the metal system.

There are multiple ways to model metal ions and each of these have their associated pros and cons. The process of decision-making will depend on how we want to define our system and the answers one is wishing to extract.

1.2.1.1 Covalent Bond Approach

This method considers bonding between the metal centre and the ligands around, treating them as springs and using the harmonic oscillator potential to describe their behaviour.

This approach predefines covalent bonds between the metal and ligands, therefore failing to describe ligand and geometry exchange. The covalent bond method has other relevant drawbacks such as the large number of parameters to be defined and double counting of the electrostatic and Lennard-Jones interactions.

Ligand exchange generally is not described, yet in some cases it occurs at a time scale of 5·10-4 microseconds. Taking into account that MD time steps are per usual around the femtosecond order, the bonded model is an effective approach to study these processes. This occurs due to the fact that the time step is much larger than the ligand exchange mechanism.

1.2.1.2 Non Bond Approach

In contrast to the method described in the section above, there is the non bond approach where metal–ligand interactions are described through electrostatic and van der Waals potentials. Thus, allowing for ligand exchange.

On the downside, it appears to be inadequate when it comes to more complex situations such as multinuclear metal center systems, closely located metal ions and transition metals. In the latter case, the challenge is to obtain a parameter set that can simultaneously reproduce both solvation free energies and metal–water distances.

Withal, it is still a very functional method to study simple systems without need for detailed parametrisation.

1.2.1.3 Hybrid Methods. Cationic Dummy Atom

In the end one realises that there are two sides of the coin, a good description of the electronic interaction or a strong bonded model. Even though these approaches clash, one can still wonder if it would be possible to have an in between both.

Currently, the solution is the cationic dummy atom model. This model places charges between the metal ions and surrounding residues to mimic covalent bonds and offer a more sophisticated electrostatic model. However, it still requires the tedious parametrisation of not just the metal centre, but also the charge and VdW on each site.

1.3 Cationic Dummy Atoms Software

CaDAS implements all advantages cited before, and as a result composes a Mollecular Dynamics input, whilst removing the time-wise handicap. All this is done through a light and simple interface adapted to users of any level, for the most frequently found metal centre geometries (tetrahedral, square planar, octahedron and square pyramid).

CaDAS consists of several dummy atoms that bind to the metal centre with the metal charge evenly transferred between them. This metal-dummy attachment is purely electrostatic even though it affixes the dummies to a static position with respect to the metal, while still emulating a covalent bond behaviour. In order not to loose the dummy geometry during simulation, these are attached to one another creating a constraints that keeps all of them packed as in figure 1.2.

The spreading of the charges between the surrounding dummies tries to copy the charge distribution density inside the metal atom orbitals, making the non covalent character of the system more realistic.

The largest part of the software is smoothly ran on the background creating all mass, bond, angle, dihedral and charge parameters the model needs. Hence, it removes the most laborious part of the approach simplifying the process for users.

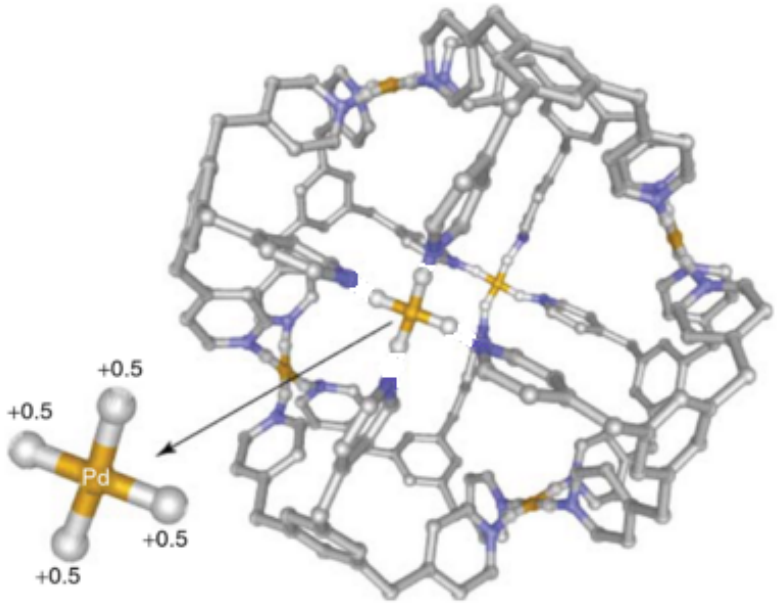

FIGURE 1.2: Pd(II) using cationic dumy atoms (CaDA).

To implement CADAS, Chimera UCSF and Tleap have been used and wrapped around python language. Both of them having a set of unique characteristics that make them essential to this software.

However, the launch of the MD simulation will be performed by an external high performance software called OpenMM.

1.3.1 UCSF Chimera

UCSF Chimera is a highly extensible program for interactive visualization and analysis of molecular structures and related data, including density maps, supra-molecular assemblies, sequence alignments, docking results, trajectories, and conformational ensembles.

In addition, UCSF Chimera gives users the opportunity to work through a python shell. The UAB research group Insilichem seized this characteristic to create PyChimera, which provides access to Chimera's modules from any Python 2.x interpreter keeping the interface out of the equation.

This last tool will be one of the two major external softwares used in CaDAS, which will extract any necessary information of the system in order to create the search and orientation algorithms that will be seen later on in the methodology.

1.3.2 Tleap

Amber is the collective name for a suite of programs that allow users to carry out molecular dynamics simulations, particularly on bio molecules. The various parts work together reasonably well, and provide a powerful framework for many common calculations.

tLEaP is the primary program to create a new system in Amber, or to modify existing systems. Unlike xLEaP, tLEaP works without interface, only using a simple command line.

Therefore, tLEaP will be in charge of building up the system coordinates and topology, in other words, the MD input.

1.3.3 OpenMM

OpenMM is a high performance toolkit for molecular simulation that can be used as a library, or as an application. Its extreme flexibility through custom forces and integrators, as well as its extreme performance through GPU Acceleration, with optimizations for AMD, NVIDIA, and Intel Integrated GPUs make the software on of the most powerful MD tools around the scientific community.

However, OpenMM is a complicated library to use and sometimes not user friendly for beginners. Hence, a new software called ommprotocol was created by Jaime Rodriguez and Jean Didier Marechal to make OpenMM MD launches simpler.

Consequently, this last program will be the one handling the MD launch.

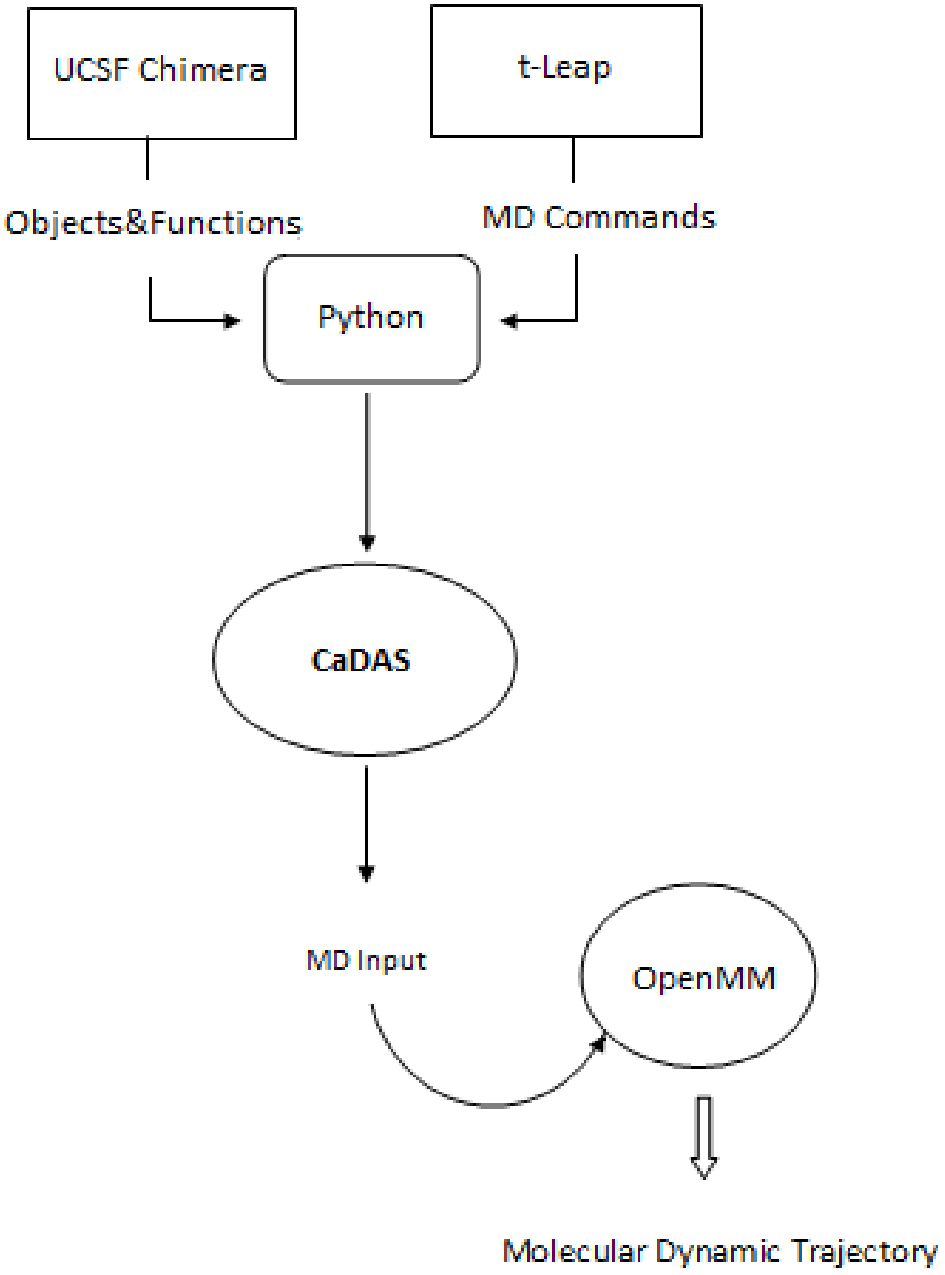

FIGURE 1.3: Software Overall Structure

Chapter 2

Methodology

The aim of this section is to display how this project was conducted by following the guide lines shown in the previous section.

2.1 CaDAS

In order to do that, the structure of the software will be broken down into three main parts. Prior to delving into the description of the software's structure, it would be helpful for one to understand first the general flowchart of the program.

2.1.1 Flowchart

The flowchart is subdivided into three main sections, as we can see in figure 2.1.

In the first place, the given input is a .pbd file usually obtained from the protein data bank or from another program's output, which contains the topology and coordinates of the system to be studied.

Next, two different processes are applied to the system. The aim of the first process, shown in orange in figure 2.1, is the addition of the dummies while conserving the chemical properties of the system. If one included the desired dummies directly, neither the mass or charges of the metal system would be the same as the original ones. This will be explained in more depth in the next section.

The second, in blue, is the acquisition of the system's physical properties such as masses, coordinates, bond length, Van der Waals radius and number and position of its dummies.

As it was mentioned before, the anatomy of the system is fundamental in MM, in order to get the right description of the system which is being modelled.

Last but not least, shown in green, is the final phase of the software, where the program recollects the system's anatomy and its properties from the steps before, producing the simulation input.

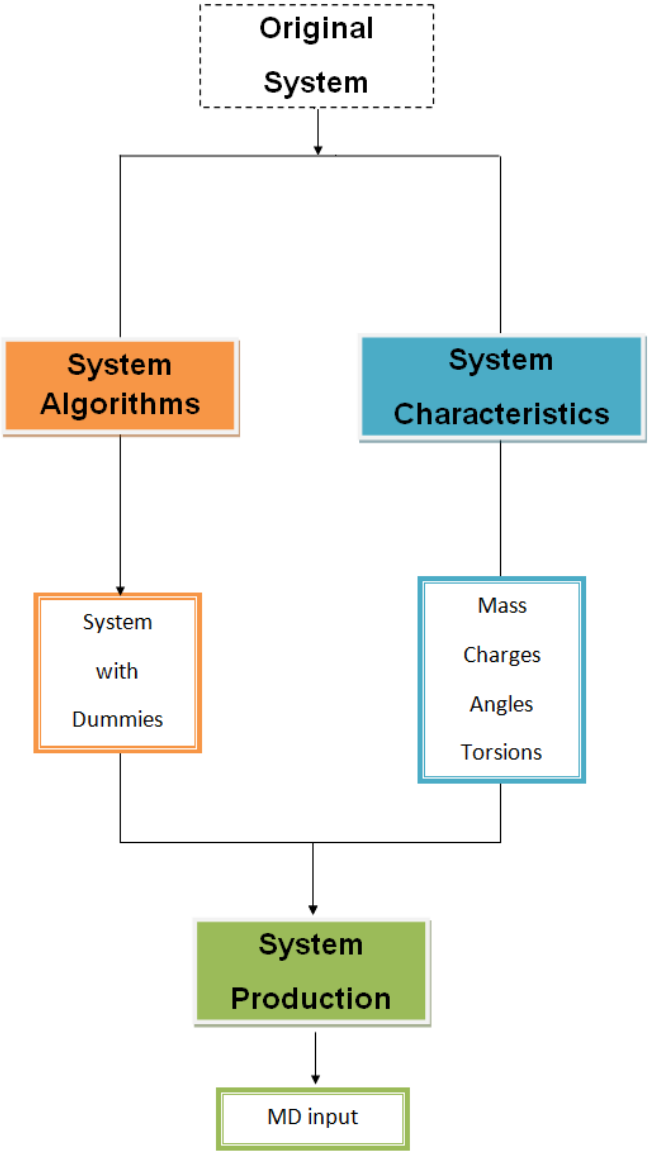

FIGURE 2.1: CaDAS Flowchart

2.1.2 System Algorithms

Prior to the dummies being included, their optimum position must be known to avoid problems at the very first stages of simulation. Finding out these coordinates is the main goal of this section.

In the initial step, the metal center ligands will be identified out of all the other atoms in the system. Having done that, the most optimum orientation of the dummies in regards to RMSD against the initial ligand position, will be determined taking into account the user's given geometry and the ligand positions obtained in the step before.

Ultimately, every dummy will be carefully placed in the system while atom types are assigned to them. This is explained more in depth later on in this section.

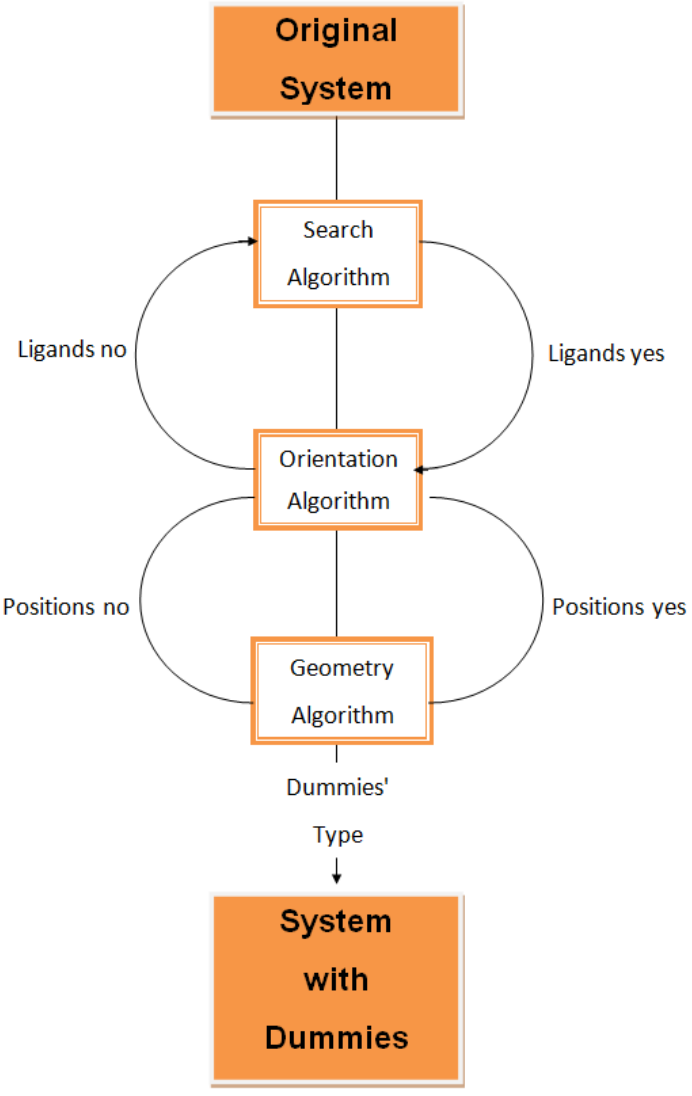

FIGURE 2.2: System Algorithms Flowchart

2.1.2.1 Search Algorithm

The heart of the program is a ligand-search algorithm around the metal using UCSF Chimera objects and wrapping them with python.

The algorithm starts by looking for the closest atoms around the metal. Once this search is completed, for the list of atoms selected, the software does the following. First, for every atom in the list, the conditions stated below are checked in the order which they are written. Once one of these conditions is not fulfilled the atom is discarded as a possible ligand and the function moves on to the next atom. Finally, the algorithm retrieves the chosen ligands.

- The atom can not be another metal. Therefore, CaDAS excludes metal-metal bridges.
- The atom valence must be greater than 5 after withdrawing the electron pair which will coordinate to the metal center. Otherwise, this pair of electrons will be placed in an inner shell being unable to participate in the formation of a chemical bond with the metal center.

```
valence = (atomic number - 2)%8 < 5
```

- The atom must have a lone pair or it will not be able to have any interaction with the metal center.
- The atom can not be a hydrogen since H atoms do not act as ligands.
- The coordinating distance between metal and ligand must be smaller than 4A or the interaction between both will not be strong enough to create a chemical bond.
- The atom must still have a free coordinating place to bond with the chosen metal.
- The angle of the atom must be close to the requested geometry. Otherwise it will not be possible to create the desired model.

```
for neighbour in neighbours_candidates:
    if angle(neighbour, candidate_to_ligand, metal)< angle_Cutoff:
    break
```

2.1.2.2 Orientation Algorithm

The atoms that fulfil the conditions from above will be treated as ligands from the studied metal center. They will undergo an orientation process to find which is the best position to place the dummies, taking into account the desired geometry and the positions of the ligands.

This part of the code is mostly geometrical, as half of the dummies are fixed to the right position (on the line connecting the metal center and the ligand) and the others are rotated around the metal while calculating the RMSD between the ideal position and the one found for each conformation. Once all of them are explored, only the best one is reported.

Lastly, the algorithm returns the exact position of all dummies for this geometry as seen in the pseudo code below.

```
for direction in optimum_directions:
    direction.length = dummy_metal_bondlenght #user input
    dummyposition = self.metal.labelCoord() + direction
    dummies_xyz.append(dummyposition)
return dummies_xyz
```

2.1.2.3 Geometry Algorithm

The final step is to obtain the right dummies' position with the right atom types.

An atom type is a label that defines how an atom will interact with its neighbours. In order to do that every atom type must bring angles, torsions and hybridisation associated. A mistake on these labels could lead to an undesired system topology.

To achieve the correct correlation between dummy position and type, CaDAS labels a dummy with a given type. The next step is to find the dummies that are equivalent in type, which means they are found at a specified geometry with respect to the labeled atom. The atoms that fulfilled this condition are labeled with the same type as the reference atom, as seen in figure 2.3. This is achieved by iterating through all the other dummies coordinates searching for an approximated 180 degree angle. In case there is a unique atom (only atom within its type) as in the square pyramid, a unique label is assigned to it.

This works for all the geometries of the software except the tetrahedral one, which does not need atom type assignation, due to the fact that all atoms are equivalent and consequently the same type.

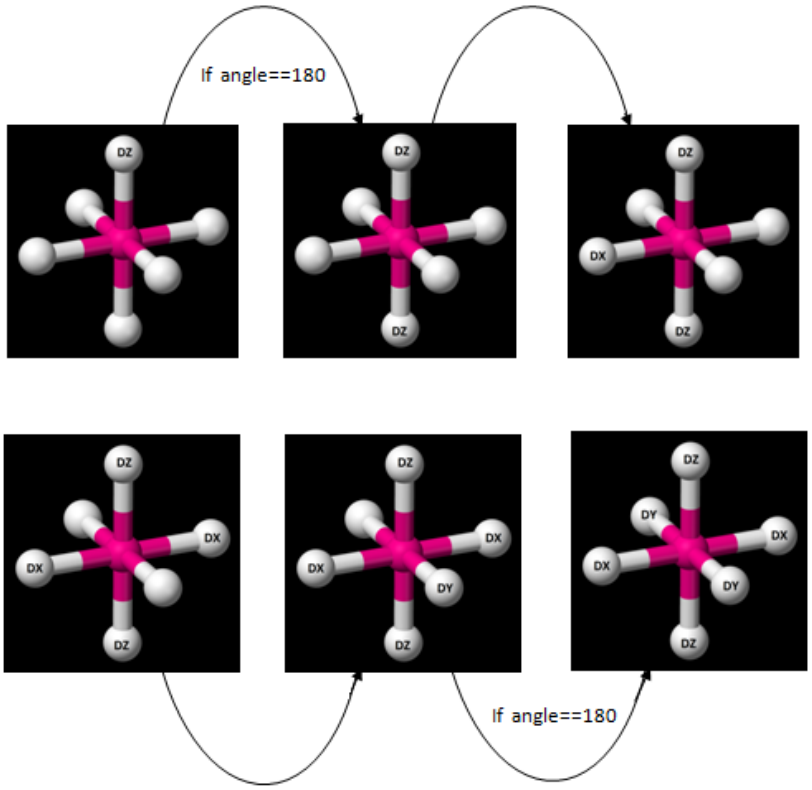

FIGURE 2.3: Demonstration of the Geometry Algorithm for an octahedron

Once all dummies are set into place, the final topology of the system is extracted as a pdb file. Having done that, the system is now ready to be parameterized in the next section of the software.

2.1.3 System Properties

This section intends to yield the metal centre properties and not the whole system properties. To do that, two processes are applied to the original system.

First, a new virtual system including the metal centre and the dummy atoms is created on the computer's RAM memory, aside of the original system. In the following step, the virtual system is modelled to the users preference by including charge on the dummies, emulating orbital density, the metal-dummy connectivity as well as bonds between dummies, to keep them packed together. Finally, a t-leap command will record all the preceding information in a .lib file.

Second, a template file of format .frcmod, which contains all the metal system parameters such as mass, angles and torsions is filled in with the user's inputs. Initially setting little portions of mass to the dummies, making them susceptible to be affected by Newtonian forces (which otherwise would be null and no movement would occur) and then, describing all the bonds, angles and torsions for each dummy type.

It is important to remember the fact that every time some portion of mass or charge is added to a dummy, it is, at the same time, deducted from the metal centre in order to not change the system's properties.

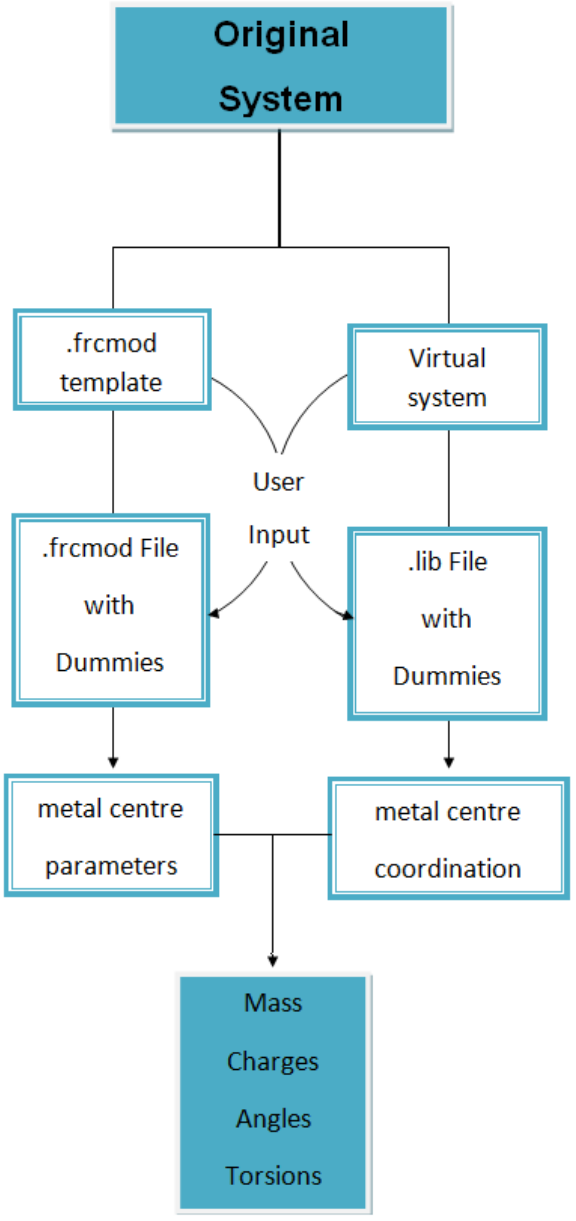

FIGURE 2.4: System Characteristics Flowchart

2.1.4 System Creation

The final phase of the software gathers the .lib, .frcmod and .pdb files in one single file ready to be launched by the simulation program OpenMM.

All of the above is performed with t-Leap, which combines topology, coordinates, charges and all the remaining parameters of the system to create a unique simulation input divided into two files, the final topology and the system coordinates. In addition, a .log error file is outputted in case the software lacks information on the system. All the previous is summarised in the figure 2.5.

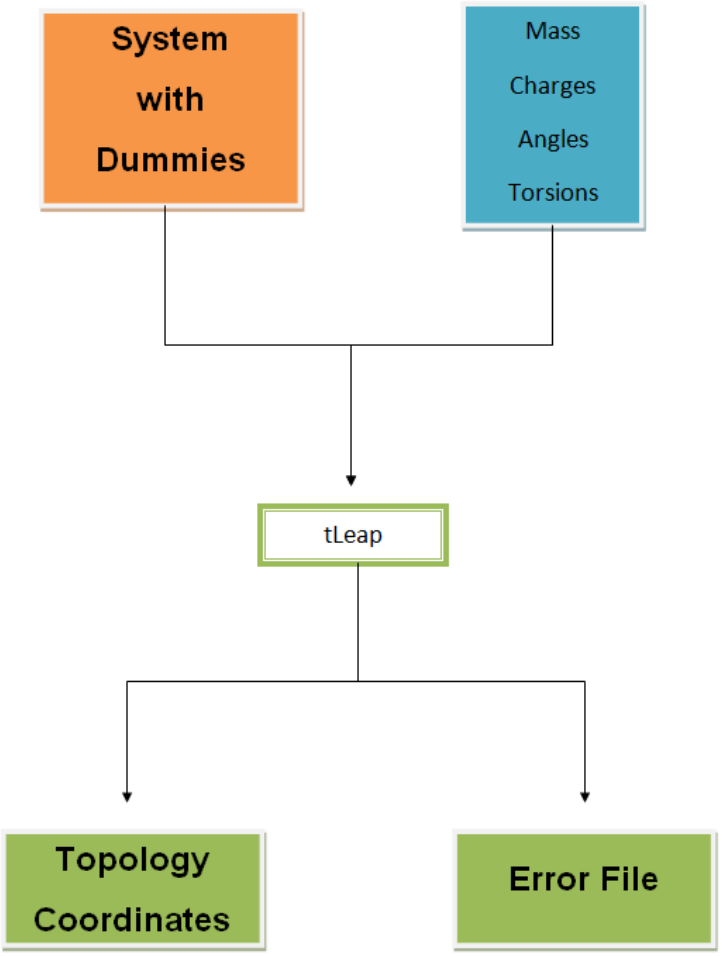

FIGURE 2.5: System Creation Flowchart

Chapter 3

Results

This chapter sets out the results of the project. Initially assessing the qualitative variables to establish whether or not the software is functional. This will be followed by a quantitative analysis of all simulations. Finally ended by a critical discussion of all findings and their linkage with future for improvements.

3.0.1 Zinc System

3.0.2 Catechol Sytem

3.0.3 Cage System

Appendix A

An Appendix

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus at pulvinar nisi. Phasellus hendrerit, diam placerat interdum iaculis, mauris justo cursus risus, in viverra purus eros at ligula. Ut metus justo, consequat a tristique posuere, laoreet nec nibh. Etiam et scelerisque mauris. Phasellus vel massa magna. Ut non neque id tortor pharetra bibendum vitae sit amet nisi. Duis nec quam quam, sed euismod justo. Pellentesque eu tellus vitae ante tempus malesuada. Nunc accumsan, quam in congue consequat, lectus lectus dapibus erat, id aliquet urna neque at massa. Nulla facilisi. Morbi ullamcorper eleifend posuere. Donec libero leo, faucibus nec bibendum at, mattis et urna. Proin consectetur, nunc ut imperdiet lobortis, magna neque tincidunt lectus, id iaculis nisi justo id nibh. Pellentesque vel sem in erat vulputate faucibus molestie ut lorem.

Quisque tristique urna in lorem laoreet at laoreet quam congue. Donec dolor turpis, blandit non imperdiet aliquet, blandit et felis. In lorem nisi, pretium sit amet vestibulum sed, tempus et sem. Proin non ante turpis. Nulla imperdiet fringilla convallis. Vivamus vel bibendum nisl. Pellentesque justo lectus, molestie vel luctus sed, lobortis in libero. Nulla facilisi. Aliquam erat volutpat. Suspendisse vitae nunc nunc. Sed aliquet est suscipit sapien rhoncus non adipiscing nibh consequat. Aliquam metus urna, faucibus eu vulputate non, luctus eu justo.

Donec urna leo, vulputate vitae porta eu, vehicula blandit libero. Phasellus eget massa et leo condimentum mollis. Nullam molestie, justo at pellentesque vulputate, sapien velit ornare diam, nec gravida lacus augue non diam. Integer mattis lacus id libero ultrices sit amet mollis neque molestie. Integer ut leo eget mi volutpat congue. Vivamus sodales, turpis id venenatis placerat, tellus purus adipiscing magna, eu aliquam nibh dolor id nibh. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Sed cursus convallis quam nec vehicula. Sed vulputate neque eget odio fringilla ac sodales urna feugiat.

Phasellus nisi quam, volutpat non ullamcorper eget, congue fringilla leo. Cras et erat et nibh placerat commodo id ornare est. Nulla facilisi. Aenean pulvinar scelerisque eros eget interdum. Nunc pulvinar magna ut felis varius in hendrerit dolor accumsan. Nunc pellentesque magna quis magna bibendum non laoreet erat tincidunt. Nulla facilisi.

Duis eget massa sem, gravida interdum ipsum. Nulla nunc nisl, hendrerit sit amet commodo vel, varius id tellus. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc ac dolor est. Suspendisse ultrices tincidunt metus eget accumsan. Nullam facilisis, justo vitae convallis sollicitudin, eros augue malesuada metus, nec sagittis diam nibh ut sapien. Duis blandit lectus vitae lorem aliquam nec euismod nisi volutpat. Vestibulum ornare dictum tortor, at faucibus justo tempor non. Nulla facilisi. Cras non massa nunc, eget euismod purus. Nunc metus ipsum, euismod a consectetur vel, hendrerit nec nunc.

Bibliography

[1] H. Polymers. Morawetz. The origins and growth of a science. *J. Wiley&Sons: New York*, pages 7–15, 1985.

[2] T. W. G. Solomons. Organic chemistry, 3rd edition. *J. Wiley&Sons: New York*, 1984.

[3] DHR Barton. Experienta. 1950 6 316.

[4] AW.; Rosenbluth MN.; Teller AH. Teller E Metropolis, N.; Rosenbluth. J. chem. phys. 1953 21 1087.

[5] Louise N. Johnson Martin E. M. Noble, Jane A. Endicott. Protein kinase inhibitors: Insights into drug design from structure. *Science*, 2004. URL http://science.sciencemag.org/content/303/5665/1800.