Molecular Modelling – How Reliable are Commonly Used Force Fields for Conformer Searching?

4th year project report, submitted to the University of Manchester in partial fulfilment of the MPharm degree

John Brown
9467009

Supervisor Richard Bryce

11/05/2018

School of Health Sciences
Division of Pharmacy and Optometry
Name of supervisor

List of Contents

Declaration and Acknowledgement

I understand the nature of plagiarism and that it is serious academic offence. I confirm that no material in the project has been plagiarised

Date:10/05/2018

.

Firstly, I would like to give thanks to the post graduate students who provided help and support with my data collection. I would also like to give thanks to my project supervisor Richard Bryce, his guidance and support has greatly helped me over the past year and I would have deeply struggled without it.

Abstract

Molecular mechanical force fields offer a fast and efficient way of calculating molecular properties and are often used prior to costly quantum mechanical calculations for rapid conformer searching to produce an ensemble of the lowest energy conformers. Despite being commonly used, the reliability of force fields is questioned by growing literature with potentially negative implications. This study compares the performance of several commonly used force fields (MMFF94, MMFF94s, AMBER12: EHT, AMBER94, AMBER99, OPLS-AA, CHARMM27) with a higher level of quantum theory than previous works (M06-2X). Semi-empirical approaches boast a major reduction in computational cost with accurate results, however, their reliability has also been questioned. Therefore, as a secondary aim this study will assess the performance of a semi-empirical approach (PM6). This study finds several factors which influence the performance of force fields, however, the most significant factors appear to be the presence of so called 'problem structures' and the degree of conformational change between conformers. The force fields tested show instances whereby they perform reasonably well, agreeing almost perfectly with M06-2X reference data, however, also show instances of poor performance. The poor performance of the force fields was somewhat predictable based on the presence of these influencing factors; however, the semi-empirical approach performed much more erratically and produced the highest errors in the whole study and as such suggested to be avoided in conformer searching. This study concludes that force fields can be used reliably for conformer searching, however, an informed decision considering the factors influencing performance needs to be made when selecting a method.

Chapter 1: Introduction

1.1 Introduction to Molecular Modelling

Drug design is an expensive and convoluted process which employs a variety of methods, starting with the initial step of lead compound identification[1]. This compound may be an endogenous ligand or naturally obtained, e.g., opiate drugs from the opium poppy, *Papaver somniferum*; however, it could be a completely new compound[2]. Computational methods of drug design are now mainstay due to the recent advances in computer software and hardware, allowing even commercial computers to calculate complex calculations with minimal expertise[3,4].

Molecular modelling is often used in structure-based drug design, describing several methods used to build the three-dimensional structure of molecules, with the aim of predicting the physical and chemical behaviours of the molecule in question[3,5]. Models can be visual, e.g. ball and stick models, mathematical, e.g. equations describing a molecule, or computational, e.g. computer software employing algorithms and parameterisations to calculate the properties of a molecular system[4]. Additionally, the complexity of models also varies, for example, ranging from simple Lewis dot structures illustrating atoms and electrons, to greatly complex molecular systems described using high levels of quantum mechanics[4,5]. Computational methods of modelling can be further categorised as either molecular mechanics or quantum mechanics[4].

Molecular mechanical approaches treat molecules as atoms connected by springs and calculates potential energy as the total energy from bonds, bond angles, non-bonded interactions, torsion angles, charge-charge interactions and other cross terms; this potential energy is called the force field[4,6,7]. Quantum mechanical methods consider the interactions between nuclei and electrons and assumes an electron's movement is independent of other electrons and nuclei have no motion[4,8]. Methods based on quantum theory can be further categorised as either semi-empirical or *ab initio*[4,8].

Although quantum approaches provide highly accurate results they are costly and time consuming in terms of computational power, considering this as well as the fast and efficient nature of molecular mechanical calculations and their ability to describe systems with millions

of atoms, they are used for 'cheap' experiments and preliminary conformer analysis to quickly identify an ensemble of low energy conformers[4-6]. Despite their wide use, there are several limitations associated with molecular mechanics, for example, their inability to describe chemical reactions and limited use within their parametrisation, furthermore, their reliability has been brought into question by recent literature[6-9].

1.2 Molecular Mechanics

The reliability of molecular force fields is based on several assumptions and their associated validity, for example, the Borne-Oppenheimer approximation which, due to the fast and independent movement of electrons, calculates potential energy based on nuclear position[7,10]. Since force fields consider nuclear position as opposed to electrons, they are unable to describe electronic properties, more specifically chemical reactions, one of the major limitations of molecular mechanical force fields[7]. In addition, 'transferability' which translates as the parameters of a force field being derived from a limited number of molecule applied to a larger one, e.g. a larger molecule set containing the same functional groups[7]. The forces considered by force fields can be described as the bonded and non-bonded interactions and are calculated with energy functions and terms, as well as deviation from reference values[4,6,7].

1.2.1 Bonded Interactions

Bonded interactions are the strain energy of mechanical bonds within a molecule, one of these is bond stretching[6,7]. Bond stretching is the energy associated with elongation of a bond and can be calculated using the Morse potential, although it is seldom used due to computational cost, therefore, Hooke's Law is often used instead[6,7]. Hooke's Law encompasses that as a bond is pulled or pushed energy will increase, however, provides a harmonic U-shaped energy curve so that a continual energy increase does not result in bond breakage[6].

Another force is bond angles, in its simplest form it can be calculated by squaring the deviation from reference values, though, alike bond stretching, Hooke's Law is commonly used[6,7]. This method is sufficient for most systems as deviation is only likely to be small, however,

modification of this can increase accuracy in certain instances, e.g., when a molecule is vastly strained[6,7].

Rotation around bonds often leads to major conformational changes in flexible molecules, as highlighted by the different conformational phases of alkanes; as such, sufficient understanding of torsional angle energy is needed to implicate parametrisation for accurate calculations[6,7]. Although torsional potentials are not always used, for molecules with a high degree of torsional rotation a cosine series expansion calculation of all nonbonding quartets is carried out[7].

Another consideration is cross terms, these are other terms which calculate interactions that arise from two other terms, for example, bond stretching-bending or stretching-torsional[6,7].

1.2.2 Non-Bonded Interactions

Van der Waals interactions are an example of a non-bonding force that are considered in molecular mechanics, it describes the distance dependant repulsive and attractive forces that occur between molecules and atoms[6,7]. A quantum calculation of these forces would be costly; therefore, the Lennard-Jones potential function is often used[7]. Modifications of this function may also be used, e.g., the Hill and Buckingham potentials[7].

Molecular mechanical force fields were originally developed for hydrocarbons, as such charge-charge forces were of little significance, however, parametrisation now contain terms which account for partial charge, e.g., Coulomb's Law[6].

1.3 Commonly Used Force Fields

1.3.1 MMFF

The 'Merk Molecular Force Field' or MMFF is a group of force fields produced by Merk for a range of drug-like and bioactive molecules; the first MMFF developed, MMFF94, was derived from high-level quantum theory[11,12]. Halgren produced several papers investigating the two MMFF force fields, MMFF94 and MMFF94s, the latter being a 'static' version of MMFF94 which includes parameters for delocalised planar nitrogen atoms[12-18]. During this series, interaction energy, conformational energy and geometries calculated by the two MMFF force fields were

compared with data from several other force fields, e.g., CHARMM, OPLS and AMBER[12-18]. Both MMFF94 and MMFF94s consistently show accuracy comparable with varying levels of quantum derived geometries and conformational energies for a range of molecules[12-18]. Furthermore, these two force fields outperform AMBER, CHARMM27 and MM2 in reproducing intermolecular energies and amino acid energies[19,20]; and have been validated for us in ligand-docking software[21-23], estimating binding and strain energy[24-26], and with amino acids[20,26].

Despite the promising results produced by Halgren[12-18] several papers suggest the reliability of MMFF force fields may not be as good as previously thought, for example, when testing the ability of MMFF94 to reproduce amino acid conformational energies, bias towards specific conformations was found, with approximately 10 kj/mol average error[27]; furthermore, MMFF94 reproduced less than half the low energy conformations identified by density functional theory, B3LYP/6-31G[20]. Studies investigating strain energy, found MMFF94 produced significantly different geometries than those produced by HF/6-31G[26] and showed difficulty reproducing key drug fragment energies[25]. Furthermore, the poor performance of MMFF94 has been shown in research investigating conformational energy changes associated with ligand-protein binding[24], positioning accuracy of ligand docking calculations[22] and condense phase simulations of organic liquids[28]. Kanal et al.[9] questioned the reliability of several force fields, including MMFF94, ranking energies calculated by the force fields and comparing them to quantum derived data, concluding that force fields are inherently unreliable and should be used with caution for conformer searching.

1.3.2 AMBER

AMBER or 'Assisted Model Building with Energy Refinement' is a group of force fields commonly used for biomolecular simulations, the idea of AMBER first started in the late 70s, there are now several AMBER force fields[29,30]. There is a vast amount of literature testing AMBER force fields, for example, estimating intermolecular interactions[31,32], hydration energies[33,34], and for proteins and peptides[35-37]. Throughout current literature, AMBER force fields are shown to describe intermolecular interactions well, for example AMBER99SB accurately reproduces experimental binding affinities of oseltamivir and pandemic influenza

strains[31]. Furthermore, AMBER99 shows reasonable performance when reproducing interaction energies between amino acid analogues and carbon nanotubes when compared with density functional theory[32]. Furthermore, Mobley et al.[33] suggested if instances of force field inadequacies, e.g., difficulty describing hypervalent phosphorous and sulphur containing molecules, are taken into consideration then force fields such as AMBER can perform well in reproducing hydration free energies of drug-like molecules. Similarly, AMBER94 produced a root mean square deviation of 1.35 kcal/mol with free energy of amino acid side chains[34]. AMBER force fields also outperform other force fields, for example, AMBER94 outperformed MMFF94 when reproducing thermodynamic properties of organic liquids[28] and two AMBER force fields, AMBER14sb and AMBER99sb, showed superiority over OPLS-AA when reproducing experimental data for peptide backbones[35-37].

As with the two MMFF force fields, several papers question the reliability of AMBER force fields, for example, AMBER94 and AMBER99 performed poorly when reproducing conformational energies of amino acids[20, 27], nucleic acids[38], protein backbones[39], as well as small peptides[40-42]. Furthermore, Halgren[18] found AMBER force fields poorly described conformational energies of small organic molecules in comparison with two MMFF force fields. During analysis of Taxol conformers, conflicting and mixed results were produced by several force fields, including an AMBER force field[43]. There is little literature regarding AMBER12: EHT, although one study found this force field to reproduce key drug fragments poorly[25]. Despite the lack of literature, AMBER12: EHT is widely used for minimisation in several studies[44-47]. The most notable results of current literature regarding force fields is from Kaminský and Jensen's[27] study, whereby eight force fields tested, including AMBER94 and AMBER14sb, produced large errors of approximately 10 kj/mol.

1.3.3 CHARMM

'Chemistry at Harvard Macromolecular Mechanics' also known as CHARMM is group of force fields and a simulation programme developed by a Harvard university research group with the project lead Martin Karplus, developed for an array of biomolecules such as proteins, nucleic acids and lipids, as well as small 'drug-like' ligands[48,49]. CHARMM27 is an all-atom force field

and is the most recent CHARMM based on the previous CHARMM22 force field and quantum data[50-52]. Halgren[18] compared MMFF94 and MMFF94s with other force fields, including CHARMM22, finding CHARMM22 to perform similarly to MMFF94 when reproducing intermolecular forces, however, noted minor inaccuracies with cyclic alkanol energies. Furthermore, CHARMM22 has been shown to outperform both OPLS and GAFF force fields in reproducing energies for a range of organic molecular structures[53]. A significant improvement over the previous CHARMM22 force field was demonstrated by CHARMM27 when investigating the structural properties of nucleic acids[50,51]. Both CHARMM22 and CHARMM27 calculate intermolecular force with reasonable accuracy, for example, with oseltamivir-influenza binding energies[31] and amino acid-carbon nanotube interaction energies[32].

However, several papers find CHARMM force fields to perform poorly, for example, with amino acids conformers[20, 27], peptides[36] and during the development of a polarisable force field for polyols[54]. Kaminský and Jensen[20,27] found several force fields, including CHARMM27, are unreliable for calculating conformational energies and geometries of amino acids; furthermore, finding that CHARMM27 only performed slightly better than the worst performing force field[20]. Fluitt and Pablo[36] investigated polyglutamine peptides simulations, finding that CHARMM27 was outperformed, noting under and over stabilisation of beta-sheets and alpha-helices, respectively. Additionally, during the development of a drude oscillator force field, CHARMM force field struggled to reproduce quantum potential energy profiles and relative energies of polyols.

1.3.4 OPLS

'Optimised Potential for Liquid Simulations' or OPLS describes a group of force fields developed for simulations, OPLS-AA denotes the 'all-atom' version and OPLSA-UA the 'united-atom' version[55,56]. Generally, studies evaluating OPLS-AA all agree that it performs well[19,20,28,31,32,34-36,57,58], with only limited research suggesting otherwise[27,37]. When applied to amino acids, OPLS-AA outperforms AMBER and CHARMM in reproducing geometries and relative energies[20], as well as giving a root mean square deviation of 0.85kcal/mol[34]. Furthermore, OPLS-AA outperformed other force fields in describing polyglutamine aggregation and folding in peptide simulations[36], showing reasonable agreement with experimental data of proline and glycine

peptides[35]. Paton and Goodman[19] evaluated several force fields in their ability to reproduce stabilisation energies and geometries of biochemical complexes, finding OPLS-AA outperformed AMBER, MM2, and MM3, agreeing reasonably with high level quantum data. When predicting thermodynamic properties, e.g., density and vaporisation energies, OPLS-AA is the most accurate compared with CHARMM, GAFF, and MMFF94[28,57]. Furthermore, OPLS-AA produced errors of under 1.6% and was deemed appropriate for condense phase simulations[28]. Both Nyugen[31] and Yang et al.[32] agree that OPLS-AA can calculate intermolecular interactions to a degree comparable with experimental and quantum data, the former investigating oseltamivir binding affinities[31] and the latter testing amino acid-carbon nanotube interaction energies[32]. Lennox et al.[58] conducted a force field validation study for 'organic-metal framework', measuring the adsorption of methane with several force fields, finding OPLS-AA to agree reasonably with density functional theory.

There is only limited research showing OPLS-AA inaccuracy, for example, when investigating conformational energies of peptide backbones, OPLS-AA produced major inconsistencies[37]. Furthermore, Kaminský and Jensen[27] found that several force fields, including OPLS-AA, produced errors of approximately 10kcal/mol when investigating conformational energies of amino acids.

1.4 Introduction to Quantum Mechanics

Quantum mechanical methods of molecular modelling directly consider the electrons in a system and are based on quantum physics[8,10]. Since electrons are directly considered, electronic properties can be calculated, more specifically, chemical reactions which molecular mechanics cannot determine[8,10]. *Ab initio* is a Latin term meaning 'from the beginning'[59] and is an approach to quantum mechanics that is based on the Schrödinger equation, this mathematical formula describes the movement of matter over time through a given space, calculating this equation allows one to describe molecular properties and thus preamble to quantum calculations[8,10]. Calculating the exact quantum nature of even the smallest molecules would require a massive amount of computational power and considering that many of the

systems calculated with quantum mechanics are complex polyatomic structures, several approximations are made[8,10], this gives rise to different methods of quantum mechanics

1.4.1 Semi-Empirical Methods

One of the most time-consuming steps in an Hartree-Fock *ab initio* calculation is computing integrals, therefore, semi-empirical approaches focus on reducing this by making approximations with electrons, i.e., by only considering valence electrons[8,10]. Complete neglect of differential overlap or 'CNDO' is one of the first semi-empirical methods, developed by Pople et al.[60] CNDO has *ab initio* derived parameters[8,10]. There are now many different semi-empirical approaches, for example, PM6[61] and PM7[62] which are derived from modification of NNDO or 'neglect of diatomic differential overlap'[63].

1.4.2 Density Functional Theory

Density functional theory is an approach to quantum mechanics which uses functionals to calculate electron density and distribution, hence the name 'density functional theory', this approach considers electron density as opposed to molecular orbitals[10,64].

B3LYP is a hybrid functional in density functional theory originally used to investigate dichroism and vibrational absorption[65]. B3LYP is now a widely used alternative to costly Hartree-Fock calculations and has be shown to produce accurate results with reduced computational cost[66], for example, for predicting thermochemical properties[67], redox potentials[68] as well as gas-phase acid and base predictions[69]. Furthermore, this hybrid functional is used as a benchmark for comparative studies, e.g. for force field validation[9,20,38].

M06 represents a group of functionals used in density functional theory, these include M06-L, a 'local' functional with no HF-exchange[70], M06-HF with complete HF-exchange[71], and the hybrid functionals M06 and M06-2X with the latter having double the HF-exchange of the former[72]. M06-2X has been validated by several papers[73-76], for example, in hydrogen-bond interactions[73], isomerisation reactions[74], reaction energy of hydrocarbons[75] and conjugation reactions[76].

Although B3LYP has been shown to produce results with accuracy comparable to higher levels of quantum theory[65-69] no method is without its limitations and several papers have shown B3LYP to be outperformed by M06-2X[73-76] and M05-2X[77-79].

1.4.3 Semi-Empirical versus Density Functional Theory

Semi empirical approaches can be calculated at a fraction of the computational cost of higher theory methods, such as density functional theory, with reasonable level of accuracy[8,10], however, this may this reduction in computer time could potentially be at the expense of accuracy. When comparing quantum and molecular mechanical methods in their ability to reproduce strain energies, Sellers et al.[25] found two semi-empirical methods, PM3 and AM1, had even worse accuracy than the force fields they tested when compared to *ab initio* calculations. Furthermore, PM6 has been shown to calculate significantly different geometries when compared with B3LYP for fullerenes[80], zinc and magnesium containing compounds[81], as well as palladium-ligand complexes[82]. Furthermore, the semi-empirical methods PM6, PM3 and RM1 produced significant errors in heat of cyclisation when compared with B3LYP[83]. Similar errors are seen with electronic and thermochemical properties of palladium-ligand complexes[82], ligand proton and dissociation energies[81], as well as ionisation potentials and electron affinities[80].

Although semi-empirical approaches may not be as robust as density functional theory, studies have shown they have better accuracy than several force fields[9,84]. Kanal and Hutchison[9] found that although PM7 had noticeable discrepancies, e.g., poor correlation with B3LYP optimised geometries, it was more reliable than the force fields they tested, concluding that it could be used for conformer searching with reasonable accuracy. Li et al.[84] compared multiple quantum methods in their ability to calculate non-covalent interactions, finding that PM6 and PM7 produced less accurate results, however noted that these calculations were substantially faster than density functional theory methods. These studies suggest that although the accuracy of semi-empirical methods may not be as good as higher theory methods, the benefit gained from

reduced computational cost outweighs the loss of accuracy. Furthermore, the PMx family are commonly used in studies as a benchmark for comparison[21-23].

1.5 Aims and Objectives

The reliability of force fields for conformer searching is often questioned throughout current literature, as such, the main aim of this project is to assess the use of commonly used force fields as applied to conformer searching and to identify any potential problems associated with their use. As a secondary aim this study will assess the performance of two lower levels of quantum theory, including a semi-empirical approach as well as a density functional theory method. This will be achieved by measuring and comparing the error of each method relative to quantum theory data. Furthermore, conformational analysis of conformers in the sample as well as linear regression analysis will highlight structures which may be problematic to force fields. Current literature regarding force field performance often uses lower level quantum methods such as density functional theory approach, B3LYP; therefore, this study will use a higher level of theory, M06-2X. Furthermore, a larger range of force fields than previous literature will be used.

Chapter 2: Methods and Materials

2.1 Molecule and Conformer Selection

Kanal and Hutchison[9] identified thirteen molecules with significantly poor R^2 values labelling them 'problem molecules', of these thirteen molecules 10 were selected for analysis. The conformers in Kanal and Hutchison's study[9] were generated by Ebejer et al.[85] however based on previous work by[86,87]. Using the simplified molecular input line entry specification, also known as 'SMILES'[88], provided in Kanal and Hutchison's[9] supporting information the 10 problem molecules were identified and extracted from a TAR file, also obtained from the supporting information file. For each of the 10 problem molecules their corresponding conformers present in the TAR file were used, each molecule had 10 conformers. The problem molecules used in the present work can be seen in figure 1 below

2.2 Computational Details

Open Babel[89] was used to convert each of the conformer files from MDL molfile format (.mol) to a Gaussian input file (.gjf). All single point energy calculations for the quantum mechanical methods used, i.e., M06-2X[72], B3LYP[65,90,91] and PM6[61], were carried out using Gaussian version 09[92]. The key word corresponding to the method being tested was entered into the bash script and the multiplicity set to 1. The charge was set corresponding to the overall net charge of the molecule in question, for example, a neutral molecule would be set to 0 and a molecule with a net charge of -1 would be set to -1, etc. For the two density functional theory methods, i.e., M06-2X and B3LYP, the basis set DEF2-SVP[93,94] was used, this was done by placing the key word 'DEF2SVP' adjacent to the method keyword in the bash script. Potential energy calculations for the seven molecular mechanical methods, i.e., MMFF94[12-18], MMFF94s[12-18], AMBER94[95], AMBER99[96], AMBER12:EHT[97], CHARMM27[50-52] and OPLS-AA[55,56], were carried out using MOE version 2016.08[98]. The default parameters in MOE[98] were used for each method and dielectric constant epsilon set to 1.

2.3 Data Analysis

All quantum mechanical data obtained from Gaussian[92] and force field data from MOE[98] was put into an excel spreadsheet. Quantum single point energies were converted from Hartree (HF) to kcal/mol using the conversion factor 627.5095. With the quantum mechanical derived data now converted, all the data was ranked lowest to highest and relative energy calculated by subtracting each value from the lowest energy value as identified by the corresponding method. Absolute energy was calculated with reference to M06-2X data. M06-2X data was used as reference data as it is the method with the highest level of quantum theory used in this study; furthermore, it has been shown to outperform other density functional theory approaches, for example, B3LYP, by several studies[73-76]. Furthermore, ranges in both relative and absolute energy are present, this is the difference in energy between the lowest and highest energy conformer as identified by that method. Ranges in relative and absolute energy as well as R^2 values can be seen in appendix D, furthermore, ranking of relative an absolute energy can be seen in appendix B and C.

Root mean square error (RMSE) was used to quantify the 'error' of a method relative to M06-2X reference data, this was performed in excel for each molecule by calculating the mean squared difference between a methods energy an corresponding M06-2X data, followed by a square root function. Using both the absolute and relative energies, XY graphs were plotted and the R^2 values as well as trend lines were calculated. Conformer analysis was performed using MOE[98] and chemical structures drawn using both ChemDraw 14 and eMolecules[99]. Structures of the molecules used in the sample can be seen in figure 1 and conformer structures can be seen in appendix A. Furthermore, the average root mean square error was calculated by averaging the sum of a methods error over the 10 molecules and the resulting value ranked to find the best and worst performing methods. Root mean square error produced by a method for each molecule can be seen in appendix D.

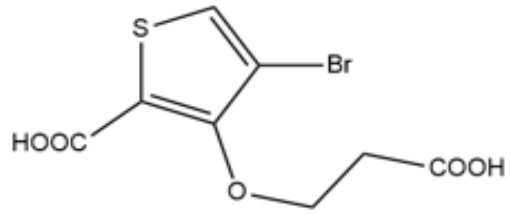
2.4 Justification of Methods

Current literature evaluating the performance of force fields commonly uses quantum approaches such as B3YLYP, like in Kanal and Hutchison's[9] study, therefore a higher level of quantum theory was used in this study, i.e., M06-2X. Furthermore, ten molecules taken from this paper[9] were used as they caused a problem for three commonly used force fields, i.e. GAFF, UFF and MMFF94, and I wanted to investigate this finding further. As well as including the MMFF94 force field used in Kanal and Hutchison's[9] paper, six other commonly used force fields were tested so that a wider comparison between molecular mechanical methods could be made. A lower level of quantum theory, B3LYP, that was used in Kanal and Hutchison's[9] study was also tested; this was included to make comparison of error between the two varying levels of quantum theory. Furthermore, several studies[73-76] have shown M06-2X to outperform B3LYP, as such; I wanted to explore this finding further. The seven force fields included in this study were tested as they are some of the most commonly used throughout molecular modelling literature; furthermore, testing their performance would allow for a greater comparison with current literature. PM6 was tested in this study to allow for comparison between a semi-empirical approach and molecular mechanics.

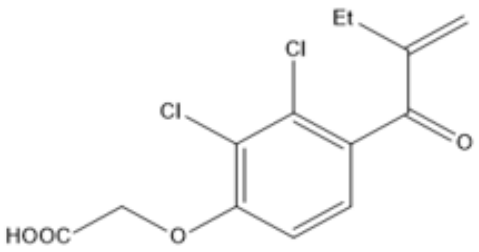
'Molecular Operating Environment' also known as MOE is a computer program which allows one to visualise molecules in three dimensions, it is widely used for computer based drug design as it allows for a number of functions to be performed, for example, calculating potential energy, as well as performing docking and dynamic simulations[98]. The MOE software was used in this study as it was readily available; furthermore, it is used by many studies, for example, conformational sampling studies[107], transition metal analysis[108,109], as well as quantitative structure activity relationship analysis studies[110].

Gaussian [92] refers to computer software which is used to calculate a variety of molecular properties such as single point energies; furthermore, several quantum and molecular mechanical methods can be used with the software. Gaussian version 09[92] was used in this study as it was readily available to me and is commonly used in computational drug design as well as being a method included in a vast amount of literature[20,26,27,31,40,54].

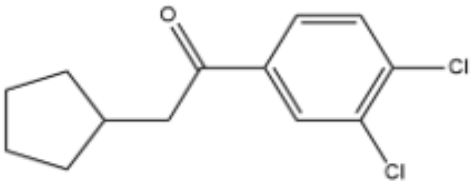
Quantum single point energies were converted from Hartree to kcal/mol to allow for better handling of the data. Furthermore, relative and absolute energies allow for comparison between different methods and their ranges tell us the difference between the lowest and highest energy conformer. Root mean square error was used as it makes it possible to quantify error and allows for easy comparison, also averaging these results over the 10 molecules and ranking this is useful as it shows us which methods did the best and worst. R^2 values are commonly used in statistical analysis and describe the relationship between two variables, in this case the relationship between relative and absolute energy, an R^2 value of 1 would be a perfect correlation between the two variables and conversely a value of 0 would mean no correlation.



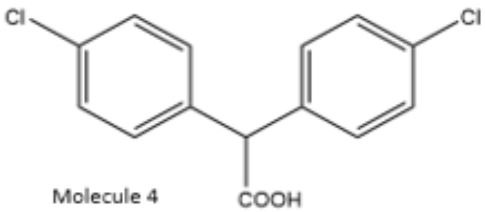
Molecule 1



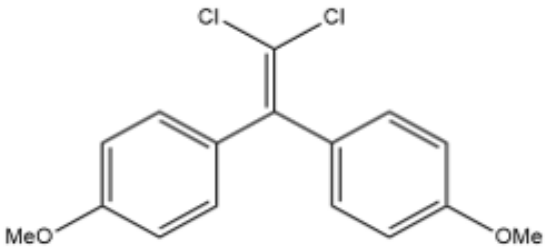
Molecule 2



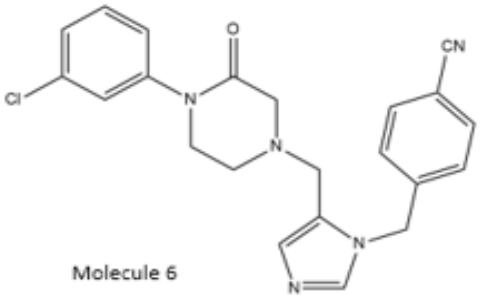
Molecule 3



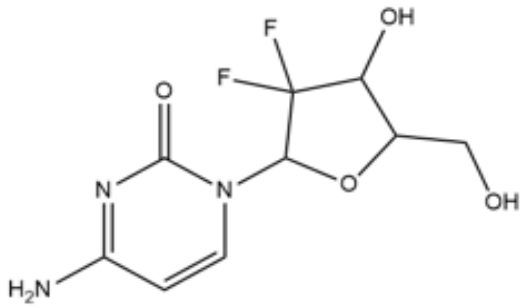
Molecule 4



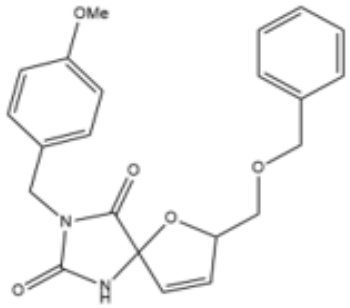
Molecule 5



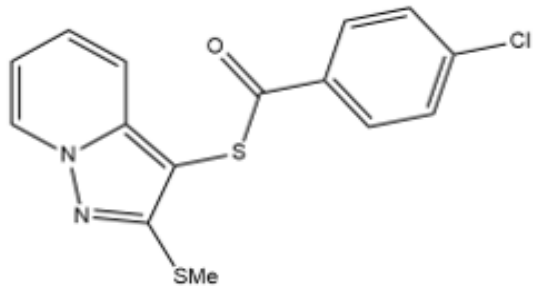
Molecule 6



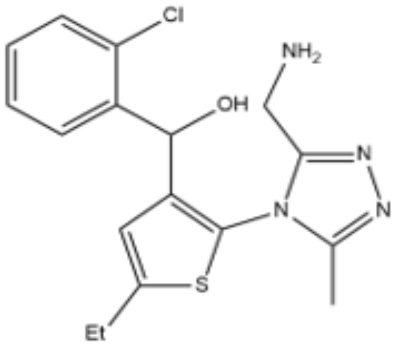
Molecule 7



Molecule 8



Molecule 9



Molecule 10

Figure 1: Sample of Ten Problem Molecules

Chapter 3: Results

3.1 Molecular Mechanical Force Fields

3.1.1 Average Root Mean Square Error

Through ranking the average root mean square error as seen in table 1 and 2, OPLS-AA ranks the best of the force fields giving the lowest average RMSE of 4.25 kcal/mol; conversely, MMFF94s ranks last giving the highest average RMSE of 5.81 kcal/mol. Although the ranking of average root mean square error shows the best and worst force field on average, it is important to note there is only a difference of 1.56 kcal/mol between the lowest and highest value.

Of the three AMBER force fields, AMBER12: EHT performs the best giving the lowest average of 4.51 kcal/mol and ranking second out of the seven force fields. The next best performing AMBER force field, AMBER94 produces an average RMSE of 5.08 kcal/mol, ranking fifth, and the worst performing AMBER force field AMBERR99 ranks sixth with an average RMSE of 5.22 kcal/mol.

MMFF94 performs the best out of the two MMFF force fields tested, producing an average RMSE of 4.81 kcal/mol and ranking fourth. However, MMFF94s performs the worst of all the force fields in terms of average RMSE, producing the highest value of 5.81 kcal/mol and ranking last.

The remaining force field, CHARMM27 performs in a similar fashion to AMBER12: EHT giving an average RMSE of 4.68 kcal/mol, raking third.

Rank	Method	Average RMSE (kcal/mol)
1	B3LYP	3.89
2	OPLS-AA	4.25
3	AMBER12: EHT	4.51
4	CHARMM27	4.68
5	MMFF94	4.81
6	AMBER94	5.08
7	AMBER99	5.22
8	MMFF94s	5.81
9	PM6	10.89

Rank	Method	Average RMSE (kcal/mol)
1	OPLS-AA	4.25
2	AMBER12: EHT	4.51
3	CHARMM27	4.68
4	MMFF94	4.81
5	AMBER94	5.08
6	AMBER99	5.22
7	MMFF94s	5.81

Table 1 (left): Ranking of all Computational Methods by Average Root Mean Square Error (RMSE, in kcal/mol) Over 10 Molecules

Table 2 (right): Ranking of Molecular Mechanical Methods by Average Root Mean Square Error (RMSE, in kcal/mol) Over 10 Molecules

3.1.2 Problem Molecules

Through analysis of individual RMSE values and ranges in relative and absolute energy, two problem molecules have been identified, molecules 2 and 6. These molecules gave RMSE values significantly higher than other molecules and seemed to pose a problem for all the computational methods tested, including the density functional theory approach B3LYP. Individual RMSE values and ranges in relative and absolute energies can be seen in appendix B, C and D.

By considering the large values and range in RMSE produced for molecule 2, it is apparent that this molecule proved challenging for the force fields. The best performing force field for this molecule, AMBER94, gave the lowest RMSE of 17.93 kcal/mol and the worst, both MMFF force fields, gave the highest RMSE value of 24.36 kcal/mol. Although AMBER94 performed the best,

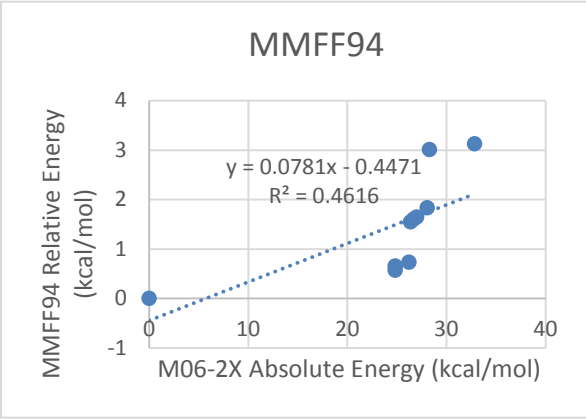
it is important to note that all force fields gave a RMSE within a 6.43 kcal/mol range. The poor performance of the force fields with molecule 2 is further emphasised by the large ranges in relative and absolute energy, for example, AMBER94 gave a large range in relative energy of 11.64 kcal/mol and the two MMFF force fields gave the largest range in absolute energy of 29.75 kcal/mol.

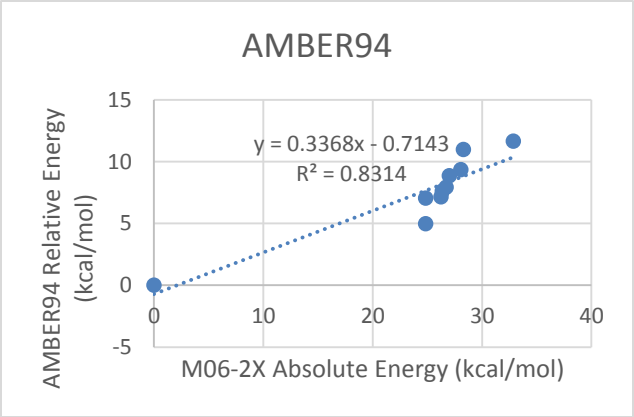
The seven force fields also seemed to show difficulty with molecule 6, however to a lesser extent than molecule 2 as the lowest and highest RMSE values for molecule 2 were 10.6 kcal/mol and 5.02 kcal/mol higher than molecule 6, respectively. The best performing force field for molecule 6, CHARMM27, gave the lowest RMSE value of 7.33 kcal/mol and the worst performing force field, MMFF94s, gave the largest RMSE of 19.34 kcal/mol. Although the RMSE values produced for molecule 2 were significantly higher than molecule 6, it is important to note that the ranges between best and worst force field is somewhat higher for molecule 6, i.e. a range of 12.01 kcal/mol between the best and worst RMSE for molecule 6 compared to a range of 6.43 kcal/mol for molecule 2. This larger range seen in molecule 6 is due to MMFF94s which performed significantly worse than the other force fields, for example, producing a RMSE 7.79 kcal/mol higher than the next worst performing force field, OPLS-AA.

Further analysis of the two problem molecules 2 and 6 through linear regression and conformer analysis has identified some individual conformers which may explain such high RMSE values. Calculated R^2 values can be seen in table 6 and conformer structures can be seen in appendix.

Although all seven force fields struggled with several conformers of molecule 2, it seems conformers 1 and 2 posed the most trouble. Furthermore, conformer 3 was misrepresented by the two MMFF force fields as well as CHARMM27 and OPLS-AA. Despite performing the best for molecule 2, AMBER94 incorrectly identified the lowest energy conformer, i.e. suggested conformer 4 had the lowest energy however conformer 2 was identified as having the lowest energy by M06-2X. Furthermore, all the force fields except AMBER99 also incorrectly identified the wrong lowest energy conformer. The poor performance of the seven force fields with molecule 2 is further highlighted by the R^2 values, with the two MMFF force fields giving the lowest values (R^2=0.46) and AMBER94 giving the highest (R^2=0.83). The low R^2 values given by

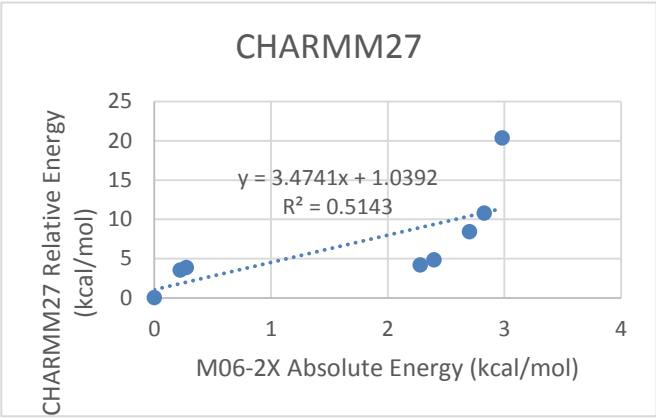
the MMFF force fields suggests a poor correlation between relative and absolute energy, this can be seen in the graphs below. Furthermore, it is important to note that R^2 values produced by MMFF94 and MMFF94s are some of the lowest in the whole study.

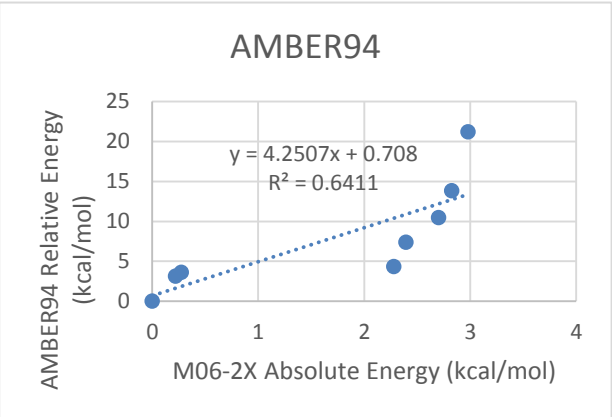




Graph 1 and 2: Relationship between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 2

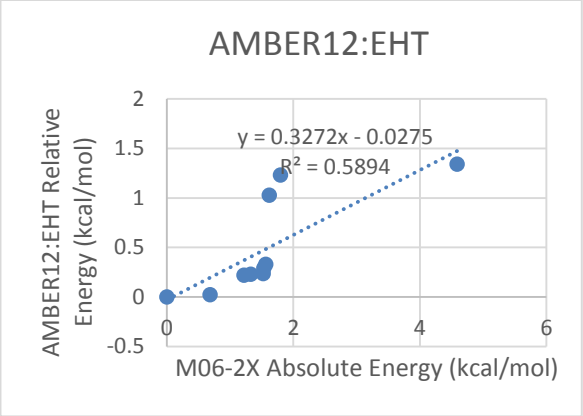
Several conformers of molecule 6 caused a problem for the force fields, however, conformer 7 seemed to be largely misrepresented by most of force fields, as well as conformer 8 which AMBER94, the two MMFF force fields as well as CHARMM27 and OPLS-AA struggled with. The best performing force field for molecule 6, CHARMM27, agreed with M06-2X that conformer 7 is the lowest energy conformer, however failed to correctly identify conformer 9 as the highest energy conformer. Despite producing the lowest RMSE for molecule 6, CHARMM27 had the lowest R^2 value of 0.51 and AMBER94 had the highest value, 0.64. The relationship between relative and absolute energy can be seen in the graphs below.

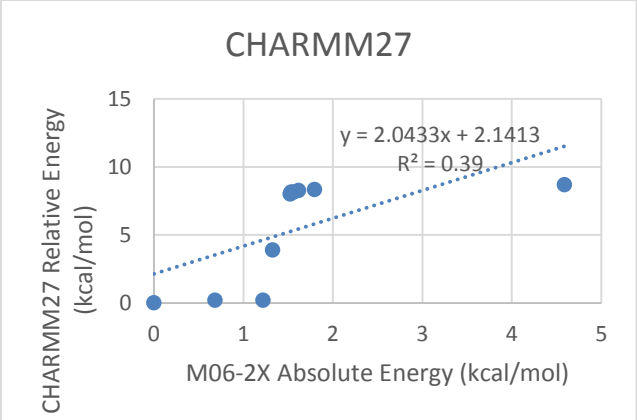




Graph 3 and 4: Relationship Between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 6

At first glance, molecule 10 may not seem to be a problem molecule considering that the RMSE values are not as large as with molecules 2 and 6, however analysis of R^2 values and conformers may suggest otherwise. CHARMM27 produced the lowest R^2 value of 0.39 and AMBER12: EHT produced the highest value, 0.59. It is important to note that these values are some of the lowest R^2 values in the whole study, the graphs showing this can be seen below. Further analysis of molecule 10 data suggests that although several conformers were misrepresented, conformer 11 posed a significant problem for all force fields except AMBER12: EHT, as well as conformer 12 which the AMBER force fields as well as CHARMM27 and OPLS-AA struggled with. Furthermore, the two MMFF force fields and AMBER12: EHT had difficulty with conformer 13. Surprisingly, the two MMFF force fields identified conformer 14 as having the lowest energy where as M06-2X calculated it as the highest energy conformer.





Graph 5 and 6: Relationship Between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 10

3.1.3 Other Molecules

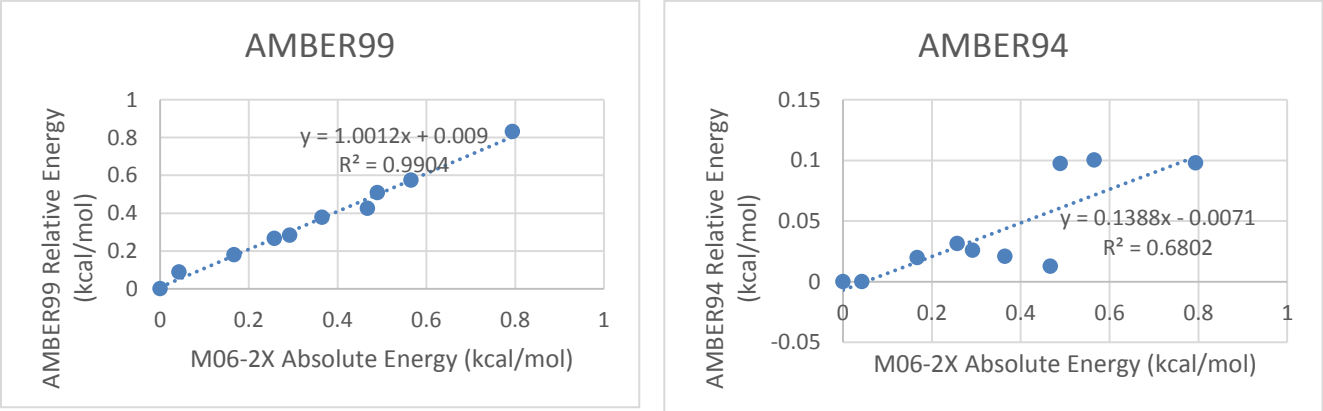
The seven force fields showed the most apparent ease with molecules 4 and 5 when considering the low RMSE values calculated for them.

AMBER99 performed the best of the force fields for molecule 4 giving the lowest RMSE of 0.02 kcal/mol and OPLS-AA giving the highest value of 0.61 kcal/mol. Furthermore, it is important to note these RMSE values are some of the lowest in the whole study. Both AMBER99 and OPLS-AA produced rather small ranges in relative and absolute energy when in comparison with other molecules, for example, AMBER99 produced a relative and absolute energy range of 0.83 kcal/mol and 0.05 kcal/mol for molecule 4, respectively, however as a comparative example the same force field produced significantly larger ranges for molecule 6. Although the best and worst force field have been identified based on RMSE, it is also important to point out that all the force fields performed within a range of 0.59 kcal/mol which is significantly lower than with other molecules.

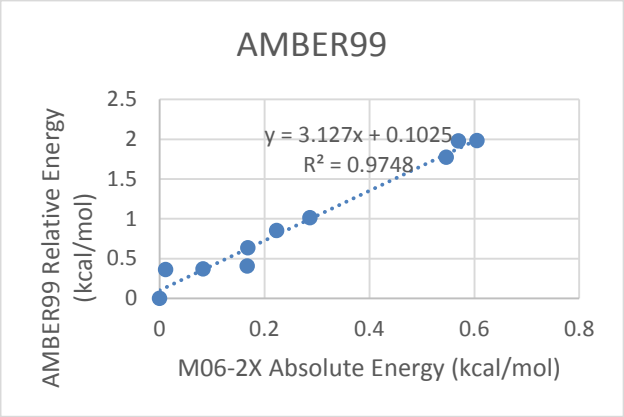
CHARMM27 performed the best of the seven force fields for molecule 5 producing the lowest RMSE value of 0.12 kcal/mol and AMBER99 performed the worst giving the largest RMSE value of 0.82 kcal/mol. CHARMM27 also produced the lowest range in absolute energy, 0.0.21kcal/mol, and gave the second lowest range in relative energy of 0.45 kcal/mol. AMBER99

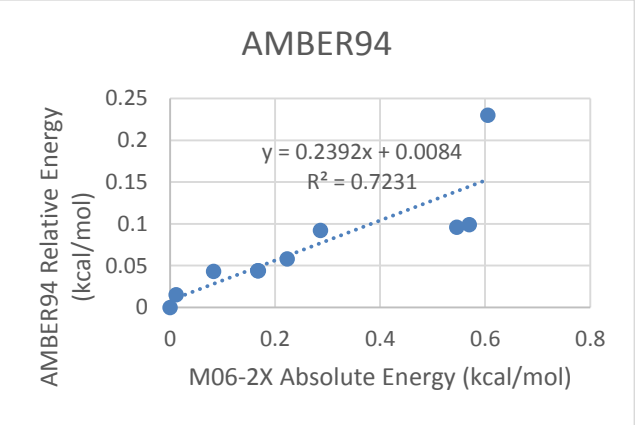
gave the highest ranges in both relative and absolute energy, however, it is should be noted that these ranges a significantly smaller than with other molecules.

Although the force field error calculated for molecules 4 and 5 is relatively low in comparison with the problem molecules, further investigation of individual conformers and linear regression analysis has identified a few instances of minor inadequacies. For example, conformer 15 of molecule 4 and conformer 16 of molecule 5 seemed to be slightly misrepresented and may have led to a 'dip' in R^2 values, especially with AMBER94. For example, the highest and lowest R^2 values for both molecule 4 and 5 were produced by AMBER99 and AMBER94, respectively. Although the best and worst values have been identified, these R^2 values can be considered relatively high and describe a positive correlation between relative and absolute energy. Furthermore, only three of the force fields, MMFF94, CHARMM27 and OPLS-AA, were able to correctly identify the lowest energy conformer of molecule 4.



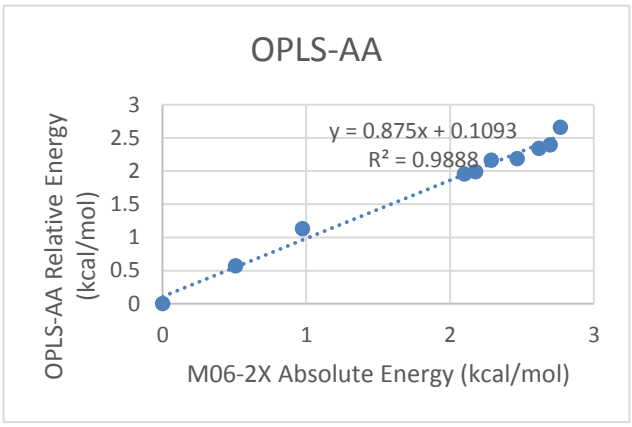
Graph 7 and 8: Relationship Between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 4

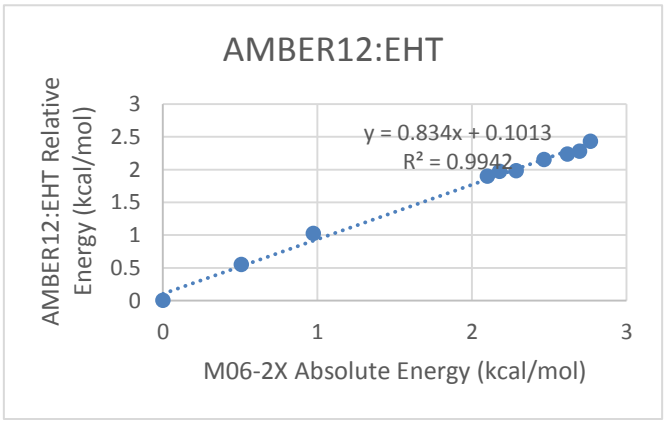




Graph 9 and 10: Relationship Between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 5

Although the force fields did not produce the lowest RMSE values or energy ranges for molecule 9, upon further investigation it is apparent that certain conformers proved seemingly easy for the force fields. Energies for conformers 17, 18, 19, 20 and 21 were in a reasonable agreement with M06-2X with many if not all the force fields, this is further emphasised by the R^2 values, i.e. all the force fields produced values of 0.98 or above suggesting a near perfect correlation between relative and absolute energy.



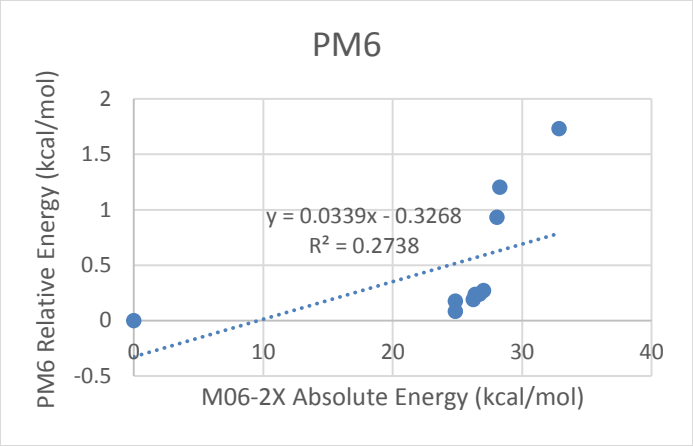


Graph 11 and 12: Relationship Between Force Field Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 9

3.2 Semi-Empirical Approach

The semi-empirical approach, PM6, performed the worst of all the methods based on average RMSE, giving the highest value of 10.89 kcal/mol. This average RMSE is 5.08 kcal higher than the worst performing force field and 6.64 kcal/mol higher than the best. This significantly higher average RMSE is more than likely due to poor performance of PM6 with molecules 2 and 7, whereby the semi-empirical approach produced massive RMSE values of 25.38 kcal/mol and 76.88 kcal/mol, respectively. These values are particularly significant as they are the two highest RMSE values in the whole study; to put it simply, the error produced by PM6 for molecule 7 is over eleven times larger than the worst performing force field for this molecule. This is further emphasised by the massive ranges in relative and absolute energy produced by PM6 for this molecule, i.e. ranges of 114.82 kcal/mol and 101.60 kcal/mol, respectively. As with the two RMSE values for molecules 2 and 7, these two energy ranges for are the highest in the whole study.

Although it is apparent that PM6 struggled with many of the conformers of molecule 2, further analysis has identified that the semi-empirical approach particularly struggled with conformers 5 and 1, but more so conformer 6, which seemingly provided a massive hurdle for PM6. Furthermore, PM6 produced the lowest R^2 values of all the methods for molecule 2 and can be seen in the graph below. As well as producing significantly large RMSE values, PM6 also failed to correctly identify the lowest and highest energy conformer. When looking at all the molecules, PM6 only correctly identified the lowest energy conformer for six of the ten molecules failing with molecule 2, 4, 6 and 9. Furthermore, PM6 only correctly identified the highest energy conformer for three of the ten molecules, 1, 6 and 7.



Graph 13: Relationship Between PM6 Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 2

When analysing the error produced by PM6 for each molecule it seems there is only three instances were PM6 performed better than all the seven force fields, however, for two of these instances PM6 was out performed by B3LYP, i.e. molecules 1 and 10. Apart from these instances, PM6 was outperformed by both B3LYP and at least one force field, with three instances of PM6 being outperformed by three or more force fields, i.e. molecules 3, 4 and 9.

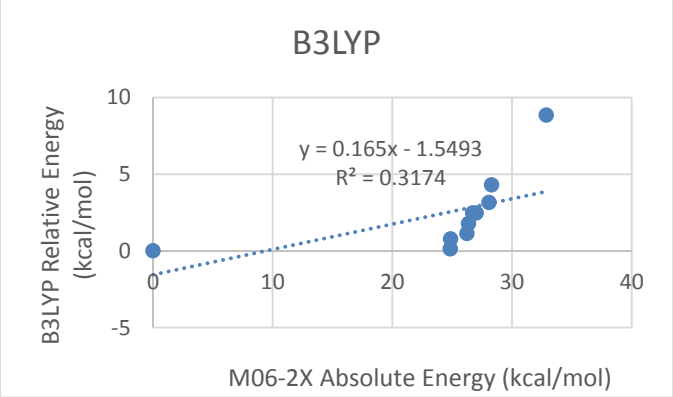
3.3 Density Functional Theory

Not surprisingly, B3LYP performed the best of all the methods producing the lowest average RMSE of 3.89 kcal/mol. Furthermore, B3LYP performed particularly well with molecules 1, 3-5, and 8-10, whereby it produced an RMSE value equal to or less than 0.5 kcal/mol, including the lowest reported RMSE value of 0.09 kcal/mol for molecule 5. B3LYP generally performs well, however, much like the force fields this approach also fell prey molecules 2 and 6, whereby it produced significantly higher errors of 23.23 kcal/mol and 11.34 kcal/mol, respectively. When factoring out the data for these two problem molecules we see a vast improvement in B3LYP's average RMSE, i.e. a reduction of 3.35 kcal/mol to an average value of 0.54 kcal/mol.

When looking at all the molecules in the test selection, B3LYP outperforms every other method for five of the ten molecules, as well as ranking within the best three methods for two of the remaining molecules and within sixth/seventh for the remaining three molecules. However, it is

important to note that the instances where B3LYP was out performed by many other methods referred to the problem molecules 2 and 6, whereby most of the methods performed as poorly as each other. Furthermore, the density functional theory approach correctly identified both the lowest and highest energy conformer for eight of the ten molecules.

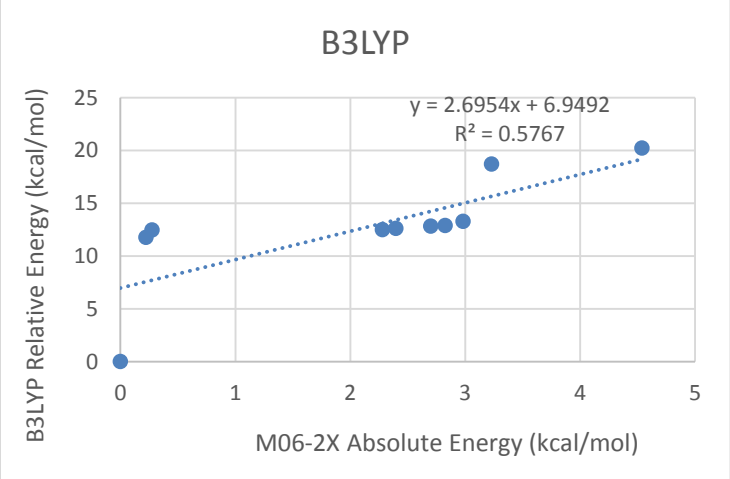
Further analysis of molecule 2 has identified individual conformers which B3LYP may have struggled with, for example, in a similar fashion to the force fields B3LYP showed difficulty in producing energies close to M06-2X for conformer 1. However, conformer 4 also seemed to pose a problem for B3LYP but did not seem to overtly cause the same issue for many of the force fields. Furthermore, B3LYP produced quite high ranges in relative and absolute energy for molecule 2, i.e. 8.85 kcal/mol and 25.12 kcal/mol, respectively. The significantly low R^2 values produced by B3LYP for molecule 2 serve to highlight the extent of the problem this molecule caused, as seen in the graph below. The poor results produced by B3LYP for molecule 2 seemingly paint this method in a negative light, as such, it is important to note that B3LYP did correctly identify the lowest and highest energy conformer for this molecule and many other methods failed this.



Graph 14: Relationship Between B3LYP Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 2

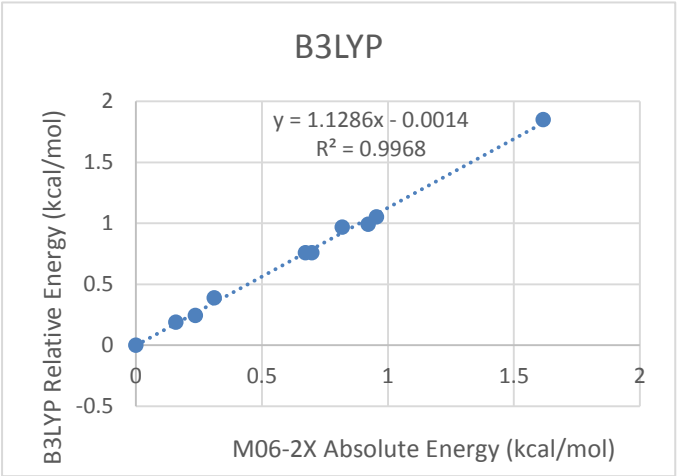
Molecule 6 posed much less of a problem to B3LYP when looking at the lower RMSE values produced, however, it is still significantly larger than other molecules and large ranges in relative and absolute energies were also seen. Further analysis of molecule 6 suggests that

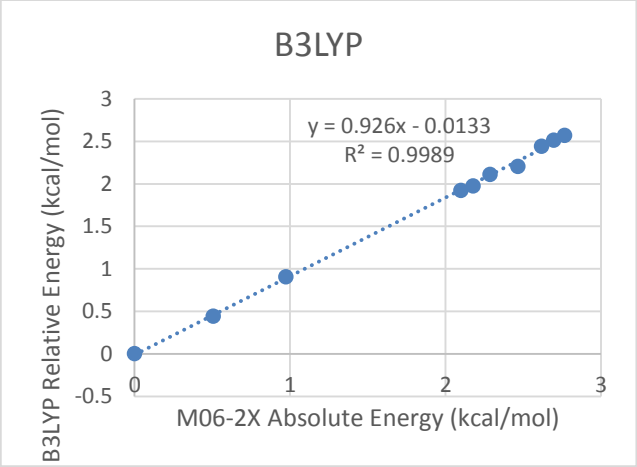
although several conformers may have contributed to the raised RMSE value, it was in fact conformers 7 and 10 which posed the most difficult for B3LYP to handle. Comparing the R^2 values of molecule 2 and 6 suggests that the latter may not have caused as much of a problem, i.e. a larger R^2 value with molecule 6, however, these values are still somewhat lower than with other molecules and describe a poorer correlation between relative and absolute energy, as seen in the graph below.



Graph 15: Relationship Between B3LYP Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 6

As mentioned before, a lot of this data when looked at individually may be seemingly negative and suggest B3LYP to be inaccurate, therefore, it is important to consider that there are several instances where B3LYP performs significantly better than other methods, as well as cases where B3LYP relative energy is extremely close with M06-2X absolute energy producing high R^2 values, as can be seen in the graphs below.





Graph 16 and 17: Relationship Between B3LYP Relative Energy and M06-2X Absolute Energy (kcal/mol) for Molecule 1 (Left) and Molecule 9 (Right)

Chapter 4: Discussion

4.1 Problem Molecules

4.1.1 Molecule 2

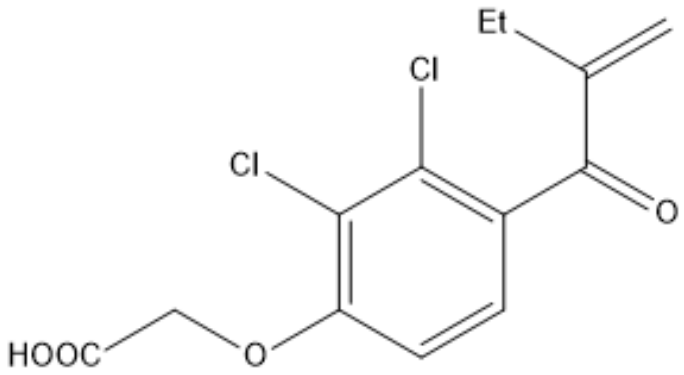


Figure 2: Chemical Structure of Molecule 2

There are no overt chemical differences which set molecule 2 apart from the others; however, analysis of its conformers highlights some potential problems which may have resulted in the high RMSE values and energy ranges; molecule 2 can be seen in figure 2. Comparison of the lowest and highest energy conformers suggests that this problem may have been caused by a closer interaction between aromatic chlorine atoms and carbonyl group as well as the terminal

carbon of the alkyl chain, as seen in figure 3 below. Furthermore, the highest energy conformer appears to be subject to allylic strain, whereby the rigidity of the SP^2 hybridised carbon-carbon double bond causes substituent groups to interact closer occupying a less favourable position, this can be seen in figure 4 and 5.

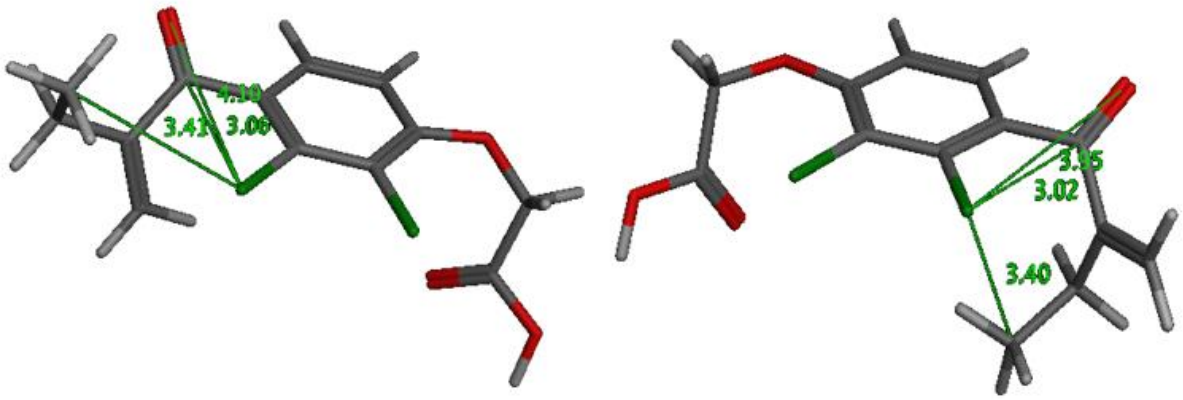


Figure 3: Showing Closer Interaction Between Aromatic Chlorine and Substituent Groups, from the Lowest Energy Conformer on the Left to the Highest on the Right.

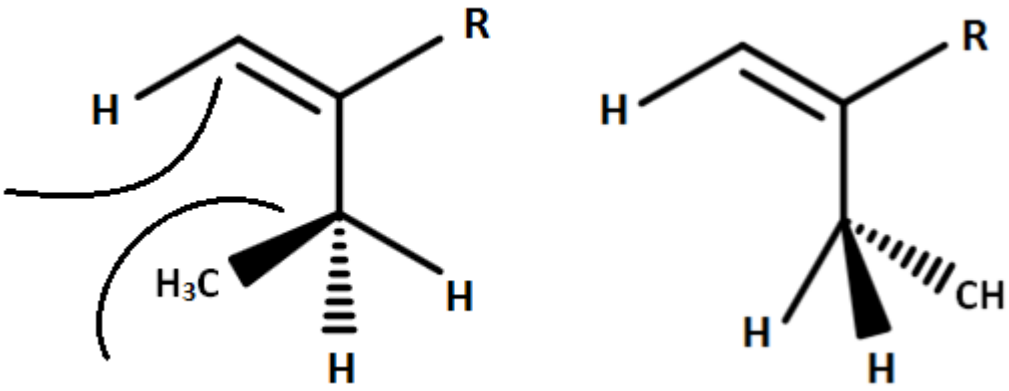


Figure 4: Showing Allylic Strain in Highest Energy Conformer (Left) Compared to the Lowest Energy Conformer (Right)

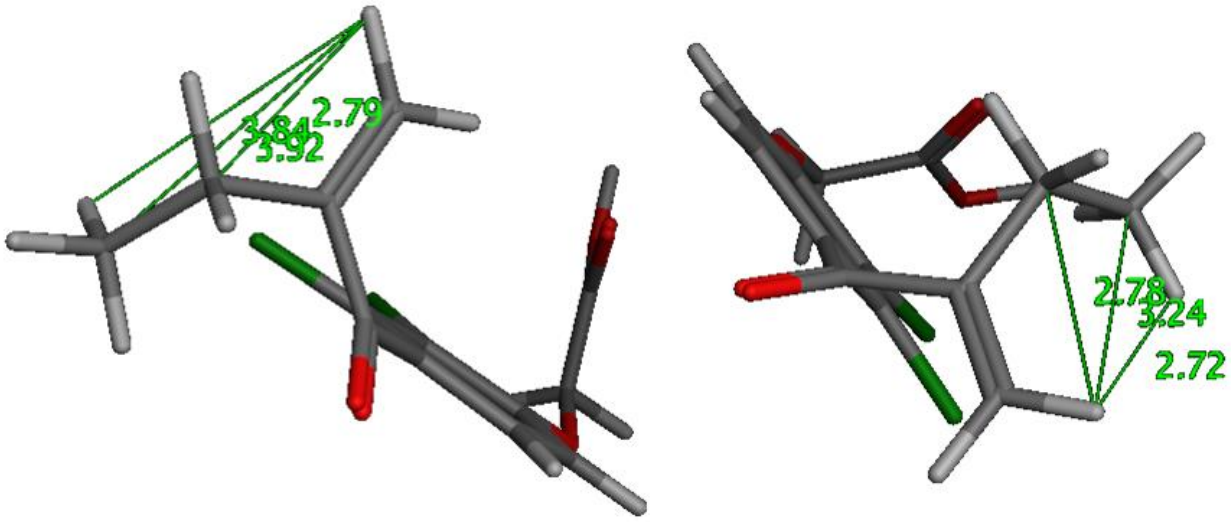


Figure 5: Showing the Closer Interaction Between Substituent Groups of the Carbon-Carbon Double Bond, from Lowest Energy Conformer on the Left to the Highest on the Right

Despite being the best performing force field for molecule 2, AMBER94 identified the wrong lowest energy conformer as calculated by M06-2X. Superimposing these conformers shows them to be nearly identical, only differing by a single rotation around a carbon-carbon bond, as seen in figure 6. This minor conformational change is reflected by the 0.692 kcal/mol difference in potential energy as calculated by AMBER94, considering the large energy ranges this is a relatively small difference.

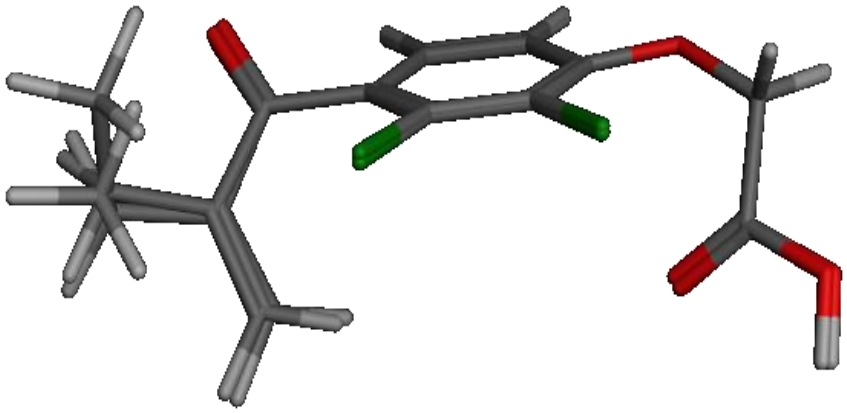


Figure 6: Lowest Energy Conformers as Identified by AMBER94 and M06-2X Superimposed

The performance of AMBER94 with molecule 2 is surprising as it is a force field used for peptides and proteins[29,30,95], furthermore, it performs poorly with small organic molecules in several studies[18,20,27]. However, AMBER94 has shown to perform reasonably with polar molecules[29,30,95] such as molecule 2. The fact that MMFF94 was outperformed is also surprising as it is a force field for small 'drug-like' ligands and several papers consistently show it to perform well with similar molecules[12-18].

It is important to remember the fundamentals of molecule mechanics, for example, nuclear position is considered rather than electrons and as such approximations are made, therefore, the validity of these approximations influences the validity of force fields[6-8,10]. Considering this, it is not irrational to suggest that a molecule with multiple polar groups and electronegative halogens may pose a problem to force fields. Furthermore, Lakdawala et al.[43] found Taxol conformer energy was vastly misrepresented by several force fields, deeming this as a result of the highly polar nature of Taxol. However, the density functional theory approach, B3LYP, which does consider electrons also performed poorly and was outperformed by many force fields, suggesting it may be the molecule which caused the problem as opposed to intrinsic force field problems.

PM6 also fell prey to molecule 2, producing the highest RMSE and some of the highest energy ranges in the whole study, as well as incorrectly identifying both the lowest and highest energy conformer. Comparing the highest energy conformer identified by PM6 and M06-2X suggests there was similar, if not more, allylic strain present in the highest energy conformer identified by PM6. Furthermore, close interaction between atoms in this problem conformer may have resulted in Van der Waals strain, whereby atoms are forced closer than their Van der Waal radius would normally allow[100].

4.1.2 Molecule 6

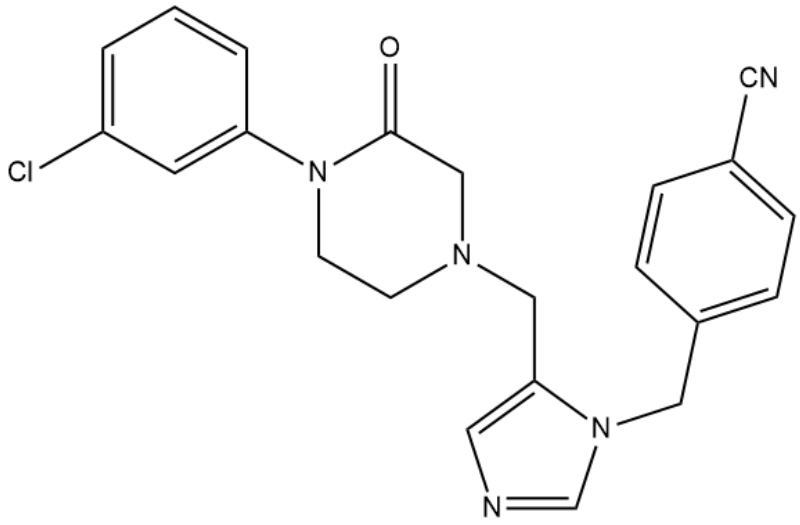


Figure 7: Chemical Structure of Molecule 6

Molecule 6 contains a 6 membered 'piperazine-like' ring with two amines and an attached carbonyl group, analysis of molecule 6 conformers shows that this structure appears distorted to some extent in all the conformers and could be described partly as 'half chair' shaped. The most stable conformation of cyclohexane is the 'chair' position and although this 'piperazine-like' ring is structurally different to cyclohexane, this distorted conformation is likely to be energetically unfavourable. Structurally and chemically molecule 2 is quite different from the others, for example, containing the 'piperazine-like' ring, as well as an aromatic nitrile and imidazole ring, these structures can be seen in figure 8 below.

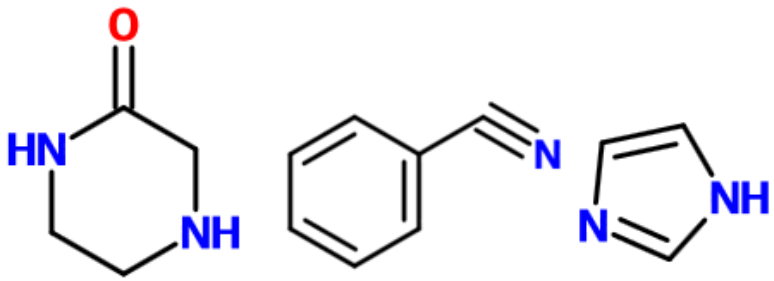


Figure 8: Showing the Chemical Groups of Molecule 6, the 'Piperazine-Like' Ring (Left), Aromatic Nitrile (Middle) and Imidazole Ring (Right)

Despite being the best performing force field, CHARMM27 incorrectly identified the highest energy conformer, however correctly identified the lowest. Comparing the highest energy conformer identified by CHARMM27 and M06-2X shows they both contain the distorted 'piperazine-like' ring, furthermore, the highest energy conformer identified by CHARMM27 has a close interaction of 1.65 angstroms between two atoms likely resulting in Van der Waals strain[100]. This close interaction is also present in the other problem conformers of molecule 6.

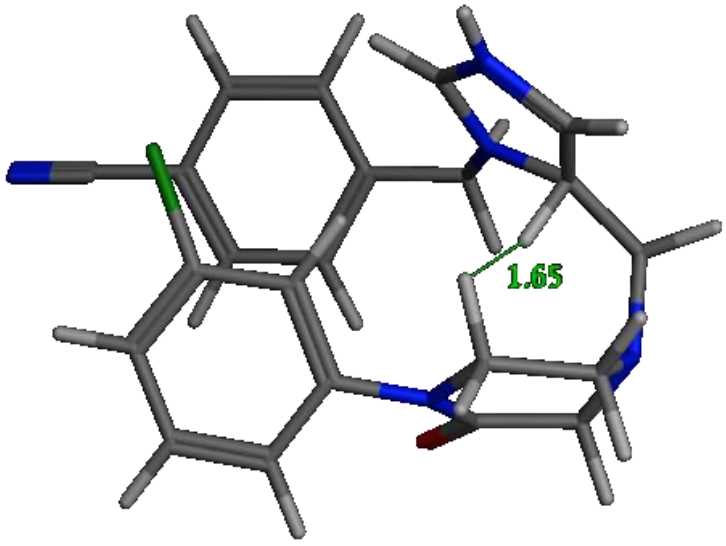


Figure 9: Showing Close Interaction between Two Hydrogen Atoms in the Highest Energy Conformer Identified by CHARMM27

MMFF94 was the only force field to incorrectly identify the lowest energy conformer of molecule 6, by superimposing the lowest energy conformers as calculated by MMFF94 and M06-2X we can see very similar conformation with a 108-degree bond rotation difference between the two conformers giving rise to two isomers, this can be seen in figure 10 and 11 below. Despite identifying the wrong lowest energy conformer, MMFF94 calculated a relative small potential energy difference of 1.17 kcal/mol between the two low energy conformers.

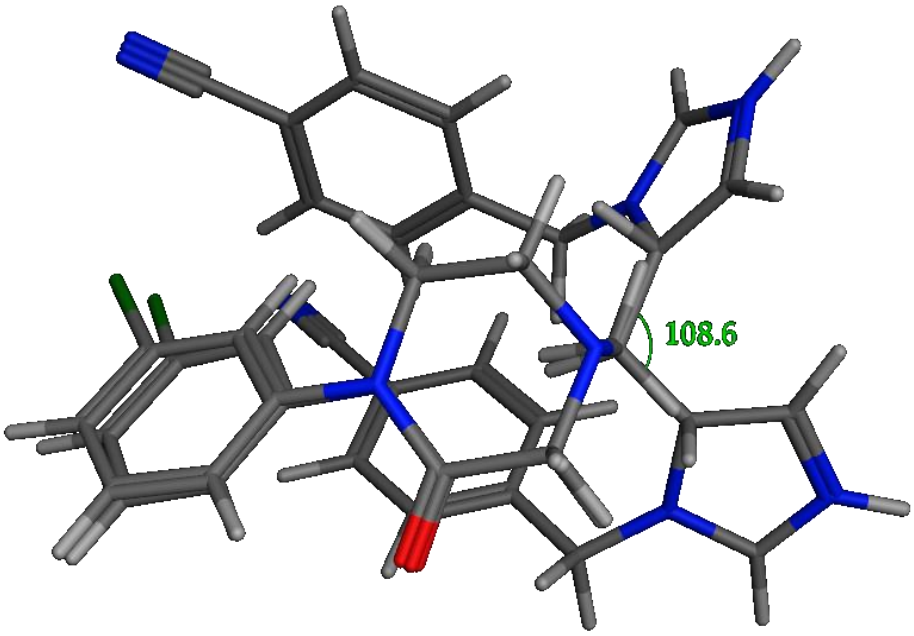


Figure 10: Lowest Energy Conformers Identified by MMFF94 and M06-2X Superimposed Showing Bond Rotation

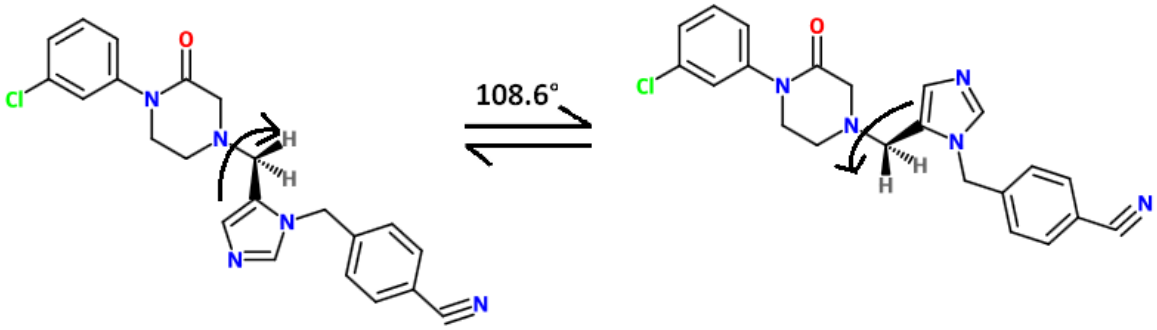


Figure 11: 2D Chemical Structures of Lowest Energy Conformers Identified by MMFF94 and M06-2X Showing Bond Rotation

Molecule 6 caused fewer problems for PM6 as it outperformed all methods, producing the lowest RMSE, around six times smaller than the best performing force field. Despite this, PM6 failed to correctly identify the lowest energy conformer. B3LYP showed fewer struggles with

this molecule compared to molecule 2, however, still produced large errors and struggled with similar conformers to the force fields.

4.1.3 Molecule 7

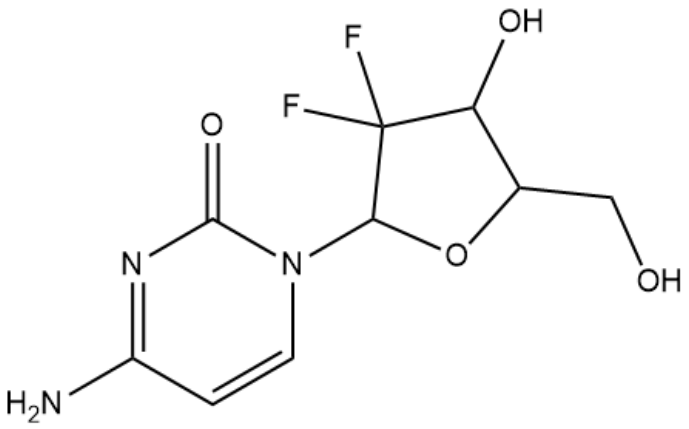


Figure 12: Chemical Structure of Molecule 7

PM6 showed severe difficulty with molecule 7, producing the highest RMSE and energy ranges in the whole study, with the RMSE over eleven times larger than the worst performing force field. PM6 surprisingly produced large R^2 values despite calculating relative energies over ten times larger than benchmark M06-2X data. Molecule 7 caused much less problems for the other methods; however, raised errors for this molecule were seen. Similar findings where PM6 overestimates results are seen in several studies, for example, when calculating energy of halogen bonds[104], estimating adsorption energies[105], as well as stabilisation energies of hydrogen bonded complexes[106].

Molecule 7 is the only molecule to contain cytosine, a pyrimidine derived nitrogenous base found in DNA and RNA. Similarly, it is the only molecule to contain two fluorine atoms and a tetrahydrofuran ring.

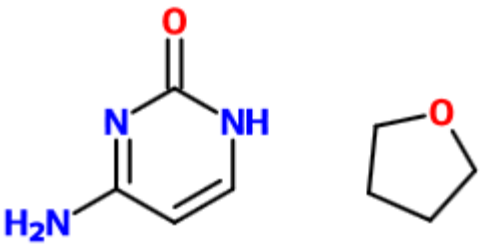


Figure 13: Showing the Chemical Groups in Molecule 7, Cytosine (Left) and Tetrahydrofuran Ring (Right)

Based on the large errors and energy ranges produced by PM6, it is likely that this method struggled with the majority of its conformers. This molecule is similar to molecule 2, in that it contains electronegative halogens; however, the fluorine atoms present in molecule 6 are more electronegative than the chlorine atoms in molecule 2. Furthermore, this molecule also contains similar heterocyclic rings seen in other problem molecules, however, there are less major conformational changes between conformers and for this reason the force fields may have handled these conformers better, as we can see in the figure below there is only a single bond rotation difference between the lowest and highest energy conformer. The presence of the two halogen atoms may have contributed to the poor performance of PM6 with molecule 7, as current literature has shown this method to perform poorly when calculating the energy of halogen bonds[105]; furthermore, PM6 also performs poorly with molecule 2 which also contains two halogen atoms.

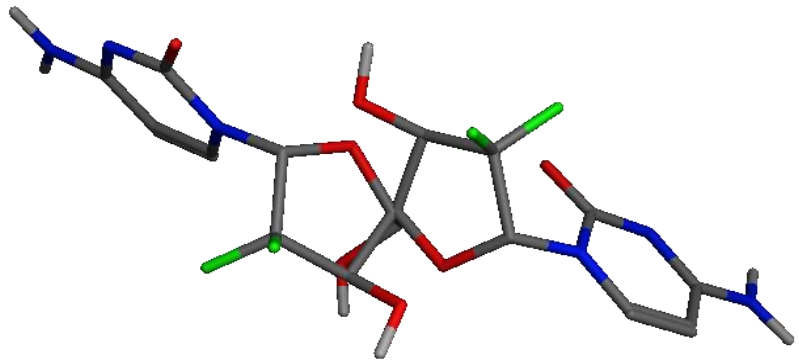


Figure 14: Lowest and Highest Energy Conformer of Molecule 7 Superimposed, Showing Single Bond Rotation

4.1.4 Molecule 10

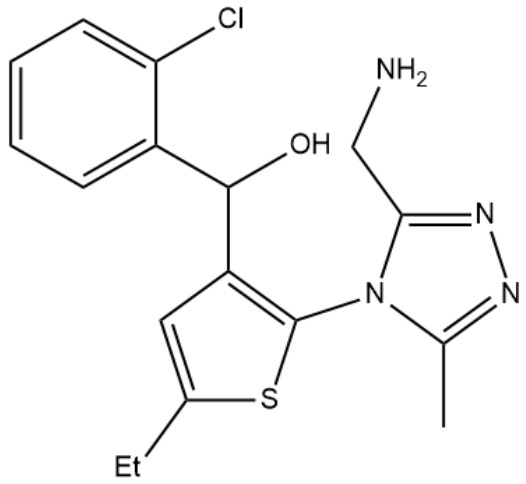


Figure 15: Chemical Structure of Molecule 10

Although molecule 10 does not have the largest errors or energy ranges, the R^2 values clearly suggest this molecule caused a problem. Molecule 10 is the only one to contain a triazole ring and is one of two to contain a thiophene ring; as such, the cause of the problem may be due to the presence of these structures. Marion et al.[101] suggested that even though force field parameters for triazole compounds do exist they 'are still lacking', this is further emphasised by growing studies developing triazole force field parameters[101-103]. This suggests that force fields may struggle with these types of compounds and as such may have been a hindrance with molecule 10.

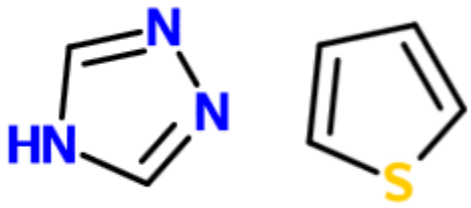


Figure 16: Showing the Chemical Groups in Molecule 10, Triazole Ring (Left) and Thiophene (Right)

MMFF94 produced a reasonable RMSE for this molecule considering that other force fields performed significantly worse. MMFF94 may have outperformed the other force fields as it has been shown by Halgren[12-18] to produce low errors in intermolecular energies and conformer

energies associated with thiophene compounds, therefore may have aided the force field with this thiophene containing molecule.

AMBER12: EHT is the best performing AMBER force field, producing an error over three times less than its counterparts. This can partly be explained by the parameters of AMBER94 having little to no changes since its early development[29,95], furthermore AMBER12: EHT is the most recent of the three which may reflect on the quality of its parameters. Furthermore, lone pairs are not considered by AMBER94 and as such it may have struggled with the sulphur of the thiophene ring[29,30,95].

Despite being the best performing force field, MMFF94 surprisingly calculated the highest energy conformer, as identified by M06-2X, as having the lowest energy. The correct highest energy conformer has multiple instances of Van der Waals strain and through comparison with the lowest energy conformer it is apparent that it does not contain these instances, therefore this finding is extremely puzzling.

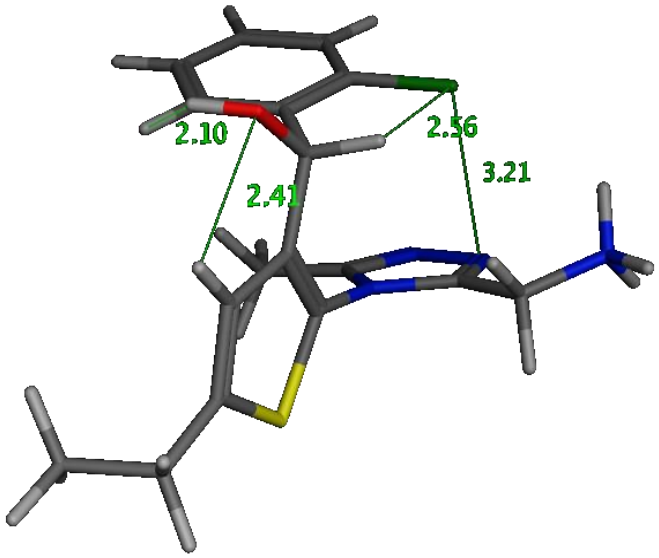


Figure 17: Showing Close Interactions Between Atoms in the Highest Energy Conformer Identified by M06-2X

4.2 Other Molecules

4.2.1 Molecule 1 and 3

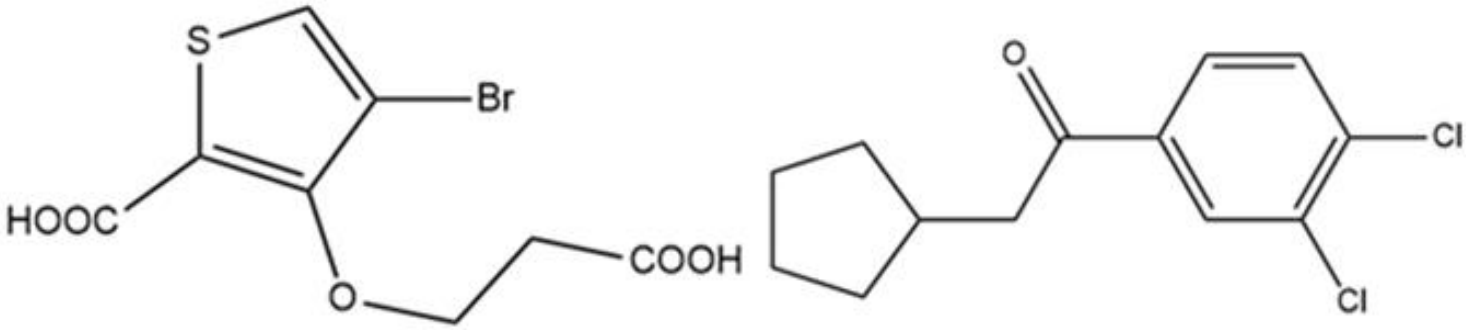


Figure 18: Chemical Structure of Molecule 1 (Left) and Molecule 3 (Right)

Molecule 1 and 3 appear to be handled seemingly well by the force fields, with the exception of AMBER94 and AMBER99 with molecule 3. These two molecules are relatively simple molecules when compared with others, for example, we can see that molecule 1 is quite small in comparison with molecule 10 and molecule 3 is composed of little more than an halogenated aromatic chlorine ring linked to cyclopentane. By comparing the conformers of each molecule we can see there is little conformational change between them, this is further reflected by the small differences in energy. Superimposing the lowest and highest energy conformer of each molecule shows only a single bond rotation difference, as seen in figure 18. As molecule 1 contains a thiophene ring we would expect this to cause a problem, however, unlike the other thiophene containing molecule, molecule 1 has only minor conformational differences between conformers.

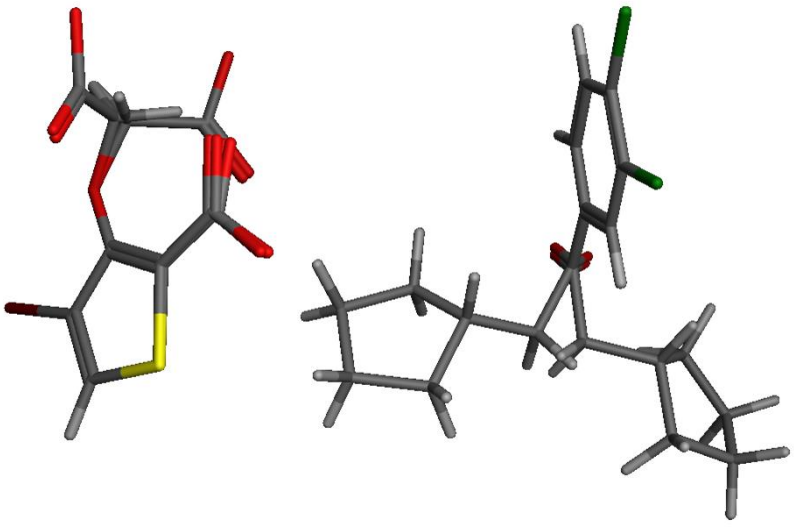


Figure 19: Lowest and Highest Energy Conformer of Molecule 1 and 3 Superimposed, Showing Single Bond Rotation

4.2.2 Molecule 4

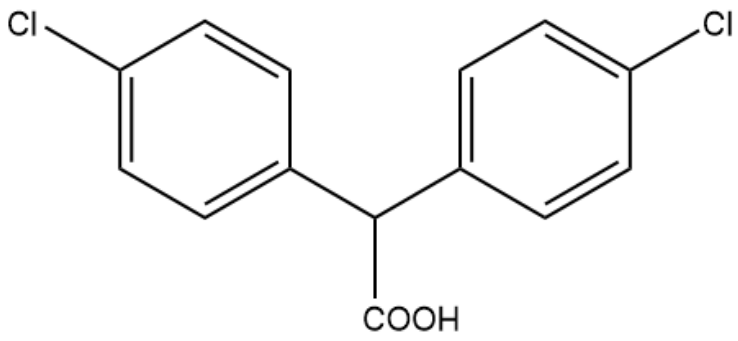


Figure 20: Chemical Structure of Molecule 4

The performance of the force fields with molecule 4 is clearly reflected by the large R^2 values and low energy ranges. Visually, molecule 4 is structurally simple and similar to molecule 5, consisting of two chlorinated aromatic rings joined in a 'para' position by a carbon link with an attached carboxyl group. Conformer analysis shows that there is little conformational changes between conformers which is reflected by the small changes in potential energy. Comparison of the lowest and highest energy conformer, as identified by M06-2X, shows only a 95 degree bond rotation difference, this can be seen in figure 20. AMBER99 produced the lowest RMSE value for molecule 4 and one of the lowest seen in the whole study. The combination of the molecules simple structure and only minor conformational changes between conformers likely allowed AMBER99, as well as the other force fields, to perform well. Both PM6 and B3LYP performed in a similar manner to the force fields.

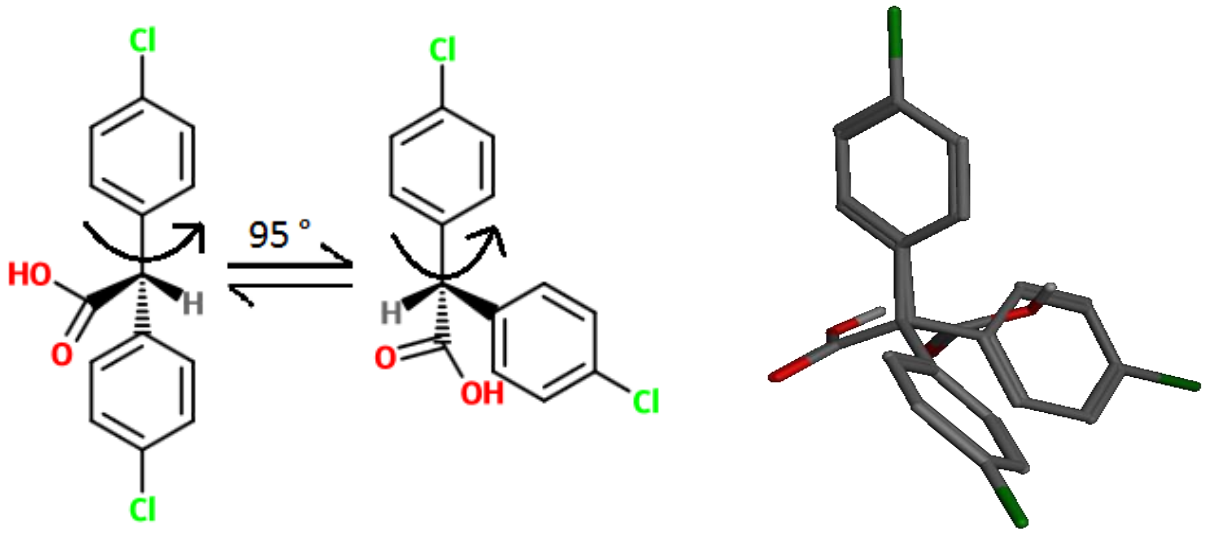


Figure 21 and 22: Showing Bond Rotation Between Lowest and Highest Energy Conformer

4.2.3 Molecule 5

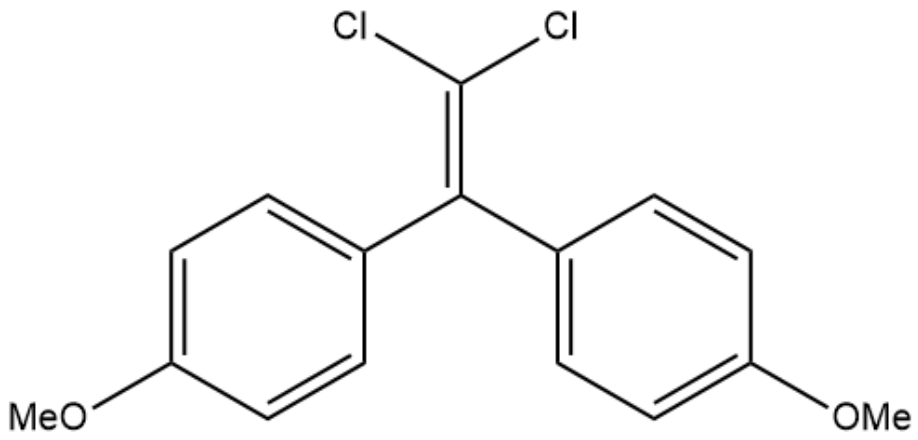


Figure 22: Chemical Structure of Molecule 5

Molecule 5 is also structurally simple, consisting of little more than two aromatic rings joined by a carbon link with two chlorine atoms attached. As with molecule 4, there are is an absence of major conformational changes between conformers, this can be seen by superimposing the lowest and highest energy conformer. Only AMBER12: EHT was able to correctly identify the lowest energy conformer; however, this can partly be explained by the small change in

potential energy between conformers. Furthermore, we can see that molecule 4 and 5 are symmetrical and this likely contributed to the performance of the force fields.

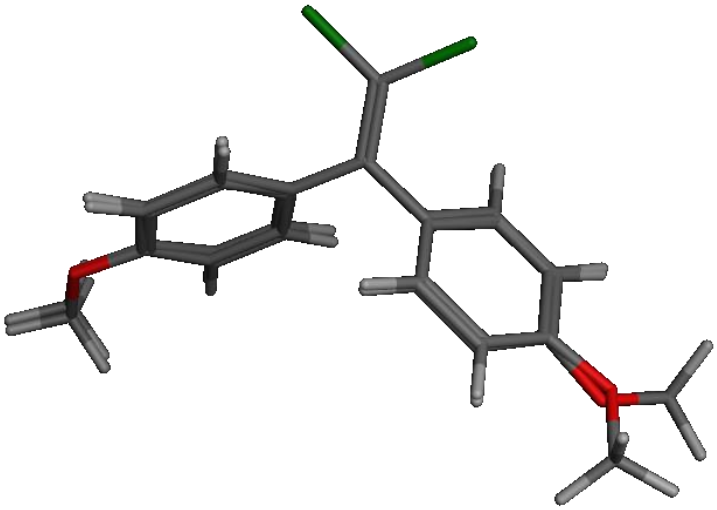


Figure 23: Lowest and Highest Energy Conformer of Molecule 5 Superimposed, Showing Minor Conformational Change Between Conformers

Despite performed well for molecule 5, there was an instance were a conformer may have led to a 'dip' in the R^2 of AMBER94, however, analysis of this conformer reveals little insight as it is almost identical to several conformers, including the lowest energy conformer.

The best performing force field, CHARMM27, likely outperformed the worst, AMBER99, as it is the newest of the two force fields and this likely reflects on the performance of its parameters[29,30,48-52]. Furthermore, AMBER99 may be better suited to biological molecules such as proteins as opposed to small drug like ligands such as molecule 5[29,30]. Additionally, CHARMM27 was designed for an array of molecules including ones similar to molecule 5[48,49] and performs well with similar molecules in numerous studies[18,50,51]. Furthermore, studies have found CHARMM27 to outperform other force fields, for example, outperforming both OPLS-AA and GAFF in reproducing conformer energy of organic compounds[53].

4.2.4 Molecule 9

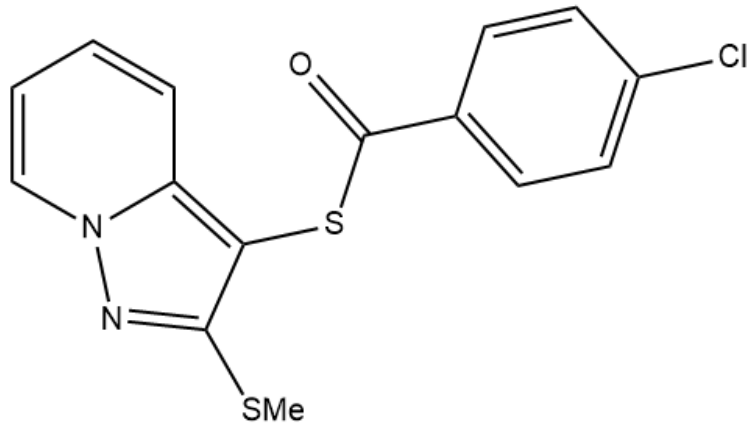


Figure 24: Chemical Structure of Molecule 9

Although molecule 9 did not have the lowest RMSE or energy ranges, it is clear from the R^2 values that the force fields performed reasonably well with this molecule. Even more so, the R^2 values produced for this molecule are some of the best in the whole study. Further analysis suggests the force fields had particular ease with three conformers.

Molecule 9 is structurally more complex than other molecules in the sample, furthermore, it is the only molecule to contain a 'pyridine pyrazole-like' ring and a methylsufanyl group, it also contains a chlorinated aromatic ring. Despite the presence of these structures the force fields did surprisingly well, as similar heterocyclic structures previously caused problems. This is likely due to the absence of major conformational changes between conformers, this can be seen in figure 25 where we can see that there is more changes between conformers in molecule 6 than molecule 9. This is further reflected by the relatively low changes in potential energy between conformers. One of the most important finding for this molecule is that all the forces correctly identified both the lowest and highest energy conformer, this is the only instance of this and both B3LYP and PM6 failed to do this for molecule 9.

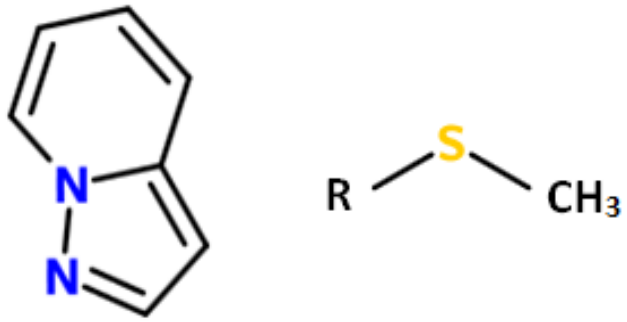


Figure 25: Showing Chemical Groups in Molecule 9, 'Pyridine Pyrazole-Like' Ring (Left) and Methylsufanyl Group (Right)

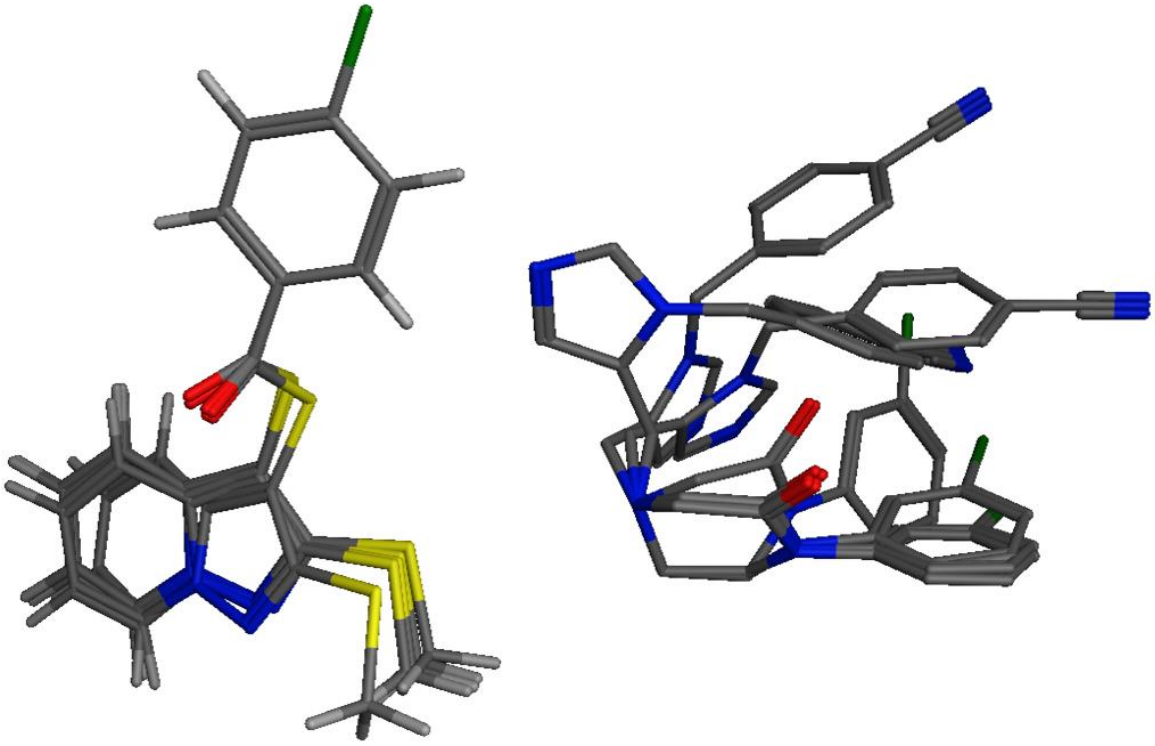


Figure 26: Comparison of Lowest and Highest Energy Conformers of Molecule 9 (Left) and 6 (Right), Showing Degree of Conformational Change.

4.3 Additional Thoughts

Based on these findings there are some recurring patterns with regards to force field performance. Firstly, close interactions between certain functional and atoms likely resulting in Van der Waals strain[100] was present in nearly all the problem molecules. Secondly, the presence of certain chemical structures deemed 'problem' structures, for example, the triazole in molecule 10, were also present in nearly all the problem molecules. Thirdly, the molecules which the force fields performed well with tended to be structurally simple and in cases, symmetrical. Finally, one of the more important observations was degree of major conformational changes between conformers.

The degree of major conformational changes between conformers seemed to be one of the most influential factors which affected force field performance, for example, molecules which performed well tended to have only minor differences between the lowest and highest conformer. This point is highlighted by molecules 3, 4 and 5 which only had a single bond rotation difference between their lowest and highest energy conformer. Furthermore, molecules which were structurally more complex and contained the 'problem structures' but had little conformational changes between conformers tended to perform better than the problem molecules, for example, molecule 9 which contained a heterocyclic 'piperidine pyrazole-like' ring however still produced some of the highest R^2 values in the whole study. When looking at molecules which contained these 'problem structures' as well as a high degree of conformational changes between conformers we can see the extent of the problem it causes, for example, molecule 6 which had some of the largest errors as well as ranges in relative and absolute energy of the whole study. This finding is somewhat like Kanal and Hutchison's[9], whereby they found that the problem molecules had a higher degree of torsional rotation as well as the presence of halides.

4.3.1 Impact of Findings

Although this study does have its limitations and the sample size is likely to limit the generalisability, the findings have huge implication as they go to show there are instances were force fields are extremely unreliable. As with Kanal and Hutchison's[9] study, these results

question how reliable force fields are and granted, it is known that force fields may not provide the most accurate results due to the fundamentals of their parametrisation, i.e., the validity of their approximations, however multiple studies are showing instances where force fields perform significantly poor[19-28,36-47] and this may lead to severe consequences. Due to the speed of force field calculations one of their more common applications is rapid conformer screening or searching, e.g., for an ensemble of the lowest or highest energy conformers, however, there was only instance in this study were all the force fields correctly identified both the lowest and highest conformer. Considering this, the reliability of force fields for this application may be undermined as these finding shows that force fields have a major problem with low and high energy conformer identification. This finding was also seen in a study investigating amino acid conformational energies, whereby, MMFF94 identified less than half of the lowest energy conformers identified by density functional theory. However, it is important to remember that the sample of molecules used in this study were deemed problem molecules by Kanal and Hutchison[9], therefore these findings may not apply to other molecules and cannot be fully generalised. Furthermore, torsional rotation and the presence of problem structures would need to be considered when using force fields for conformer searching, for example, we know that these can cause a problem and may result in the wrong ensemble of conformers therefore to counter these problems the problem structures should be avoided. Additionally, comparison of force field data with B3LYP for molecule 2 and 6 seems to suggest that the problem caused may be due to molecular factors as opposed to an intrinsic problem with the actual method.

The performance of PM6 in this study is somewhat erratic and unpredictable, for example, it performs reasonably well for some molecules but produces extremely high errors for others. This finding seems to oppose that of Kanal and Hutchison[9] whereby they deemed PM7 suitable for conformer searching. Although their semi-empirical approach is the more recent of the two PMx approaches, it still highlights the potential pathological performance of this method. As well as producing the highest errors in the whole study, PM6 was also outperformed by force fields several times. Considering the previous arguments, as well as the fact that PM6 had an overt problem with identifying the lowest and highest energy conformers, it would be advisable to use caution when applying this method to conformer searching.

Overall the impact these findings have can be considered constructive as it highlights the potential problems of force fields, therefore with this knowledge these problems can be avoided. When applying force fields to conformer searching it is important to think about the chemical and structural properties of the conformers in question, as well as the individual strengths and weaknesses of the force field. Therefore, when selecting a method for conformer searching it would be important to take all of this into account. Furthermore, this study highlights future areas of research that will aid a better understand of force fields and will ultimately lead to better parametrisation and future force field development.

Chapter 5

5.1 Suggestions for Future Research

As the sample size of this study is only ten molecules it limits the generalisability of the findings, as such it would be advisable for any future research to use a large sample size, perhaps including more molecules with a similar structure to the problem molecules found in this study. Including molecules with similar structures to the problem molecules in this study would likely provide a better understanding of the influence that these types of molecules have on force fields. Furthermore, the influence of torsional rotations and presence of these 'problem structures' would be a recommended area to investigate as it would help further force field parametrisations if the cause of the problem could be pinpointed. Also, as only the potential energy of the conformers was calculated it would be interesting to see how the force fields handled optimising these structures and whether the presence of the problem structures and torsional rotations influences optimisation as much as potential energy. It is important to highlight a potential limitation of this study in that the molecules and conformers that were used had been generated previously using an algorithm involving the MMFF94 force field[86,87] and as such this may have favoured towards the two MMFF force fields and the results regarding these two force fields may be biased. Furthermore, the force field data would have ideally been calculated using the AMBER suite, however, time did not permit this and as such this would be a further recommendation for future research.

5.2 Conclusion

The findings of this comparative study strongly agree with much of the literature regarding the poor performance of molecular mechanical force fields for conformer analysis. However, despite their poor performance with these so called 'problem molecules', this study also found instances where force fields performed reasonably well. With certain molecules, the energy produced by the force fields was in near perfect agreement with M06-2X reference data, as reflected by the large R^2 values. Considering this, it would be irrational to completely disregard the use of force fields, especially as they are an extremely efficient tool used commonly in drug discovery and development. When looking at the molecules which posed problems there are clear patterns which are present, as such, if these instances are accounted for during method selection then there is no reason why force fields cannot be used reliably and with confidence of accurate results. However, this much cannot be said for PM6 as it produced highly unpredictable and erratic results, performing seemingly well for one molecule but then butchering the next. Furthermore, this approach could not reliably identify the lowest and highest energy conformer in many instances, this finding is also seen in several other studies and PM6 is often outperformed by lower theory methods. Considering this, the application of PM6 for conformer searching should be used cautiously. This study has highlighted the potential problematic influence of torsional rotations and the presence of particular chemical groups on force field performance, as such further investigation of this influence would be of great importance to aid the development new force fields, as well as optimising current ones. The ranking of force field performance based on average root mean square error may have been potentially misleading in this study, and it is important to highlight that each individual force field has its own strengths and weaknesses; therefore, the choice of method should be well thought out and take into consideration the strengths and weaknesses of the method as well as the influence molecular factors can cause. As a concluding remark and to answer the title of this study, the reliability of force fields in conformer searching is not as clear cut as 'they are, or they are not reliable', there are many factors which affect reliability and many individual strengths and weaknesses of force fields, therefore, consideration of these factors when choosing a method for conformer searching will ultimately decide the reliability.

References

[1] Hughes J, Rees S, Kalindjian S, Philpott K. Principles of early drug discovery. British Journal of Pharmacology. 2011;162(6):1239-1249.

[2] Patrick G. An introduction to medicinal chemistry. 6th ed. Oxford University Press; 2017, pp. 197-221

[3] Leach, A. (2009). *Molecular modelling*. 2nd ed. Harlow: Pearson/Prentice Hall, pp.1-25. (chapter 1)

[4] Patrick G. An introduction to medicinal chemistry. 6th ed. Oxford University Press; 2017, pp. 349-392

[5] Ramachandran, K., Deepa, G. and Namboori, K. (2010). *Computational chemistry and molecular modeling*. Berlin: Springer, pp.1-15.

[6] Goodman J. Chemical applications of molecular modelling. Cambridge: Royal Society of Chemistry; 2002. Pp.15-30.

[7] Leach, A. (2009). *Molecular modelling*. 2nd ed. Harlow: Pearson/Prentice Hall, pp.165-252 (chapter 4)

[8] Leach, A. (2009). *Molecular modelling*. 2nd ed. Harlow: Pearson/Prentice Hall, pp.26-105 (chapter 2)

[9] Kanal I, Keith J, Hutchison G. A sobering assessment of small-molecule force field methods for low energy conformer predictions. International Journal of Quantum Chemistry. 2017;118(5):e25512.

[10] Goodman J. Chemical applications of molecular modelling. Cambridge: Royal Society of Chemistry; 2002. Pp.122-140.

[11] MMFF94 Force Field (mmff94) [Internet]. Open Babel 2.4 Documentation. [cited 9 March 2018]. Available from: http://open-babel.readthedocs.io/en/latest/Forcefields/mmff94.html

[12] Halgren T. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. Journal of Computational Chemistry. 1996;17(5-6):490-519.

[13] Halgren T. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. Journal of Computational Chemistry. 1996;17(5-6):520-552.

[14] Halgren T. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. Journal of Computational Chemistry. 1996;17(5-6):553-586.

[15] Halgren T, Nachbar R. Merck molecular force field. IV. conformational energies and geometries for MMFF94. Journal of Computational Chemistry. 1996;17(5-6):587-615.

[16] Halgren T. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. Journal of Computational Chemistry. 1996;17(5-6):616-641.

[17] Halgren T. MMFF VI. MMFF94s option for energy minimization studies. Journal of Computational Chemistry. 1999;20(7):720-729.

[18] Halgren T. MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. Journal of Computational Chemistry. 1999;20(7):730-748.

[19] Paton R, Goodman J. Hydrogen Bonding and π-Stacking: How Reliable are Force Fields? A Critical Evaluation of Force Field Descriptions of Nonbonded Interactions. Journal of Chemical Information and Modeling. 2009;49(4):944-955.

[20] Kaminský J, Jensen F. Force Field Modeling of Amino Acid Conformational Energies. Journal of Chemical Theory and Computation. 2007;3(5):1774-1788.

[21] Sulimov A, Zheltkov D, Oferkin I, Kutov D, Katkova E, Tyrtyshnikov E et al. Evaluation of the novel algorithm of flexible ligand docking with moveable target-protein atoms. Computational and Structural Biotechnology Journal. 2017;15:275-285.

[22] Sulimov A, Kutov D, Katkova E, Ilin I, Sulimov V. New generation of docking programs: Supercomputer validation of force fields and quantum-chemical methods for docking. Journal of Molecular Graphics and Modelling. 2017;78:139-147.

[23] Sulimov A, Kutov D, Katkova E, Sulimov V. Combined Docking with Classical Force Field and Quantum Chemical Semiempirical Method PM7. Advances in Bioinformatics. 2017;2017:1-6.

[24] Perola E, Charifson P. Conformational Analysis of Drug-Like Molecules Bound to Proteins: An Extensive Study of Ligand Reorganization upon Binding. Journal of Medicinal Chemistry. 2004;47(10):2499-2510.

[25] Sellers B, James N, Gobbi A. A Comparison of Quantum and Molecular Mechanical Methods to Estimate Strain Energy in Druglike Fragments. Journal of Chemical Information and Modeling. 2017;57(6):1265-1275.

[26] Fu Z, Li X, Merz K. Accurate assessment of the strain energy in a protein-bound drug using QM/MM X-ray refinement and converged quantum chemistry. Journal of Computational Chemistry. 2011;32(12):2587-2597.

[27] Kaminský J, Jensen F. Conformational Interconversions of Amino Acid Derivatives. Journal of Chemical Theory and Computation. 2016;12(2):694-705.

[28] Kaminski G, Jorgensen W. Performance of the AMBER94, MMFF94, and OPLS-AA Force Fields for Modeling Organic Liquids. The Journal of Physical Chemistry. 1996;100(46):18010-18013.

[29] Pearlman D, Case D, Caldwell J, Ross W, Cheatham T, DeBolt S et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Computer Physics Communications. 1995;91(1-3):1-41.

[30] Case D, Cheatham T, Darden T, Gohlke H, Luo R, Merz K et al. The Amber biomolecular simulation programs. Journal of Computational Chemistry. 2005;26(16):1668-1688.

[31] Nguyen T, Mai B, Li M. Study of Tamiflu Sensitivity to Variants of A/H5N1 Virus Using Different Force Fields. Journal of Chemical Information and Modeling. 2011;51(9):2266-2276.

[32] Yang Z, Wang Z, Tian X, Xiu P, Zhou R. Amino acid analogues bind to carbon nanotube via π-π interactions: Comparison of molecular mechanical and quantum mechanical calculations. The Journal of Chemical Physics. 2012;136(2):025103.

[33] Mobley D, Bayly C, Cooper M, Dill K. Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations†. The Journal of Physical Chemistry B. 2009;113(14):4533-4537.

[34] Shirts M, Pitera J, Swope W, Pande V. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. The Journal of Chemical Physics. 2003;119(11):5740-5761.

[35] Aliev A, Courtier-Murias D. Experimental Verification of Force Fields for Molecular Dynamics Simulations Using Gly-Pro-Gly-Gly. The Journal of Physical Chemistry B. 2010;114(38):12358-12375.

[36] Fluitt A, de Pablo J. An Analysis of Biomolecular Force Fields for Simulations of Polyglutamine in Solution. Biophysical Journal. 2015;109(5):1009-1018.

[37] Balaji G, Nagendra H, Balaji V, Rao S. Experimental conformational energy maps of proteins and peptides. Proteins: Structure, Function, and Bioinformatics. 2017;85(6):979-1001.

[38] Svozil D, Šponer J, Marchan I, Pérez A, Cheatham T, Forti F et al. Geometrical and Electronic Structure Variability of the Sugar−phosphate Backbone in Nucleic Acids. The Journal of Physical Chemistry B. 2008;112(27):8188-8197.

[39] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics. 2006;65(3):712-725.

[40] Doemer M, Guglielmi M, Athri P, Nagornova N, Rizzo T, Boyarkin O et al. Assessing the performance of computational methods for the prediction of the ground state structure of a cyclic decapeptide. International Journal of Quantum Chemistry. 2012;113(6):808-814.

[41] Valdes H, Pluháčková K, Pitonák M, Řezáč J, Hobza P. Benchmark database on isolated small peptides containing an aromatic side chain: comparison between wave function and density

functional theory methods and empirical force field. Physical Chemistry Chemical Physics. 2008;10(19):2747.

[42] Grubišić S, Brancato G, Pedone A, Barone V. Extension of the AMBER force field to cyclic α,α dialkylated peptides. Physical Chemistry Chemical Physics. 2012;14(44):15308.

[43] Lakdawala A, Wang M, Nevins N, Liotta D, Rusinska-Roszak D, Lozynski M et al. Calculated conformer energies for organic molecules with multiple polar functionalities are method dependent: Taxol (case study). BMC Chemical Biology. 2001;1(2).

[44] Galli C, Sensi C, Fumagalli A, Parravicini C, Marinovich M, Eberini I. A Computational Approach to Evaluate the Androgenic Affinity of Iprodione, Procymidone, Vinclozolin and Their Metabolites. PLoS ONE. 2014;9(8):e104822.

[45] Dokainish H, Gauld J. Formation of a Stable Iminol Intermediate in the Redox Regulation Mechanism of Protein Tyrosine Phosphatase 1B (PTP1B). ACS Catalysis. 2015;5(4):2195-2202.

[46] Ruiz-Gómez G, Hawkins J, Philipp J, Künze G, Wodtke R, Löser R et al. Rational Structure-Based Rescaffolding Approach to De Novo Design of Interleukin 10 (IL-10) Receptor-1 Mimetics. PLOS ONE. 2016;11(4):e0154046.

[47] Guan L, Ohtsuka J, Okai M, Miyakawa T, Mase T, Zhi Y et al. A new target region for changing the substrate specificity of amine transaminases. Scientific Reports. 2015;5(1).

[48] Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. Journal of Computational Chemistry. 1983;4(2):187-217.

[49] Brooks B, Brooks C, Mackerell A, Nilsson L, Petrella R, Roux B et al. CHARMM: The biomolecular simulation program. Journal of Computational Chemistry. 2009;30(10):1545-1614.

[50] Foloppe N, MacKerell, Jr. A. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. Journal of Computational Chemistry. 2000;21(2):86-104.

[51] MacKerell A, Banavali N. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. Journal of Computational Chemistry. 2000;21(2):105-120.

[52] MacKerell A, Banavali N, Foloppe N. Development and current status of the CHARMM force field for nucleic acids. Biopolymers. 2000;56(4):257-265.

[53] Nemkevich A, Bürgi H, Spackman M, Corry B. Molecular dynamics simulations of structure and dynamics of organic molecular crystals. Physical Chemistry Chemical Physics. 2010;12(45):14916.

[54] He X, Lopes P, MacKerell A. Polarizable Empirical Force Field for Acyclic Polyalcohols Based on the Classical Drude Oscillator. Biopolymers. 2013;99(10):724-738.

[55] Jorgensen W, Maxwell D, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. Journal of the American Chemical Society. 1996;118(45):11225-11236.

[56] Robertson M, Tirado-Rives J, Jorgensen W. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. Journal of Chemical Theory and Computation. 2015;11(7):3499-3509.

[57] Caleman C, van Maaren P, Hong M, Hub J, Costa L, van der Spoel D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. Journal of Chemical Theory and Computation. 2011;8(1):61-74.

[58] Lennox M, Bound M, Henley A, Besley E. The right isotherms for the right reasons? Validation of generic force fields for prediction of methane adsorption in metal-organic frameworks. Molecular Simulation. 2017;43(10-11):828-837.

[59] ab initio | Definition of ab initio in English by Oxford Dictionaries [Internet]. Oxford Dictionaries | English. 2018 [cited 17 March 2018]. Available from: https://en.oxforddictionaries.com/definition/ab_initio

[60] Pople J, Santry D, Segal G. Approximate Self-Consistent Molecular Orbital Theory. I. Invariant Procedures. The Journal of Chemical Physics. 1965;43(10):S129-S135.

[61] Stewart J. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. Journal of Molecular Modeling. 2007;13(12):1173-1213.

[62] Stewart J. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. Journal of Molecular Modeling. 2012;19(1):1-32.

[63] Pople J, Beveridge D. Approximate Molecular Orbital Theory. New York: McGraw-Hill; 1970.

[64] Kohn W, Becke A, Parr R. Density Functional Theory of Electronic Structure. The Journal of Physical Chemistry. 1996;100(31):12974-12980.

[65] Stephens P, Devlin F, Chabalowski C, Frisch M. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. The Journal of Physical Chemistry. 1994;98(45):11623-11627.

[66] Tirado-Rives J, Jorgensen W. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. Journal of Chemical Theory and Computation. 2008;4(2):297-306.

[67] Bourcier S, Hoppilliard Y. B3LYP DFT molecular orbital approach, an efficient method to evaluate the thermochemical properties of MALDI matrices. International Journal of Mass Spectrometry. 2002;217(1-3):231-244.

[68] Coskun D, Jerome S, Friesner R. Evaluation of the Performance of the B3LYP, PBE0, and M06 DFT Functionals, and DBLOC-Corrected Versions, in the Calculation of Redox Potentials and Spin Splittings for Transition Metal Containing Systems. Journal of Chemical Theory and Computation. 2016;12(3):1121-1128.

[69] Burk P, Koppel I, Koppel I, Leito I, Travnikova O. Critical test of performance of B3LYP functional for prediction of gas-phase acidities and basicities. Chemical Physics Letters. 2000;323(5-6):482-489.

[70] Zhao Y, Truhlar D. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. The Journal of Chemical Physics. 2006;125(19):194101.

[71] Zhao Y, Truhlar D. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average than B3LYP for Ground States. The Journal of Physical Chemistry A. 2006;110(49):13126-13130.

[72] Zhao Y, Truhlar D. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theoretical Chemistry Accounts. 2007;120(1-3):215-241.

[73] Walker M, Harvey A, Sen A, Dessent C. Performance of M06, M06-2X, and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions. The Journal of Physical Chemistry A. 2013;117(47):12590-12600.

[74] Luo S, Zhao Y, Truhlar D. Validation of electronic structure methods for isomerization reactions of large organic molecules. Physical Chemistry Chemical Physics. 2011;13(30):13683.

[75] Steinmann S, Wodrich M, Corminboeuf C. Overcoming systematic DFT errors for hydrocarbon reaction energies. Theoretical Chemistry Accounts. 2010;127(5-6):429-442.

[76] Rokob T, Hamza A, Pápai I. Computing Reliable Energetics for Conjugate Addition Reactions. Organic Letters. 2007;9(21):4279-4282.

[77] Zhao Y, Truhlar D. A Density Functional That Accounts for Medium-Range Correlation Energies in Organic Chemistry. Organic Letters. 2006;8(25):5753-5755.

[78] Wodrich M, Corminboeuf C, Schreiner P, Fokin A, Schleyer P. How Accurate Are DFT Treatments of Organic Energies?. Organic Letters. 2007;9(10):1851-1854.

[79] Zhao Y, Truhlar D. How Well Can New-Generation Density Functionals Describe Protonated Epoxides Where Older Functionals Fail?. The Journal of Organic Chemistry. 2007;72(1):295-298.

[80] Reis H, Rasulev B, Papadopoulos M, Leszczynski J. Reliable but Timesaving: In Search of an Efficient Quantum-chemical Method for the Description of Functional Fullerenes. Current Topics in Medicinal Chemistry. 2015;15(18):1845-1858.

[81] Lu X, Gaus M, Elstner M, Cui Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. The Journal of Physical Chemistry B. 2014;119(3):1062-1082.

[82] Hosseinnejad T, Kazemi T. Quantum chemical investigation on complexation of palladium with iminopyridyl ligands: Structural, thermochemical, and electronic aspects. Molecular Crystals and Liquid Crystals. 2016;637(1):53-64.

[83] Budyka M, Oshkin I. Comparative semiempirical and DFT study of styrylnaphthalenes and styrylquinolines and their photocyclization products. International Journal of Quantum Chemistry. 2010;:n/a-n/a.

[84] Li A, Muddana H, Gilson M. Quantum Mechanical Calculation of Noncovalent Interactions: A Large-Scale Evaluation of PMx, DFT, and SAPT Approaches. Journal of Chemical Theory and Computation. 2014;10(4):1563-1575.

[85] Ebejer J, Morris G, Deane C. Freely Available Conformer Generation Methods: How Good Are They?. Journal of Chemical Information and Modeling. 2012;52(5):1146-1158.

[86] Hartshorn M, Verdonk M, Chessari G, Brewerton S, Mooij W, Mortenson P et al. Diverse, High-Quality Test Set for the Validation of Protein−Ligand Docking Performance. Journal of Medicinal Chemistry. 2007;50(4):726-741.

[87] Hawkins P, Skillman A, Warren G, Ellingson B, Stahl M. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. Journal of Chemical Information and Modeling. 2010;50(4):572-584.

[88] Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. Journal of Chemical Information and Modeling. 1988;28(1):31-36.

[89] O'Boyle N, Banck M, James C, Morley C, Vandermeersch T, Hutchison G. Open Babel: An open chemical toolbox. Journal of Cheminformatics. 2011;3(1):33.

[90] Becke A. Density-functional exchange-energy approximation with correct asymptotic behavior. Physical Review A. 1988;38(6):3098-3100.

[91] Lee C, Yang W, Parr R. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Physical Review B. 1988;37(2):785-789.

[92] Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F.

Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

[93] F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy," *Phys. Chem. Chem. Phys.*, **7** (2005) 3297-305.

[94] F. Weigend, "Accurate Coulomb-fitting basis sets for H to Rn," *Phys. Chem. Chem. Phys.*, **8** (2006) 1057-65.

[95] Cornell W, Cieplak P, Bayly C, Gould I, Merz K, Ferguson D et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. Journal of the American Chemical Society. 1995;117(19):5179-5197.

[96] Kollman P. Advances and Continuing Challenges in Achieving Realistic and Predictive Simulations of the Properties of Organic and Biological Molecules. Accounts of Chemical Research. 1996;29(10):461-469.

[97] Hoffmann R. An Extended Hückel Theory. I. Hydrocarbons. The Journal of Chemical Physics. 1963;39(6):1397-1412.

[98] *Molecular Operating Environment (MOE)*, 2016.08; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2018**.

[99] eMolecules [Internet]. Emolecules.com. 2018 [cited 5 May 2018]. Available from: https://www.emolecules.com/

[100] Bondi, A. (1964). Van der Waals Volumes and Raii. *The Journal of Physical Chemistry,* 68(3), pp.441-451.

[101] Marion, A., Góra, J., Kracker, O., Fröhr, T., Latajka, R., Sewald, N. and Antes, I. (2017). Amber-Compatible Parametrization Procedure for Peptide-like Compounds: Application to 1,4- and 1,5-Substituted Triazole-Based Peptidomimetics. *Journal of Chemical Information and Modeling*, 58(1), pp.90-110.

[102] Eike, D. and Maginn, E. (2006). Atomistic simulation of solid-liquid coexistence for molecular systems: Application to triazole and benzene. *The Journal of Chemical Physics*, 124(16), p.164503.

[103] Cadena, C. and Maginn, E. (2006). Molecular Simulation Study of Some Thermophysical and Transport Properties of Triazolium-Based Ionic Liquids. *The Journal of Physical Chemistry B*, 110(36), pp.18026-18039.

[104] Řezáč J, Hobza P. A halogen-bonding correction for the semiempirical PM6 method. Chemical Physics Letters. 2011;506(4-6):286-289.

[105] Gomez D, Combariza A, Sastre G. Quantum-chemistry calculations of hydrogen adsorption in MOF-5. Physical Chemistry Chemical Physics. 2009;11(40):9250.

[106] Řezáč J, Fanfrlík J, Salahub D, Hobza P. Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. Journal of Chemical Theory and Computation. 2009;5(7):1749-1760.

[107] Chen I, Foloppe N. Tackling the conformational sampling of larger flexible compounds and macrocycles in pharmacology and drug discovery. Bioorganic & Medicinal Chemistry. 2013;21(24):7898-7920.

[108] Deeth R. General Molecular Mechanics Method for Transition Metal Carboxylates and its Application to the Multiple Coordination Modes in Mono- and Dinuclear Mn(II) Complexes. Inorganic Chemistry. 2008;47(15):6711-6725.

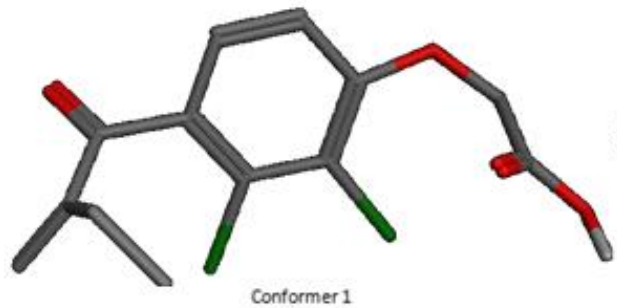
[109] Deeth R, Fey N. A Molecular Mechanics Study of Copper(II)-Catalyzed Asymmetric Diels−Alder Reactions. Organometallics. 2004;23(5):1042-1054.

[110] Karthikeyan C, Santosh D, Hari Narayana Moorthy N, Trivedi P. QSAR Analysis of HIV-1 Reverse Transcriptase Inhibitory 5-Alkyl-2- [(Aryl and Alkyloxylcarbonylmethyl) Thio]-6- (1-Napthylmethyl) Pyrimidin-4 (3H)-Ones. International Journal of Virology. 2007;3(1):19-27.

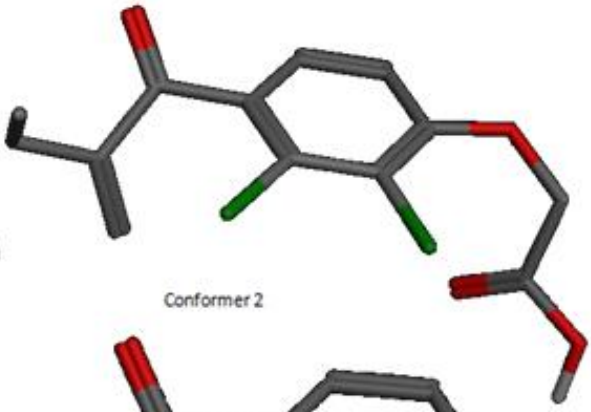
]

Appendix A: Conformer Structures

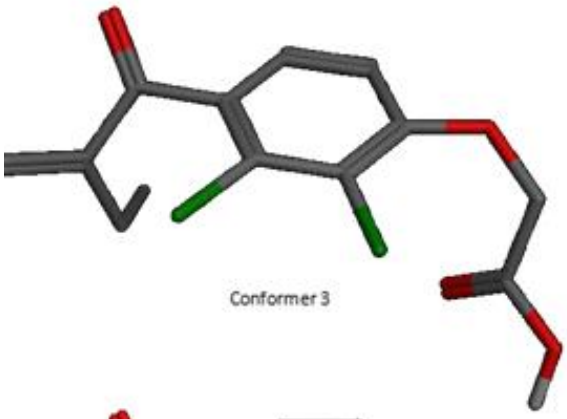
Molecule 2



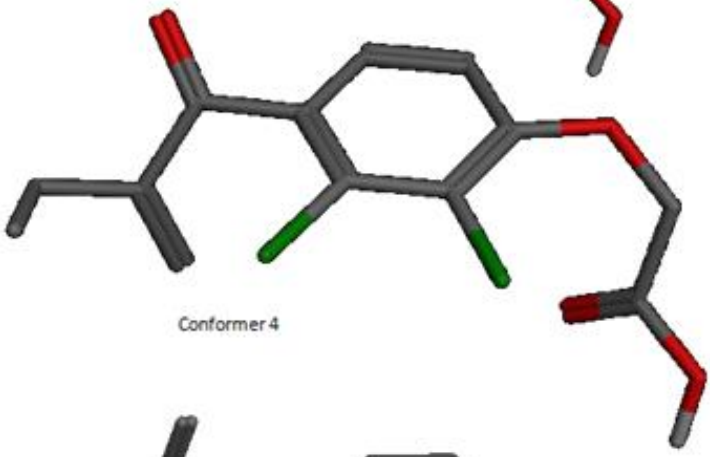
Conformer 1



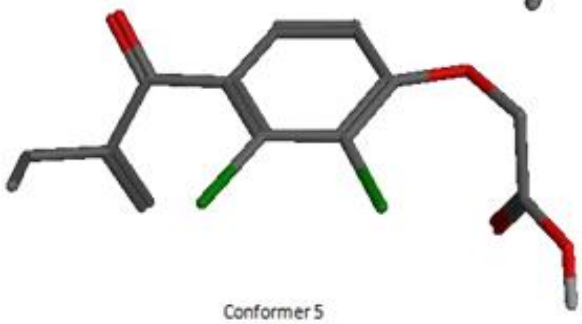
Conformer 2



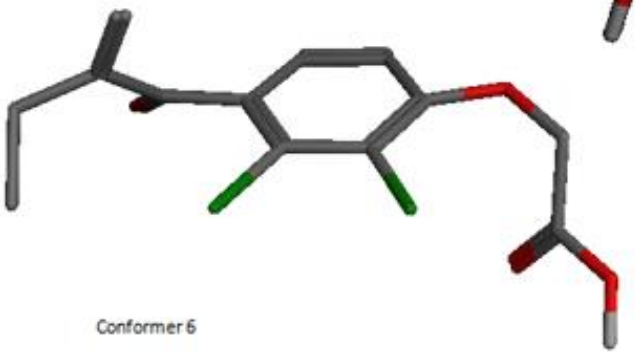
Conformer 3



Conformer 4

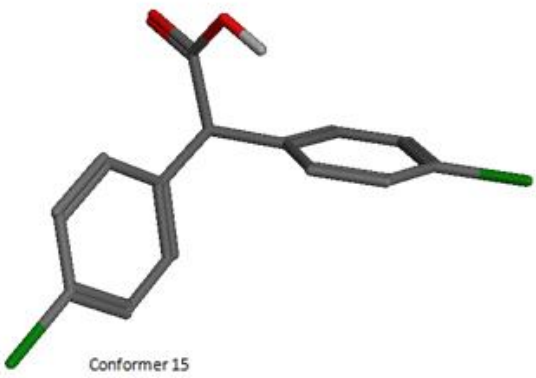


Conformer 5



Conformer 6

Molecule 4

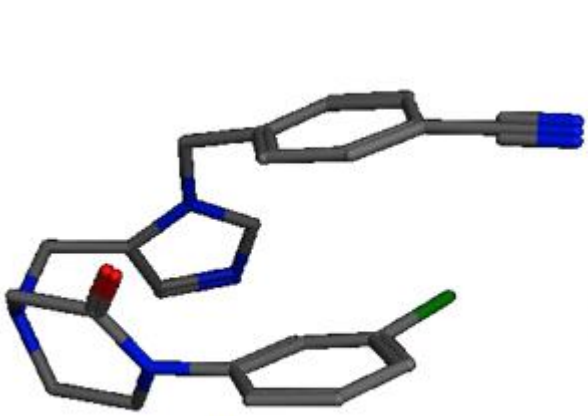


Conformer 15

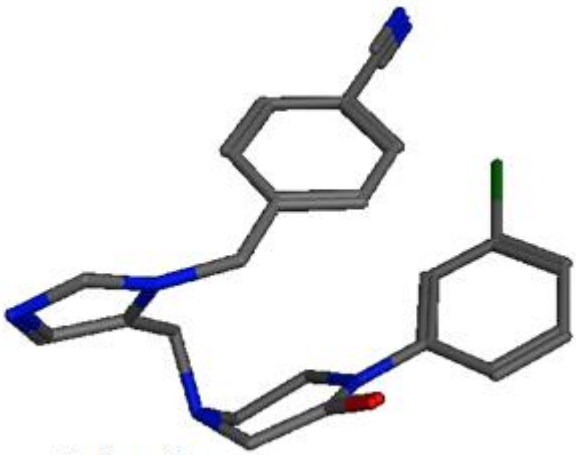
Molecule 5

Conformer 16

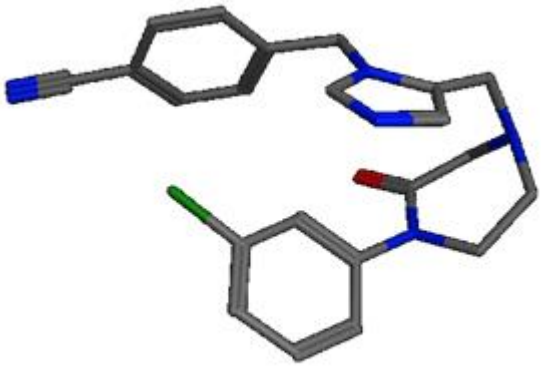
Molecule 6



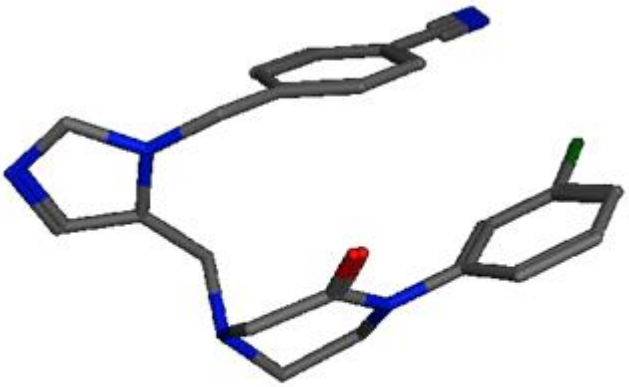
Conformer 7



Conformer 8

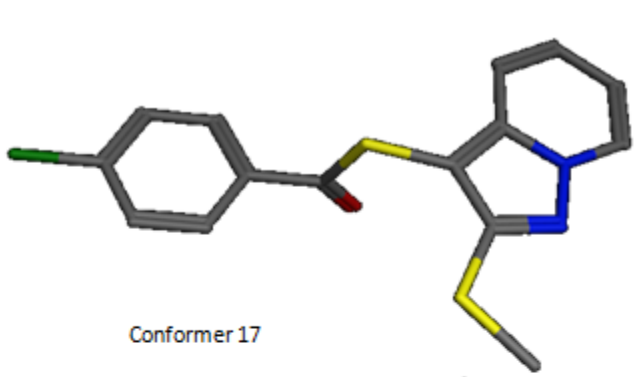


Conformer 9

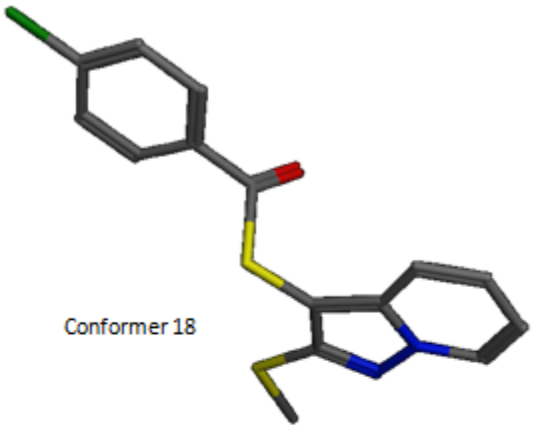


Conformer 10

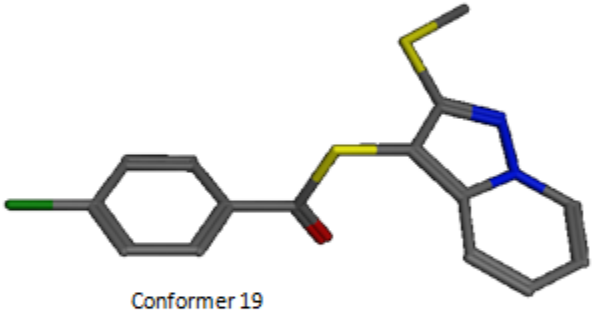
Molecule 9



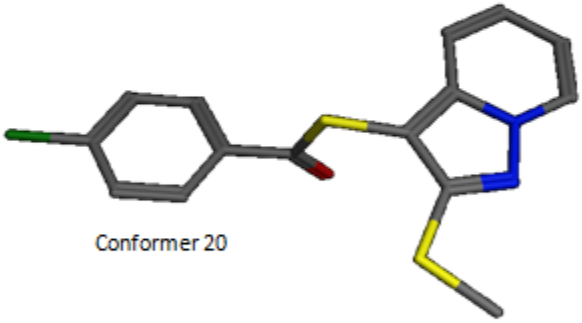
Conformer 17



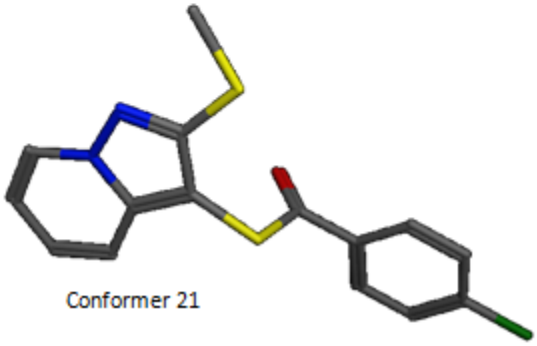
Conformer 18



Conformer 19

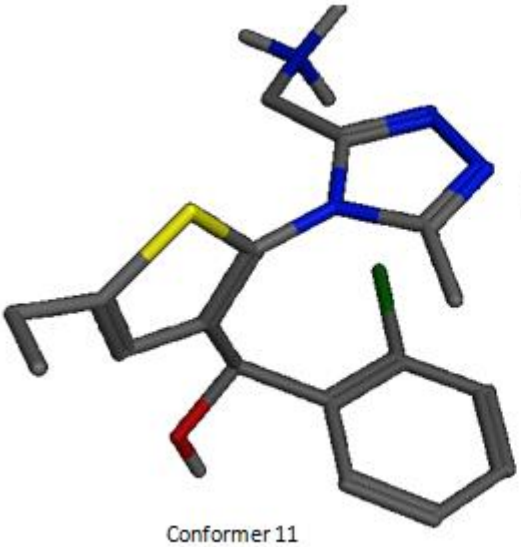


Conformer 20

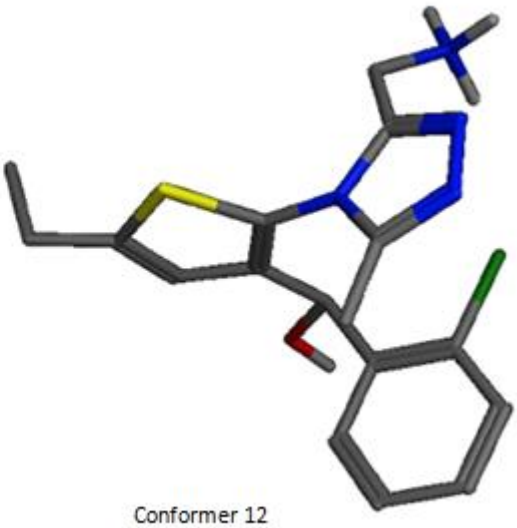


Conformer 21

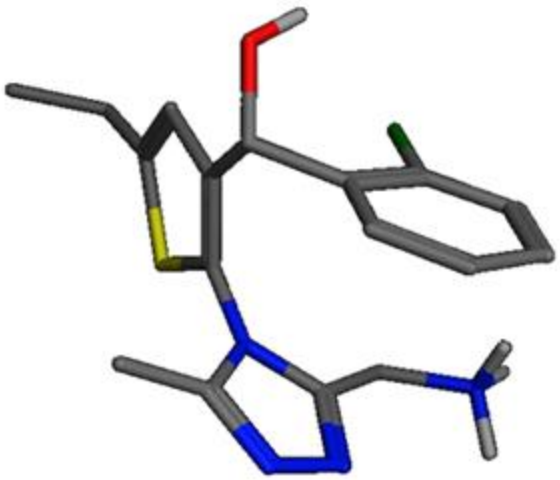
Molecule 10



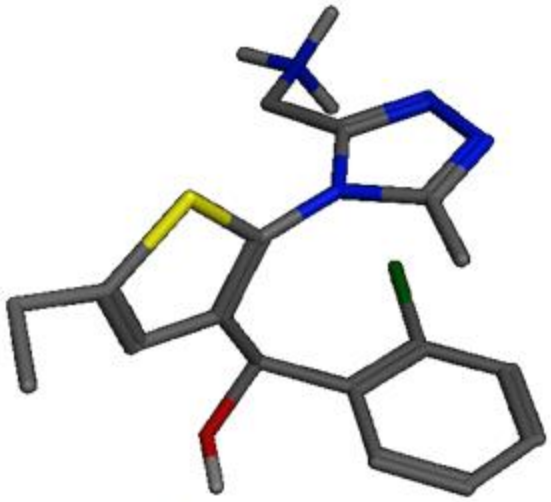
Conformer 11



Conformer 12



Conformer 13



Conformer 14

Appendix B: Relative Energy Rankings

Molecule 1

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.19	0.23	0.48	0.53	0.45	0.21	0.52	0.66	0.05
0.24	0.55	0.63	1.01	0.62	0.25	0.53	1.38	0.49
0.39	0.64	0.84	1.31	0.94	0.70	0.86	2.05	0.52
0.76	0.72	1.14	1.47	1.15	0.72	1.14	2.16	0.83
0.76	0.82	1.15	1.82	1.28	1.04	1.21	2.82	0.84
0.97	0.97	1.64	2.07	1.70	1.05	1.70	3.11	1.07
0.99	1.03	2.09	2.14	1.99	1.28	2.08	3.38	1.15
1.05	1.12	2.26	2.64	2.39	1.36	2.13	4.07	2.06
1.85	1.93	2.97	3.77	2.99	2.27	2.82	6.41	2.15

Molecule 2

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.12	0.08	1.19	4.97	1.27	0.57	0.56	0.83	0.59
0.77	0.18	1.59	7.03	3.45	0.65	1.15	1.07	0.65
1.12	0.19	2.39	7.14	3.91	0.73	2.44	2.78	0.73
1.78	0.23	2.66	7.62	4.16	1.55	3.52	3.53	1.55
2.45	0.24	3.34	7.91	4.55	1.60	4.11	3.82	1.60
2.46	0.27	3.61	8.85	4.89	1.64	4.22	4.64	1.64
3.14	0.93	3.74	9.35	5.19	1.83	4.87	5.09	1.83
4.30	1.20	5.17	10.98	5.63	3.01	6.59	7.10	3.01
8.85	1.73	5.31	11.64	10.45	3.12	7.02	8.14	3.12

Molecule 3

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.28	0.13	0.34	1.62	1.66	0.01	0.14	0.03	0.33
0.48	0.45	0.72	4.82	5.05	0.38	0.30	0.22	0.34
0.54	0.71	0.91	5.17	5.36	1.04	0.35	0.26	0.71
0.63	1.89	1.19	5.42	5.42	1.22	0.47	0.48	1.37
1.07	2.05	1.41	6.03	6.25	1.24	1.26	1.44	1.55
1.10	2.76	2.31	8.02	7.09	1.33	1.41	1.47	1.57
1.38	3.73	3.05	9.34	8.39	1.49	3.33	1.77	1.66

1.53	4.29	3.42	10.13	10.35	1.67	3.34	1.99	1.82
2.69	5.30	4.57	13.02	11.85	2.13	4.04	2.46	2.41

Molecule 4

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.21	0.05	0.03	0.04	0.09	0.01	0.33	0.54	0.03
0.35	0.09	0.11	0.19	0.18	0.13	0.45	0.62	0.16
0.45	0.14	0.12	0.29	0.27	0.41	0.47	0.63	0.23
0.51	0.19	0.13	0.32	0.28	0.47	0.52	0.66	0.44
0.54	0.20	0.21	0.39	0.38	0.49	0.58	0.76	0.50
0.63	0.29	0.27	0.48	0.43	0.50	0.74	1.15	0.52
0.80	0.30	0.33	0.59	0.51	0.54	1.14	1.44	0.53
0.86	0.31	0.47	0.67	0.57	1.09	1.14	1.52	0.57
1.19	0.41	0.66	0.89	0.83	1.09	1.15	1.56	1.12

Molecule 5

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.11	0.34	0.03	0.02	0.36	0.17	0.05	0.03	0.21
0.11	0.43	0.07	0.04	0.37	0.21	0.22	0.05	0.49
0.13	0.99	0.11	0.04	0.41	0.49	0.24	0.05	0.59
0.15	1.00	0.13	0.04	0.63	0.59	0.25	0.06	0.59
0.20	1.16	0.15	0.06	0.85	0.59	0.25	0.09	0.68
0.22	1.49	0.17	0.09	1.01	0.68	0.31	0.12	0.82
0.44	2.43	0.20	0.10	1.77	0.82	0.35	0.13	0.84
0.44	2.61	0.29	0.10	1.98	0.84	0.36	0.16	0.87
0.44	2.82	0.30	0.23	1.98	0.87	0.45	0.23	1.17

Molecule 6

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
11.78	2.40	3.65	3.13	3.35	0.96	3.51	3.41	15.59
12.47	2.51	4.82	3.61	4.08	2.72	3.84	6.24	16.75
12.49	2.51	6.33	4.32	5.10	2.86	4.18	6.36	18.46
12.62	3.08	9.65	7.36	5.38	3.19	4.81	7.85	18.93
12.84	3.19	13.49	10.45	10.88	13.22	8.39	12.93	19.64

12.90	3.23	15.01	13.80	13.05	16.82	10.78	16.52	32.16
13.28	3.28	21.76	21.16	26.33	25.37	20.34	29.10	31.29
18.71	3.30	-	-	-	-	-	-	-
20.22	3.31	-	-	-	-	-	-	-

Molecule 7

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
10.56	66.34	3.86	2.50	2.97	2.81	8.56	7.77	1.50
14.27	83.64	3.88	5.00	5.42	3.29	11.02	8.13	2.18
14.34	85.43	4.16	5.44	5.78	3.62	11.19	8.58	2.20
14.51	86.78	5.05	5.62	5.90	4.49	11.50	8.91	3.04
14.53	89.11	5.39	6.22	6.56	4.64	11.76	9.28	4.89
14.59	95.08	5.43	7.36	8.09	5.76	12.66	10.70	5.55
15.43	102.35	5.72	7.53	8.17	6.12	12.79	11.08	5.80
16.12	102.42	5.81	7.54	8.17	6.23	13.20	11.18	6.59
16.44	114.82	7.26	7.98	8.44	6.31	15.59	11.85	14.33

Molecule 8

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.31	0.04	0.03	0.09	0.59	0.21	0.14	0.49	0.32
0.60	0.12	0.77	0.25	0.59	2.61	0.87	1.27	2.82
0.82	0.16	3.37	3.63	0.65	5.40	10.99	2.19	5.43
1.12	0.19	3.59	3.81	0.69	5.57	10.99	3.22	5.85
1.35	0.22	4.33	3.89	0.87	6.06	11.03	3.25	6.33
1.74	0.27	4.37	4.25	1.26	6.51	11.40	3.45	6.69
1.87	0.39	4.72	4.96	1.77	6.66	12.24	3.88	6.86
2.23	0.44	5.02	5.82	2.21	6.86	12.41	4.05	7.20
2.70	0.45	-	-	-	-	-	-	-

Molecule 9

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.44	0.61	0.55	2.04	2.03	0.28	0.55	0.57	0.28
0.91	0.90	1.02	3.27	3.29	0.77	0.90	1.13	0.77
1.92	1.45	1.90	7.13	7.12	1.38	1.68	1.95	1.38

1.98	1.49	1.97	7.32	7.31	1.38	1.71	1.98	1.38
2.11	1.57	1.98	7.66	7.67	1.40	1.76	2.16	1.40
2.20	1.68	2.15	7.79	7.76	1.64	1.78	2.18	1.64
2.44	2.21	2.24	8.30	8.32	1.66	1.78	2.34	1.66
2.51	2.43	2.28	8.56	8.57	1.72	1.87	2.39	1.72
2.57	2.87	2.43	8.59	8.61	1.93	2.07	2.66	1.93

Molecule 10

B3LYP	PM6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF94s
0	0	0	0	0	0	0	0	0
0.73	0.07	0.02	0.06	0.21	0.45	0.19	0.18	0.45
0.99	0.26	0.22	0.25	0.52	0.74	0.20	0.38	0.74
1.65	1.54	0.23	4.06	4.05	0.79	3.89	2.36	0.79
1.80	2.03	0.24	5.98	5.97	0.87	8.01	2.87	0.87
1.81	2.11	0.29	6.09	6.10	1.17	8.14	2.97	1.17
1.87	2.23	0.33	6.39	6.33	1.19	8.14	3.12	1.19
1.90	2.33	1.03	6.46	6.49	1.25	8.25	3.26	1.25
2.02	2.78	1.23	6.53	6.52	1.37	8.33	3.28	1.37
4.50	2.79	1.34	6.79	6.86	1.44	8.68	3.38	1.44

Appendix C: Absolute Energy Rankings

Molecule 1

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.1587	0.0073	0.0445	0.3263	0.3733	0.2863	0.0159	0.2929	0.5053	0.1137
0.2361	0.0299	0.0677	0.3919	0.7699	0.3799	0.0483	0.3663	1.1449	0.1422
0.3112	0.0605	0.1068	0.4462	0.7922	0.4792	0.0492	0.4692	1.4922	0.1592
0.6728	0.0689	0.1189	0.4702	0.9958	0.5842	0.2291	0.5092	1.7408	0.2038
0.6988	0.0779	0.1559	0.5258	1.1252	0.6238	0.3402	0.5468	2.1212	0.2293
0.8189	0.0854	0.1678	0.8211	1.2213	0.8811	0.3543	0.8831	2.2941	0.2501
0.9217	0.0966	0.3115	1.1643	1.2461	1.0703	0.3838	1.1563	2.4623	0.2519
0.9551	0.1488	0.3145	1.3079	1.6869	1.3709	0.4029	1.1759	3.1149	0.5289
1.6171	0.2324	0.3266	1.3489	2.1549	1.4329	0.6519	1.2079	4.7879	1.1069

Molecule 2

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
24.8612	23.9911	24.6886	23.1233	17.3163	21.4111	24.2141	21.7073	21.1963	24.2141
24.8641	24.0145	24.7793	23.2791	17.8331	22.1315	24.2922	22.6112	22.3825	24.2722
26.2433	24.0977	26.0527	23.3892	18.1715	22.1722	24.8393	22.8045	22.8583	24.8393
26.3853	24.2715	26.1511	23.4165	18.7282	22.2283	25.1242	22.8683	22.9012	25.1242
26.7242	24.5633	26.4866	23.6692	18.7683	22.3353	25.2893	23.2032	22.9842	25.2893
27.0215	24.6076	26.7502	23.7273	18.8122	22.4227	25.3825	23.7171	23.4683	25.3825
28.0762	24.7414	27.0936	23.8513	19.1013	22.6653	25.5113	23.8003	23.7921	25.5113
28.2953	24.9356	27.1457	24.3412	19.8902	22.8862	26.2502	24.2992	24.0362	26.2502
32.8687	25.1234	31.1391	27.5577	21.2277	23.5932	29.7467	25.8477	24.7337	29.7467

Molecule 3

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.1277	0.0745	0.0049	0.0955	1.5528	1.5353	0.1137	0.0113	0.0977	0.0000
0.6235	0.1307	0.0050	0.2040	4.4995	4.4235	0.1940	0.2941	0.1300	0.0830
0.7100	0.1365	0.1745	0.2163	4.8029	4.6510	0.2405	0.3184	0.2821	0.1993
0.7641	0.1440	1.1045	0.4229	5.0204	4.6529	0.2964	0.3285	0.4015	0.2825
0.9436	0.1522	1.1245	0.4634	5.5361	5.3104	0.3039	0.3610	0.4449	0.4705
1.0261	0.1672	1.7310	1.2819	7.4854	6.0639	0.3235	0.3789	0.4470	0.5005
1.1565	0.1795	2.5701	1.8975	8.7426	7.2335	0.3310	1.7170	0.4984	0.5409
1.3495	0.2223	2.9371	2.0725	9.4262	9.0045	0.3365	1.9875	0.6175	0.6039
2.3250	0.3682	2.9764	2.2460	11.7954	9.5210	0.4589	2.1725	0.6445	0.6064

6.3.4 Molecule 4

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.0430	0.1593	0.0090	0.0140	0.0000	0.0081	0.0326	0.2085	0.3649	0.0051
0.1671	0.1622	0.0809	0.0581	0.0126	0.0085	0.0330	0.2159	0.3735	0.0063
0.2575	0.1789	0.1025	0.0977	0.0199	0.0093	0.0351	0.2259	0.3969	0.0130
0.2921	0.1817	0.1201	0.1322	0.0209	0.0129	0.0524	0.2696	0.4529	0.0225
0.3651	0.1890	0.1640	0.1385	0.0259	0.0129	0.1259	0.2879	0.4940	0.0404
0.4674	0.2170	0.1768	0.1551	0.0315	0.0184	0.1485	0.2900	0.6826	0.0536
0.4896	0.2907	0.1873	0.1596	0.0974	0.0378	0.1749	0.3528	0.7648	0.1319
0.5657	0.3133	0.2596	0.1671	0.0978	0.0424	0.2938	0.5773	0.9504	0.1439
0.7942	0.3974	0.3882	0.1934	0.1003	0.0450	0.5213	0.6504	0.9543	0.3238

Molecule 5

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.0118	0.0183	0.3300	0.0113	0.0032	0.2376	0.1237	0.0247	0.0142	0.1952
0.0833	0.0256	0.3487	0.0202	0.0403	0.2857	0.1602	0.0262	0.0330	0.2904
0.1674	0.0316	0.8271	0.0375	0.1234	0.3482	0.2608	0.0352	0.1125	0.2959
0.1685	0.0395	0.8305	0.0534	0.1245	0.4655	0.2669	0.0706	0.1134	0.4037
0.2233	0.0675	0.9378	0.0733	0.1653	0.6257	0.2744	0.0775	0.1383	0.4226
0.2868	0.0997	1.2070	0.1208	0.1948	0.7242	0.3196	0.1357	0.1678	0.4245
0.5466	0.1099	1.8882	0.2801	0.3752	1.2254	0.3697	0.1552	0.3712	0.4607
0.5701	0.1307	2.0363	0.3012	0.4506	1.3728	0.3972	0.1946	0.4121	0.5342
0.6052	0.1605	2.2116	0.3486	0.4711	1.4049	0.4215	0.2121	0.4176	0.5668

Molecule 6

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.2218	10.0774	0.0690	3.4242	2.0416	2.8226	0.5756	1.9046	3.1842	15.3632
0.2746	10.1423	0.2262	4.0486	2.9072	2.9864	0.7392	2.4134	4.0846	16.1846
2.2794	10.2109	0.2937	4.5434	3.3334	3.1282	0.7944	3.2902	5.4554	16.4764
2.3966	10.2205	0.4069	7.2484	4.9604	3.8054	2.4474	3.5664	5.9624	16.5294
2.7018	10.2988	0.4921	10.7842	7.7472	8.1732	10.5192	5.6882	10.2282	16.9402
2.8271	11.5588	0.6794	12.1779	10.9769	10.2199	13.9919	7.9489	13.6879	28.3063
2.9817	12.1947	1.2297	18.7803	18.1773	23.3513	22.3923	17.3553	26.1213	29.3359

3.2294	15.4778	2.1774	-	-	-	-	-	-	-
4.5392	15.6813	2.2308	-	-	-	-	-	-	-

Molecule 7

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
8.1284	2.4284	58.2106	4.2704	4.7176	3.9916	5.3214	0.1146	0.3564	1.1088
11.1400	2.5158	72.5043	5.9552	4.7836	4.1386	6.1926	0.1160	1.2336	6.0591
11.3056	2.5177	74.1212	6.5896	5.1031	4.4741	6.3166	0.1708	1.3682	6.5076
11.6658	2.8395	75.1146	6.6178	5.2412	4.7832	6.4201	0.2560	1.3786	6.5236
12.0140	3.0338	77.1000	6.6210	5.6284	5.1544	6.9132	0.4316	1.4651	6.6334
12.0786	3.1232	83.0054	6.6446	5.7950	5.4510	7.1788	0.4844	2.7226	7.1240
12.3096	3.1289	89.7700	6.8391	5.8696	5.5216	7.3750	0.5519	2.7330	8.6278
12.6451	3.2222	90.0425	7.1446	6.0498	5.7190	7.6876	0.5774	2.7548	8.9620
13.2182	3.4775	101.5997	7.2570	6.1440	5.7648	7.8460	2.3708	3.0140	9.1016

Molecule 8

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.4975	0.1021	0.3853	0.2624	0.2556	0.0254	0.2835	0.3595	0.0105	0.1795
0.5026	0.1884	0.4369	0.4685	0.4115	0.0360	2.1114	0.3624	0.7654	2.3124
0.6140	0.2105	0.4511	2.7590	2.9856	0.0691	4.7840	10.1276	1.5770	4.8110
0.6229	0.4512	0.4588	2.9651	3.0130	0.0904	4.9511	10.3489	2.3486	5.2281
0.8994	0.4962	0.6758	3.3249	3.1841	0.0905	5.1556	10.3671	2.3989	5.4266
1.0481	0.6495	0.7566	3.4266	3.2039	0.2069	5.3819	10.3760	2.5769	5.6389
1.1428	0.6904	0.7814	3.5389	3.8152	0.6312	5.4609	10.9319	2.5991	5.7132
1.4761	0.7225	1.0402	3.5772	4.3439	0.7299	5.5152	11.0932	2.7372	5.7219
2.0462	0.7584	1.5925	-	-	-	-	-	-	-

Molecule 9

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.5075	0.0661	0.0744	0.0405	1.5285	1.5235	0.2072	0.0425	0.0605	0.2072
0.9742	0.0687	0.0994	0.0468	2.2938	2.3178	0.2245	0.0722	0.1089	0.2245
2.0998	0.1746	0.1012	0.2018	5.0312	5.0242	0.7228	0.4218	0.1272	0.7228
2.1776	0.1767	0.2628	0.2066	5.1454	5.1344	0.7966	0.4646	0.1478	0.7966
2.2862	0.1780	0.4100	0.3052	5.3208	5.2978	0.8242	0.5262	0.1548	0.8242
2.4662	0.1826	0.6470	0.3132	5.3688	5.3838	0.8339	0.6872	0.1936	0.8339
2.6179	0.1975	0.6846	0.3359	5.6831	5.6991	0.8892	0.7009	0.2809	0.8892

2.6964	0.2020	0.7171	0.3829	5.8231	5.8481	0.9599	0.8314	0.2812	0.9599
2.7669	0.2624	0.7903	0.4154	5.8626	5.8706	0.9744	0.8379	0.3064	0.9744

Molecule 10

M06-2X	b3lyp	pm6	AMBER12:EHT	AMBER94	AMBER99	MMFF94	CHARMM27	OPLS-AA	MMFF9s
0	0	0	0	0	0	0	0	0	0
0.6852	0.0474	0.2146	0.5674	0.6282	0.4792	0.2402	0.4972	0.5072	0.2402
1.2198	0.0936	0.5064	0.5948	0.9718	0.6968	0.3703	1.0218	0.8398	0.3703
1.3272	0.2202	0.5753	0.6642	2.2022	2.2692	0.3738	2.5658	1.0378	0.3738
1.5232	0.2304	0.6157	1.0008	2.7318	2.7258	0.3761	4.0882	1.2098	0.3761
1.5353	0.2745	0.6626	1.0992	4.4558	4.4448	0.4274	6.4828	1.3428	0.4274
1.5641	0.2755	0.7054	1.2361	4.5537	4.5667	0.4808	6.5336	1.4337	0.4808
1.6198	0.2814	0.9577	1.2473	4.7316	4.7246	0.5342	6.5729	1.4796	0.5342
1.7964	0.3080	0.9805	1.2872	4.8289	4.7669	0.6542	6.6017	1.5539	0.6542
4.5898	0.3265	1.7993	3.2498	4.8382	4.8702	3.1498	6.6342	1.6402	3.1498

Appendix D

RMSE

Molecule	B3LYP	PM6	AMBER12: EHT	AMBER 94	AMBER 99	MMFF 94	CHARMM 27	OPLS-AA	MMFF94s
1	0.10	0.20	0.81	1.19	0.84	0.32	0.77	2.35	0.42
2	23.23	25.38	22.84	17.93	21.28	24.36	22.25	21.98	24.36
3	0.18	1.72	1.23	6.80	5.95	0.29	1.10	0.41	0.40
4	0.23	0.19	0.13	0.06	0.02	0.21	0.36	0.61	0.12
5	0.09	1.29	0.18	0.26	0.82	0.29	0.12	0.24	0.39
6	11.34	1.11	9.48	8.36	10.01	10.10	7.33	11.55	19.34
7	2.79	76.88	6.16	5.22	4.78	6.50	0.82	1.98	6.75
8	0.50	0.78	2.68	2.83	0.33	4.32	8.61	1.98	4.50
9	0.17	0.48	0.27	4.66	4.66	0.73	0.55	0.19	0.73
10	0.23	0.84	1.36	3.51	3.50	1.07	4.90	1.21	1.07

Relative Energy Ranges

Molecule	B3LYP	PM6	AMBER12: EHT	AMBER 94	AMBER 99	MMFF 94	CHARMM 27	OPLS-AA	MMFF94s
1	1.85	1.93	2.97	3.77	2.99	2.27	2.82	6.41	2.15
2	8.85	1.73	5.31	11.64	10.45	3.12	7.02	8.14	3.12
3	2.69	5.30	4.57	13.02	11.85	2.13	4.04	2.46	2.41
4	1.19	0.41	0.66	0.89	0.83	1.09	1.15	1.56	1.12
5	0.44	2.82	0.30	0.23	1.98	0.87	0.45	0.23	1.17
6	20.22	3.31	21.76	21.16	26.33	25.37	20.34	29.10	31.29
7	16.44	114.82	7.26	7.98	8.44	6.31	15.59	11.85	14.33
8	2.70	0.45	5.02	5.82	2.21	6.86	12.41	4.05	7.20
9	2.57	2.87	2.43	8.59	8.61	1.93	2.07	2.66	1.93
10	4.50	2.79	1.34	6.79	6.86	1.44	8.68	3.38	1.44

Absolute Energy Ranges

Molecule	M06-2X	B3LYP	PM6	AMBER12: EHT	AMBER 94	AMBER 99	MMFF 94	CHARMM 27	OPLS -AA	MMFF94s
1	1.62	0.23	0.33	1.35	2.15	1.43	0.65	1.21	4.79	1.11
2	32.87	25.12	31.14	27.56	21.23	23.59	29.75	25.85	24.73	29.75
3	2.32	0.37	2.98	2.25	11.80	9.52	0.46	2.17	0.64	0.61
4	0.79	0.40	0.39	0.19	0.10	0.05	0.52	0.65	0.95	0.32
5	0.61	0.16	2.21	0.35	0.47	1.40	0.42	0.21	0.42	0.57
6	4.54	15.68	2.23	18.78	18.18	23.35	22.39	17.36	26.12	29.34
7	13.22	3.48	101.60	7.26	6.14	5.76	7.85	2.37	3.01	9.10
8	2.05	0.76	1.59	3.58	4.34	0.73	5.52	11.09	2.74	5.72
9	2.77	0.26	0.79	0.42	5.86	5.87	0.97	0.84	0.31	0.97
10	4.59	0.33	1.80	3.25	4.84	4.87	3.15	6.63	1.64	3.15

R^2 Values

Molecule	B3LYP	PM6	AMBER12: EHT	AMBER 94	AMBER 99	MMFF 94	CHARMM 27	OPLS-AA	MMFF94s
1	1.00	0.95	0.95	0.96	0.95	0.97	0.95	0.96	0.87
2	0.32	0.27	0.58	0.83	0.55	0.46	0.46	0.46	0.46
3	0.97	0.87	0.90	0.94	0.92	0.87	0.78	0.82	0.85
4	0.97	0.97	0.94	0.68	0.99	0.89	0.85	0.86	0.94
5	0.97	0.99	0.92	0.72	0.97	0.84	0.81	0.88	0.79
6	0.58	0.54	0.57	0.64	0.52	0.55	0.51	0.57	0.52
7	1.00	0.97	0.88	0.87	0.87	0.82	0.97	0.94	0.36
8	0.94	0.89	0.74	0.76	0.94	0.71	0.61	0.81	0.72
9	1.00	0.88	0.99	0.99	0.99	0.99	0.98	0.99	0.99
10	0.97	0.78	0.59	0.40	0.42	0.57	0.39	0.41	0.57

Supervision Record

4TH YEAR PROJECT SUPERVISION RECORD
2017 – 2018

Name of Candidate: John Brown

Name of Primary Supervisor: Richard Bryce

Name of Secondary Supervisor:

Title of Project: Medicinal Chemistry: Correctly Modelling the Shape of Drug Like Compounds

It is a requirement of the project that students should have regular meetings with their supervisor(s). In order to ensure a degree of consistency in the level of supervision offered across the School, a minimum of **four formal meetings** should take place and be documented by the student and supervisor(s) on this form. This form should then be submitted along with the final project report.

Meeting 1 Date: 04/10/17

The following aspects should be discussed with the student and summarised below:

Background to project
-Computational Drug Design
- Conformer Analysis
-Methods of Drug Modelling

Overall aims and objectives of project
- Compare different methods of molecular modelling
- Investigate the reliability and validity of molecular mechanical force fields
- Identify and find problem conformers

Discussion on research methodology and training
- Introduction to computer software
- Signing forms for the use of Gaussian

Discussion of project specific safety
- Risk assessment for filled out
- Attend risk assessment work shop for project room access

Aims for next meeting
-Identify any areas of uncertainty
-Discuss which molecules to use in sample
-Discuss any problems with computer software and data gathering

Student's signature: JBrown

Supervisor's signature: Richard Bryce

Meeting 2	Date: 29/11/17
Progress on the aims set at last meeting -Sample of molecules has been selected and data gathering has begun -Able to use computer software to obtain force field and quantum data -Identified areas to research for literature review Agreed aims for next meeting -Discuss data gathering progress -Obtain quantum data for half of the sample -Discuss any problems encountered	
Student's signature: JBrown Supervisor's signature: Richard Bryce	
Meeting 3	Date: 31/01/18
Progress on the aims set at last meeting - Quantum data has been gathered for 5 molecules - Data for force fields has started to be obtained - Problems with Gaussian have been discussed and rectified Agreed aims for next meeting - Gather remaining quantum data for the rest of the sample - Gather remaining force field data - Discuss data analysis and interpretation of results	
Student's signature: JBrown. Supervisor's signature: Richard Bryce	
Meeting 4	Date: 07/03/18
Progress on the aims set at last meeting -All force field and quantum data has been gathered -Data is being analysed and discussed how to interpret it -Literature review completed -Results chapter has been reviewed -Plan for discussion and conclusion has been made Agreed aims for next meeting -Discuss findings of data interpretation -Identify any problems -Finalise write up	
Student's signature: JBrown Supervisor's signature: Richard Bryce	

Risk Assessment

MANCHESTER 1824

of Manchester

School of Health Sciences
Division of Pharmacy and Optometry
Undergraduate Fourth Year Projects Risk Assessment Form

This form is to be completed and returned to Karen Purcell (Dispensing Laboratory. Tel. 52433) before the commencement of work on the project. The student should keep a copy of the form.

Name of student John Brown

Name of academic supervisor Richard Bryce

Laboratory/Area where project is carried out Stopford Building Project Room

The potential hazards of this project have been discussed in full by the supervisor and the student. The student understands these hazards and agrees to abide by the safety procedures set out in the School Safety Code and the instructions of the supervisor or person guiding his/her work.

Type of Project (*please tick as appropriate*)					
Laboratory		Survey/Interviewing		Computer Work	X

Read the relevant Standard Operating Procedure before completing this Risk Assessment Form

Title of Project

Medicinal Chemistry: Correctly Modelling the Shape of Drug Like Compounds

Brief Description of Project

Computer based project using various computer software to compare the performance of commonly used molecular modelling methods, for example, molecular mechanical force fields.

Risk Assessment

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
Work in an office environment	Slips, trips	Staff, Visitors, Cleaners - could suffer injury e.g. sprains or fractures if they fall	Reasonable standards of housekeeping maintained Trailing cables positioned neatly away from walkways Damage to floor coverings and other repairs and maintenance reported immediately to ..... for repair/replacement as necessary Floors kept clear of items, e.g. papers, bags Cabinet drawers and doors kept closed when not in use Floor cleaned regularly Adequate lighting provided	Low	A
	Manual Handling – carrying, lifting, pulling, pushing heavy loads e.g. furniture, PCs, stationary	Staff - could suffer from back pain if heavy/bulky objects carried incorrectly	Staff trained in correct manual handling techniques Untrained staff to contact ..... to arrange for items to be moved Trolley used to transport boxes of paper etc Top shelves used for storage of light objects only	Low	A
	Regular computer use	Staff - may suffer from upper limb disorders (associated with repetitive actions) from regular PC use or suffer from eyestrain/headache if lighting/screen image is poor	DSE Self Assessment of computer workstation undertaken within 4 weeks of starting employment and results forwarded to ..... All problems identified in the self assessment prompt a full assessment with the findings reported to ..... who will arrange for remedial action to be taken Staff appointed to undertake full DSE assessment as necessary Work scheduled so that staff have regular breaks from the computer Training in new software use provided as necessary	Low	A
	Electrical e.g. PC, printer, lamp, fan, heater, kettle, photocopier, shredder, extension leads	Staff and others - could suffer electrical shock or burns if equipment is faulty	All office equipment used in accordance with the manufacturers instructions and staff trained in its use by ..... Portable Appliance Testing (PAT) is carried out in accordance with the University Code of Practice on the Maintenance of Electrical Equipment Defective plugs, cables equipment etc reported to ..... for repair/replacement Sufficient power sockets provided to reduce need for extension cables Staff discouraged from bringing in own electrical equipment as maintenance cannot be assured Kettle/water heater positioned so water spills cannot contact electricity supply or equipment Liquid spills mopped up immediately	Low	A
	Fire	Staff and other building users - could suffer from smoke inhalation or burns if trapped in office	Staff induction includes fire evacuation procedures and means of raising the alarm Annual fire evacuation practice carried out Access to fire exits kept clear Regular removal of combustible waste Heaters located away from combustible materials and switched off when office is left unattended	Med	A
	Lone working	Staff - if presence not known in the event of an emergency or if there is a threat to personal security	Telephone contact available at all times Out of hours working arrangements in place to indicate the presence of people in the office Staff advised to ensure unauthorised persons do not gain access when using building out of hours Staff informed of how to contact Security	Low	A
	Falls from height	Staff - retrieving items stored at height	Stepladder/kick stool available from ..... Equipment checked annually for defects and by user before each use Staff trained in safe use of stepladder/kick stool	Low	A
	Stress	Staff - from pressure of work demands, lack of job control, insufficient support from colleagues, not knowing their role, poor relationships, or badly managed change	University Stress policy in place Work plans & objectives discussed and agreed at PDR annually or more frequently if need arises Self-referral to Occupational Health Service available	Low	A
	Environmental hazards a) thermal comfort	Staff – may feel too hot/cold or suffer other general discomfort	Building temperatures kept as reasonable as possible with supplementary heating/cooling available via Estates or School/Directorate by contacting .... when necessary Office is adequately ventilated	Low	A
	b) space	Staff and others – contact with furniture if insufficient space to move around	Space provided is sufficient to enable free movement around the office, and for carrying out tasks Individual space requirements are re-considered when additional equipment furniture is acquired	Low	A

Activity (8)	Hazard (9)	Who might be harmed and how (10)	Existing measures to control risk (11)	Risk rating (12)	Result (13)
	c) lighting	Staff and others – may suffer eyestrain if lighting is insufficient or of the wrong type	Lighting levels sufficient for the room, tasks undertaken and glare is minimised Window blinds fitted where necessary to adjust lighting levels Local lighting (e.g. lamps) is provided for close work where necessary Light switches are easily accessible	Low	A
	Hygiene & welfare	All staff & others could experience general discomfort	Toilets supplied with hot/cold water, soap and towels, any deficiencies are reported to ..... Refreshment area available with drinking water and other facilities which is cleaned daily by ..... No smoking policy implemented	Low	A
	Chemical e.g. photocopier toner/ ozone production, cleaning materials	Staff	Toner changed in accordance with manufacturers instructions. Disposable vinyl gloves worn when skin contact with toner is likely e.g. when changing cartridges Proprietary chemicals/substances used as directed by the manufacturer All spills are cleared up immediately and waste appropriately disposed of Heavily used photocopiers situated in well ventilated areas.	Low	A
	Anything else relevant to the local situation	etc	etc	etc	etc

I have read the relevant Standard Operating Procedure

Signature of student _____

Signature of supervisor _____ Date 08/05/2018