Ab initio and Force field investigations of physical hydrogen adsorption in Zeolitic Imidazole Frameworks

Huda Alqahtani

A thesis submitted to

Cardiff University

In accordance with the requirements for the degree of

Doctor of Philosophy

School of Chemistry

Cardiff University

2018

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……..…

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ……(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……(candidate) Date …….

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……..

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed …… (candidate) Date …….....

DEDICATION

This PhD research thesis is dedicated to all members of my family, especially to my: ***Mother: Aisha Al-athbah*** God rest her soul, my father: ***Aid Al-fehrah***, my husband: ***Mohammed Al-adbanee***, all my sisters: ***M****ona, **F****atima, **R****eem, **A****thbah and **H****anan*, my brothers: ***A****bdullah, **A****bdurrahman, **K****halid* and ***S****alem*, my sons: ***R****ayan, **A****bdullah, **A****lbaraa, **A****nas and **S****yhayb*, and my daughter: ***R****aseel.* Also to all my nieces and nephews, especially: ***Y****ahya, **R****aneem, **S****adeem, **S****humok* and ***J****umanah*, and to all people that made dua'a for me.

Acknowledgements

I would like to thank many people who encouraged me to complete this study. First of all, I would like to gratefully thank my supervisor, Prof. Peter Knowles, for his patience, enthusiasm, kindness, outstanding support, and for providing me the opportunity, the knowledge and the guidance to complete this project.

I will not forget King Saud University that gave me this opportunity to complete my high education and supported me financially and facilitated for me the procedures of the scholarship.

Moreover, I would like to thank Dr. James A. Plattes and Prof. Kenneth Harris for their advice, support, and encouragement during the project-process meeting.

Also, I would like to thank all academic staff of the Theoretical and Computational Chemistry at Cardiff University for their professional academic skills and their friendly personalities during my PhD studies, which encouraged me to be interesting on this field of science.

Last but not least I would like to thank my family (especially my husband Mohammed) for their endless emotional, moral, and financial support through all my studying years, and for believing in me and giving me the trust to be what I am today.

Abstract

Recent theoretical calculations and experiments have considered that metal-organic frameworks are promising for storing molecular hydrogen (H_2). Optimizing the geometry and the interaction energy of storing for enormous H_2 storage is of great current interest. In this work, we used specific category of MOFs, Zeolitic Imidazole Frameworks (ZIFs). We carried out calculations through high-accuracy electronic structure calculations (MP2, CCSD and CCSD(T)) levels of theory, with controlled errors. Also we established and calibrated a computational protocol for accurately predicting the binding energy and structure of weakly bound complexes. Then, we applied the protocol to a number of models for metal-organic frameworks. For example, we have built many systems of noncovalently bound complexes [H_2…benzene, H_2…. imidazole, CO…. imidazole, N_2… imidazole, NH_3…imidazole and H_2O …imidazole] and we have optimized geometries of these systems through calculating numerical gradients at MP2/CP level and LMP2 level of theory and extrapolated from aug-cc-PVTZ and aug-cc-PVQZ basis set to evaluate the binding energy by using Hobza's scheme to obtain correct interaction energies. We found that NH_3 and H_2O with imidazole prefer to form hydrogen bonds rather than physical adsorption (London dispersion force). Also, the perpendicular position of hydrogen has the lowest potential energy surface, while the parallel hydrogen position has the highest potential energy surface. We have confirmed that by using a high level of basis set at MP2 such as cc-pVXZ (x= Q, 5, 6) and aug-cc-pVXZ (x=D, T, Q, 5, 6), and by using the same basis sets at CCSD and CCSD(T) as the high level of theory. Also, it is clear from these results that the binding energies are sensitive to improvement of the size of basis sets. In terms of applying Hobza's scheme to obtain correct interaction energies, we found that this scheme CCSD(T)/ [34] = MP2/ [34] + (CCSD(T)/ [23] – MP2 [23]) achieved the highest accurate of interaction energy for CO...imidazole. On the other hand, this scheme CCSD(T)/ [34] = MP2/ [34] + [CCSD(T)/AVDZ– MP2/AVDZ] produced the highest accurate of interaction energy for H_2...imi, N_2…imi and H_2…Benzene. Regarding to Basis Set Superposition Error (BSSE) and counterpoise examination (CP),

we found that the MP2/CP and LMP2 methods yield very similar results at the basis set limit and the convergence of MP2 and LMP2 with increasing size of basis sets is different since the BSSE in LMP2 is reduced. Furthermore, we found that the extrapolation to the CBS limit cannot offer an alternative to the counterpoise correction where the differences in the values of bending energies are large so we need to use both techniques together to overcome the BSSE problem.

Then to confirm our result regard to the potential energy surface, we calculated corresponding potential energy surfaces using several popular force fields potential, and compare critically with best ab initio results, where we focused on the adsorption of H_2 on imidazole as the organic linker in ZIFs. We carried out ab initio calculations at the MP2/CCSD(T) levels with different basis sets, basis set extrapolation and Lennard-Jones potential for the three directions X, Y and Z for 294 positions of H_2. Also, we have fitted ab initio binding energy at the MP2/CCSD(T) levels with different basis set and basis set extrapolation to Lennard-Jones (12-6 LJ) binding energy by applying the nonlinear least squares method. Then we estimated the fitted binding energy using Hobza's schemes to reduce the errors. We found that the 12-6 LJ formula produced unreasonable fit for ab initio calculated potential energy surface PES, for both the equilibrium and attractive regions, to improve this fitting, we found the good fit is only achieved by the exponential formula of repulsion region.

It is hoped that this study could facilitate the search for a "good" application to store the H_2 molecule conveniently and safely.

Contents:

Chapter 4: Application of Potential energy surface

Chapter 5: General conclusions

Appendix

List of tables:

List of figures:

List of Acronyms

AVXZ	Augmented Correlation Consistent Polarised Valence X-tuple zeta (basis)
AO	Atomic Orbital
B3LYP	Becke-Lee-Yang-Parr (density-functional theory)
BSSE	Basis-set superposition error
CCSD	Coupled Cluster Singles and Doubles
CCSD(T)	Coupled Cluster with Single, Double, and Triple replacement
cc-pVXZ	Correlation Consistent Polarised Valence X-tuple zeta (basis)
CI	Configuration Interaction
CISD	Configuration Interaction Singles and Doubles
CP	Counterpoise Method
DFT	Density Functional Theory
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
12-6LJ	Lennard-Jones Potential

LUMO	Lowest Unoccupied Molecular Orbital
MO	Molecular Orbital
MP2	Møller r-Plesset second-order perturbation Theory
MPPT	Møller -Plesset perturbation Theory
PES	Møller -Plesset perturbation Theory
RHF	Restricted Hartree-Fock
SCF	Self-Consistent Field
UHF	Unrestricted open-shell Hartree-Fock Method

Outline of thesis

This thesis is divided into five chapters. Chapter 1 describes previous work and gives motivation for the work performed in this thesis. Chapter 2 describes the methodology used in the design of the calculations and the approach taken and the mathematical tools used in the analysis. Chapter 3 presents the theoretical results of adsorption of H_2 with imidazole and applies the established computational protocol to another system. Chapter 4 describes and discusses the application of potential energy surfaces. Finally, chapter 5 summarizes the main conclusion of the thesis and outlook for future work.

Chapter 1

Introduction

1 Introduction

1.1 Aim and scope of this thesis

The economic, efficient and safe storage of hydrogen presents significant challenges to making hydrogen a viable energy source.

Thus far, fossil fuels have been the primary source of energy, and this dependency has resulted in serious energy crises and environmental pollution on a global scale[1-4]. Each year in Hong Kong, an estimated total of 290,000 terajoules are consumed. Vehicle transportation accounts for 36% of the fossil fuels expended and generating electricity is estimated to be 48%[5]. As a result of fossil fuel combustion, approximately 40 million tons of toxic and greenhouse gases are emitted annually[6].

The environmental harm caused by fossil fuels demands the urgent creation and implementation of clean and renewable energy technologies[7]. Potential renewable sources of energy include biomass, hydrothermal, solar, tidal and wind[8-13]. Biomass, solar and wind energy are especially abundant; should their technological developments reduce the costs to competitive levels, they are potentially able to dominate the future energy market. Solar and wind energy are unreliable sources with variable output as they are intermittent and site specific. Whilst batteries do offer a mechanism to store energy, they do have many drawbacks including creating a considerable amount of waste, low storage capacity and the equipment has a short life span. In view of these shortcomings amongst various alternatives, hydrogen offers the potential of being a suitable fuel and an energy carrier for the energy industries of the future. Through electrolysis driven by solar cells or wind turbines, hydrogen can readily be produced from water[4]. Water is the only by-product that arises from the conversion of hydrogen into electricity through a fuel cell, making hydrogen environmentally friendly in terms of its life cycle.

In our project, we shall investigate - through high-accuracy electronic structure calculations - the adsorption of small molecules on organic fragments, with a view to understanding how to carry out calculations of the properties of larger systems, such as metal-organic frameworks, with controlled errors. We will also establish and calibrate a

computational protocol for accurately predicting the binding energy and structure of weakly bound complexes. Then, we will apply the protocol to a number of models for metal-organic frameworks.

1.2 Metal Organic Frameworks (MOFs)

Metal-Organic Frameworks (MOFs) are considered as a novel class of porous crystalline materials, which in the last decade have seen a significant growth and are expected to have enormous effect on the development of scientific future technologies. 1965 is the commonly assumed birth year of MOFs. Some researchers use the terms "Coordination Polymer" or "metal-ligand coordination polymers" instead of metal organic frameworks to describe this material as Férey[14-17]. Also, there are many other names that are also in usage for MOFs such as porous material by Kitagawa[18] , metal-organic materials (MOMs) by Perry[19], and hybrid organic inorganic materials or organic zeolite analogues[17].

The term MOF is used throughout this thesis to describe the extended structure based on metals and organic linking ligands.

Metal organic frameworks (MOFs) are of significant interest in the formation of porous materials with an extensive selection of functions. There are many studies in this field. Where, the number of studies which feature the term MOF has increased dramatically over the last decade, and an enormous selection of building blocks have been used for producing tuneable pore sizes and various architectures. In fact, more than 3,000 MOFs have been discovered so far[20].

Metal-Organic Frameworks are combinations composed of cluster or metal ions coordinated to mostly rigid organic molecules to form one, two, or three-dimensional structures that seem porous, see (Figure 1-1)[21-23]. These porous solids consist of metals such as chromium, zinc, copper, zirconium, aluminum and other elements linked together by organic linker chemicals (e.g. imidazolates, carboxylates and tetrazolates) to form networks of empty pores almost like the pores in a kitchen sponge and to produce strong bonds between organic and inorganic fragments. However, over 90 percent of a

MOF is composed of empty space that might be filled with carbon dioxide or hydrogen [20]. The following diagram shows how MOFs are formed (Figure 1-1).

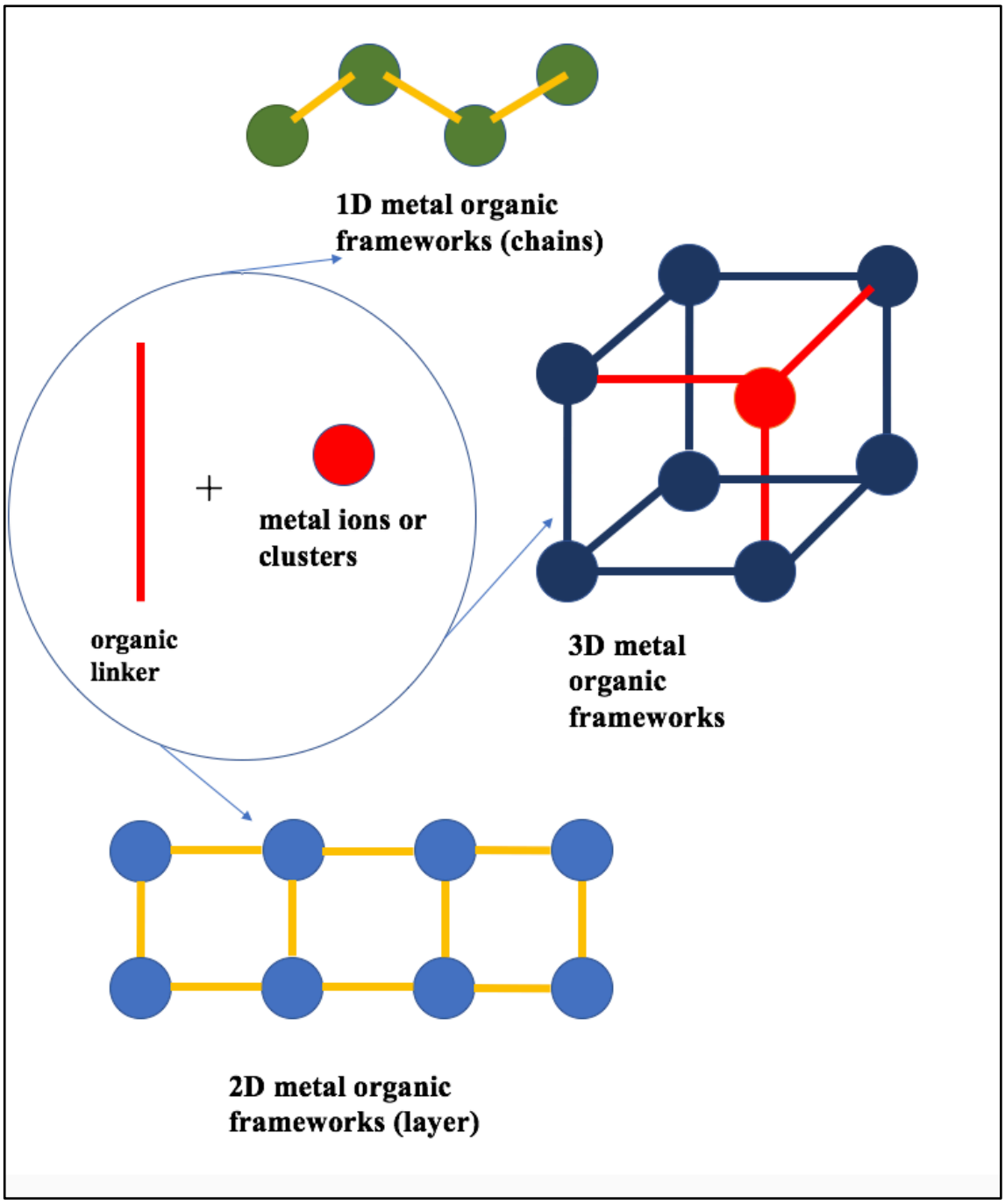


Fig. 1-1: Illustration of the formation blocks and structure of MOFs with different dimensionalities (1D, 2D and 3D)[24, 25]

Figure 1-1 shows how the MOFs build where metal ions act as a Lewis acid. Recently, MOFs have attracted a lot of attention because of their remarkable properties[26], and

their multiple applications in catalysis[27-35], magnetism, electrochemical storage[36-43], separation[22, 44-46], as hydrogen storage[20-23, 47], and CO_2 capture[44, 48-51], ion exchange[52, 53], polymerization[54], drug delivery[55] and sensing[56-59]. In all these applications, the interaction between the small guest molecules and the walls of the MOFs are controlled by the van der Waals forces[60]. Also, the structures of MOFs can be designed according to aimed properties[19, 21, 61], which have governed them by setting the geometries of the organic linkers and coordination modes of the inorganic metal ions or clusters of metal ions[62].

MOFs can exhibit properties that reflect the components, therefore, may display chemical functionality, chirality and geometric rigidity. Simple preparations of MOFs are typically high yielding and scalable with careful selection of the building block, and a level of design can be generated to produce specific products from the great number of potentially accessible MOF, determining the geometric requirements of a target framework, the process of executing the framework's design and synthesis to produce an ordered structure, are termed reticular synthesis. To achieve this, it is necessary to understand the local coordination patterns of both the organic and metal components and an ability to predict the net topologies that they will take on. MOFs can be made up from numerous varieties of chemical moieties that can be differentiated using crystallography; these offer great potential in determining the structural factors that most effectively adsorb hydrogen.

Indeed, there are many areas of interest that are related to metal-organic- frameworks (MOFs), but we shall focus on the adsorption of small molecules (on MOFs). In fact, there are many studies on the adsorption of small molecules on metals, adsorption through open-site frameworks, in addition to more theoretical studies that include characterization of coordination space in adsorption and the computational approaches of adsorption on MOFs.

In fact, the reduction in fossil fuel reserves immediately needs solutions of replacement and MOFs might be one, for their capacity to adsorb large amounts of gases like H_2, CH_4, C_2H_2, O_2, CO, CO_2, NO_x, etc. within the pores.

In MOFs, pore size and shape have a significant role in the adsorption of guest molecules onto the solid surface where it is not only limited to the interaction between guest molecules and the surfaces. Currently, the main efforts focus on H_2, CO_2, and CH_4 with a marked difference between hydrogen molecule and the others. In fact, at 77 K, MOFs adsorb large amounts of hydrogen. While, adsorption is negligible to CO_2 and CH_4, which exhibit interesting performances at 300 K and above. This low temperature adsorption of H_2 prevented for a long time, applications for its use in cars. Nowadays, technical solutions are in progress [63, 64].

To start with, Lee and coworkers to investigate small-molecule adsorption in open-site metal-organic frameworks conducted one interesting study. Special attention was paid to the systematic density functional theory of rational design[65].

The authors utilized the density functional theory, in this detailed study, to compute the binding enthalpies of fourteen various molecules including M-MOF-74 with M = Mg, Ti, V, Cr, Zn, Cu, Fe and Ni, the small molecules considered include trace gases, major flue-gas components, and small hydrocarbons for example, H_2, CO_2, CO, H_2S, H_2O, N_2, NH_3, SO_2, CH_4, C_2H_6, C_3H_8, C_2H_4, C_3H_6 and C_2H_2. Overall, the adsorption energetics of 140 systems were examined and discussed.

Initially, the adsorption energies of various systems were analyzed and outlined for further discussion. Based on the results, and as predicted, it was found that the Cu-MOF-74 tends to lend selectivity of CO_2 over H_2O to results in the separation of trace flue-gas impurities and other unwanted gas mixtures. This study comes in frame of investigation that MOFs are ideal adsorbents for gas separation.

Another study by Soares and coworkers where they examined the adsorption of small molecules in porous zirconium-based metal organic frameworks (MIL-140A), and presented a joint computational experimental approach[66].

The study covered small molecules namely H_2, CH_4, N_2, CO and CO_2. Those molecules were explored in small pore zirconium terephthalate MOF by combining force field based molecular simulations and quantum, and experiments. Density-functional theory

was used, giving very good agreement between the experimental spectroscopic (IR, NMR) and the predicted results.

The structural features of the given MOF configuration were also studied carefully. As found, the preferential adsorption sites and strength remain confirmed to the host molecules and that was confirmed with field-based Monte Carlo simulation. This means that quantum calculations help to give a clear understanding for expectation of the special adsorption sites and the strength of the interactions for all restricted molecules between host and guest in these systems. Also, furthermore, the stability of the hybrid porous material was examined in the light of its interaction with the given MOF, where the water stability of this hybrid porous solid was similarly studied as well as the interaction between the MOF and gas toxin, that is, H_2S..

Mavrandonakis and coworkers[67], studied the adsorption of small molecules on metal-organic frameworks, just like the previous study, but also included oxo-center Trimetallic $M_3^{III}(\mu_3\text{-}O)(X)\ (COO)_6$ units. Hence, the study aimed at investigating the role of under-coordinated metal ions.

The authors, uniquely, conducted combinatorial computational screening on a diverse variety of units including Al^{3+}, Rh^{3+} and Ir^{3+} in addition to several others. The screening process resulted in considering parameters including interaction energy, vibrational properties and adsorption enthalpies. It was then found that the binding energies of such small molecules are important to the process but also can be tuned through the adjustment of metal composition, though Rh and Ir were found to be the best candidates for use in H_2 or gas separation.

Furthermore, Nagaoka and coworkers carried out a theoretical study that included the characterization of coordination space studied the adsorption state and behavior of small molecules in nano-channeled metal-organic frameworks[68].

The authors claim that this area of study, namely those related to coordination space, have attracted the attention of recent scholars in the area. The authors made estimates of electronic state properties of various molecules and studied the role of MOFs in adsorption. Hybrid electronic state and molecular mechanical methods were used, along

with Monte Carlo simulations. Furthermore, various MOF combinations were synthesized, and the maximum entropy methods were determined. Also, density functional theory was also employed here. The conclusion made is that the impact of electronic potential on hydrogen adsorption in MOFs can be very significant. Additionally, since the nuclear quantum effect is also important for hydrogen adsorption at low temperature, some realistic prescription applicable to multi-scale phenomena with huge degrees of freedom must be established immediately.

Yet another interesting publication presented various computational approaches to examine the adsorption in MOFs, especially with unsaturated metal sites. The paper studies the potential of using various metals to tune the affinity of adsorbents, for certain applications[69].

Metal organic frameworks, given specified coordinated unsaturated sites could offer such opportunity to be used as a tuning material. However, conventional force fields have only limited ability to deal with such systems, and so in order to model faithfully, a mixed method that takes into account QM-based data with various classical models may offer better understanding of adsorption in relationship to conventional force fields.

In this work, we shall concentrate on Zeolitic Imidazole Frameworks (ZIFs) as a specific class of MOFs (adsorbent) and the hydrogen molecule as a small molecule (adsorbate) from a theoretical side, where we shall investigate this adsorption through high accuracy electronic structure calculations.

1.3 Zeolitic Imidazole Frameworks (ZIFs)

Zeolitic Imidazole Frameworks (ZIFs) are a new category of metal-organic frameworks[70, 71], which are linked topologically with zeolites. In this category, we found that there are connections between two transition metal ions (Zn, Co, In) with organic imidazole linkers through N atoms forming a tetrahedral shape because the angle (metal-imidazole-metal) equals 145º and this value of angle is the same as the Si-O-Si angle in zeolites (Fig. 1-2).

The crystal structure of ZIFs is based on aluminosilicate zeolite nets (Fig.1-3), where the tetrahedral $Si(Al)O_4$ sites are replaced by transition metals M (M = Zn(II), Co(II), In(III)) tetrahedrally coordinated by imidazolate ligands, and that is the reason for the term" zeolitic" in the name of ZIFs, because the structures of zeolites and ZIFs are very similar[70].

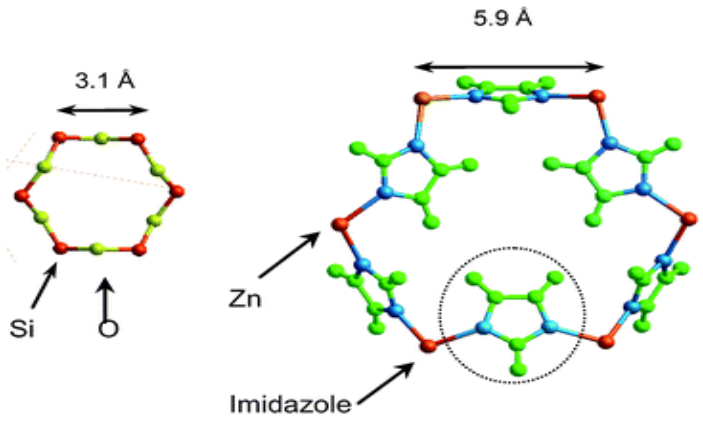


Fig. 1-2: Zeolites versus ZIFs: the sequence of Si–O–Si bonds in zeolites (left) and of Zn-IM-Zn in ZIFs (right) is illustrated for a 6-membered ring[72].

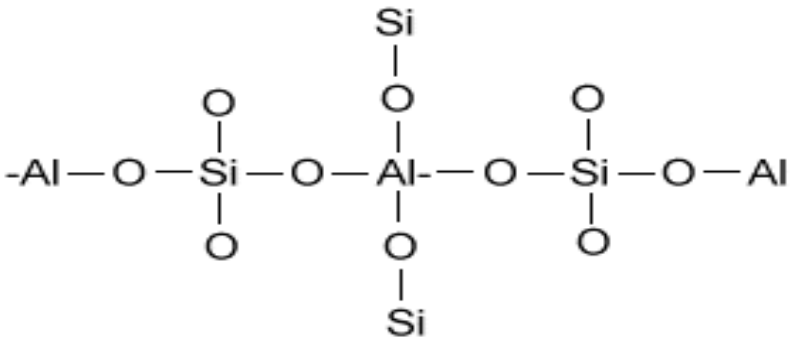


Fig. 1-3: The aluminosilicate zeolite nets, whereby the tetrahedral Si(Al) sites are replaced by transition metals M (M = Zn(II), Co(II), In(III))[73].

ZIFs are considered thermally and chemically stable materials[70], so when ZIFs are heated to high temperatures they do not decompose and can be boiled in water or any solvent for a week and remain stable. This makes them suitable for use in a hot environment to produce the energy-producing environment as power plants. In addition, Zeolitic Imidazole Framework membranes have a highly porous structure, large pore volume with fully exposed edges and the faces of the organic linkers, so ZIFs are suitable for the separation of gases process as one of the applications of ZIFs[74]. Besides the separation of gases process there are many applications for which ZIFs are important, such as carbon capture and storage. ZIFs are scrubbers for carbon dioxide where one litre of the crystals could store about 83 litres of CO_2. Also, the crystals are not toxic and they require little energy to make[75].

1.4 An overview of imidazole

In the present project, the imidazole is the organic linker in the ZIFs as already mentioned above so it is worth to present some details about it. The structure of imidazole (Im) is a five-membered planar ring, bearing 3 carbons, 2 nitrogens and 4 hydrogen atoms. Because hydrogen can be located at either of the nitrogen atoms, Im has two tautomeric forms. The nitrogen atoms function as proton donors and acceptors, enabling the molecule to form strong and directional H-bond interactions. Theoreticians and researchers alike are attracted by this unique compound and its qualities[76-87]. Imidazole's ring structure is exceptional in that either of the N atoms can pick up

protons to form cations and transport hydrogen from the other N atom in a proton transfer process (Fig. 1-4). As a result of this ability, imidazole is fundamental to a number of biological processes and it is significant in the structure of proteins and the function of enzymes[76-78]. Together with some of its derivatives, Im forms a group of general base and nucleophilic catalysts[80, 86]. It also acts as a ligand to the cobalt metal ions in B12 coenzyme[81]. The structures and relative energies of H-bonded imidazole molecules have been explored in various experimental studies[82, 83]. By studying infinite chain structures, it has been possible to calculate the electrochemical reduction potential of imidazole (-1.77 V) and its monoprotonated form in acetonitrile (-0.73 V). Studies have also demonstrated that in the presence of tetrafluoroborate ($TBABF_4$), the electrochemical reduction potential of n dimer-Im (n is the number of dimer of imidazole), is reducible with 2n electrons through 2n anions of Im (Im-) and gaseous nH_2 (Fig. 1-5). The peak can be reduced to a more positive potential, rising to -0.73 V[84], if imidazole is protonated with fluoroboric acid (HBF_4). The electroreducibility and H-bond abilities of aminobenzimidazoles with tetrafluoroborate have been studied; the protonated forms of the aminobenzimidazoles were displayed by the addition of fluoroboric acid. In oligomers, reduction arises through 2n electrons and in protonated structures, 2 electrons, releasing hydrogen gas[85]. Theoretical calculations have been used to explore hydrogen bonding of Im structures. Using *ab initio* techniques[86], Bredas and coworkers, investigated monomer, dimer and infinite chain structures of imidazole. Tafazzoli and Amini studied hydrogen bonding on imidazole and 4-nitroimidazole through calculating the chemical shielding by *ab initio* and density functional theory (DFT) at chemical shielding of ^{13}C and ^{15}N nuclei of imidazole, its 4-nitro derivative and gauge invariant atomic orbital (GIAO) condition at different temperatures. The accuracy of calculated results is improved because of the use of structures determined by neutron and X-ray diffraction methods. The results indicate that the DFT method is more reliable than the Hartree-Fock (HF) method [87, 88].

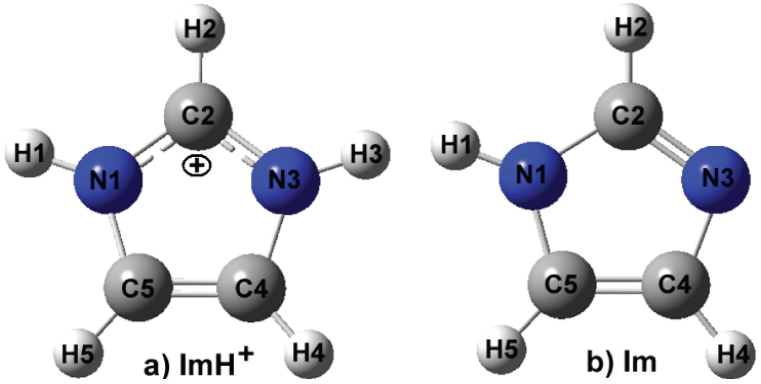


Fig. 1-4: Molecular geometries of (a) monoprotonated cation and (b) neutral structures of imidazole[88].

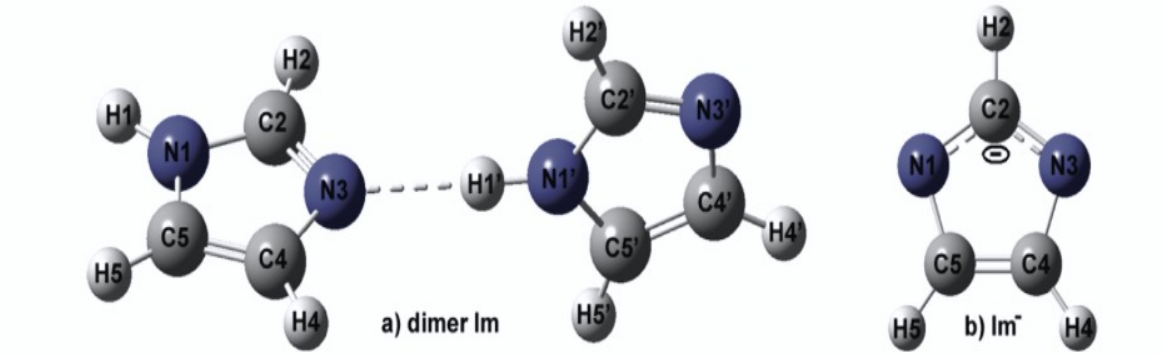


Fig. 1-5: Molecular geometries of (a) dimer and (b) anion (Im^-) structures of imidazole[88].

1.5 An overview of H_2 storage

Hydrogen storage materials can be categorised into two categories. The first involves physisorption of hydrogen molecules on support surfaces. The second is characterised by dissociation of the chemical binding as hydrides and hydrogen atoms[12, 89]. Indeed, the significance of the hydrogen economy has been acknowledged in several countries around the world, including Australia, China, Germany, Japan, Turkey, UK and USA[4]. However, to achieve hydrogen economy, the problem of storing hydrogen economically, efficiently and safely must be answered. Currently, there are eight possible hydrogen storage technologies: 1) high-pressure gas compression, 2) liquefaction, 3) metal hydride storage, 4) carbon nanotube adsorption, (5) metal clusters, (6) complex chemical hybrids, (7) metal organic frameworks, (8) nanostructure metals[4].

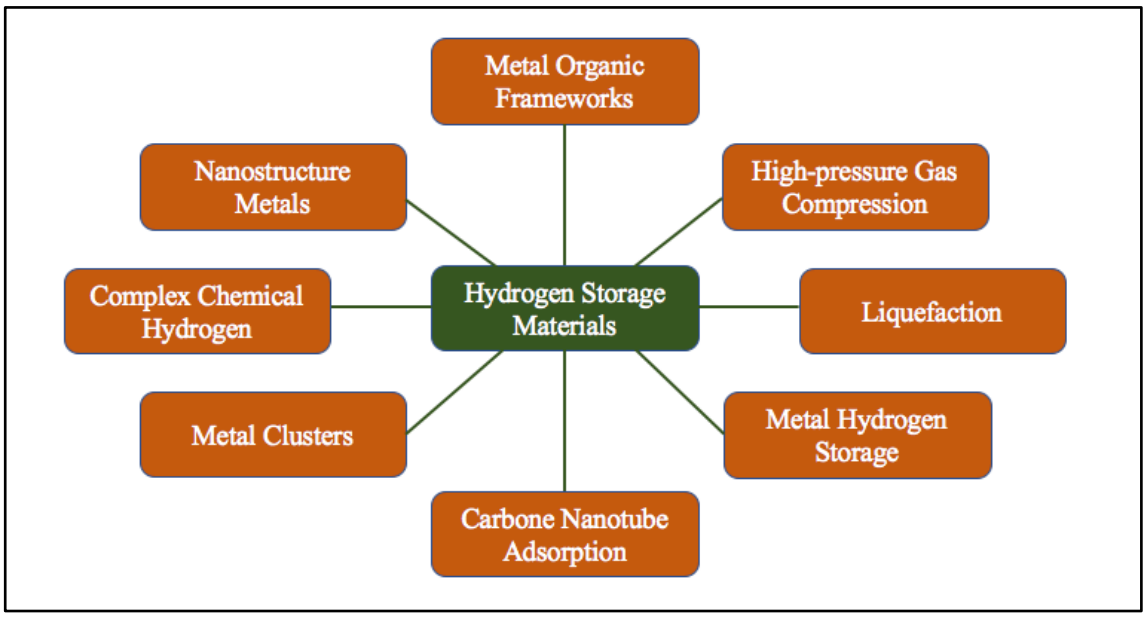


Fig 1-6: Summary of various hydrogen storage materials[4].

In the current project, we shall study MOFs as hydrogen storage materials. MOFs are crystalline compounds comprised of metal ions or clusters that are linked through molecular bridges as mentioned above. The vital problems for hydrogen adsorption in MOFs are both the number and surface of adsorption sites, and the strength of adsorption interaction (adsorption energy). The hydrogen adsorption energies on MOFs are such that adsorption occurs below 80K, but not at room temperature because an extremely high pressure is needed to make the adsorption possible[90]. To store the hydrogen at room temperature and reasonable pressures, the adsorption energy should be about 30 – 40 kJ/mol, according to the estimations of Li and coworkers[91]. In terms of the estimations of another author, the adsorption energy should be about 15 kJ/mol[92] according to Myers and coworkers, and 20 kJ/mol according to Snurr and coworkers[93].

Several strategies have been recommended in order to improve the adsorption energy of hydrogen such as incorporation of transition metals[94, 95]. Another method to develop the storage capacity is reducing the framework density by using light metal ions. One expects that the large polarizing power can give rise to relatively strong coordination bonds[38]. Many attempts have been made to find a means of storing hydrogen and

transporting it conveniently and safely. However, no method offers both perfect reliability and an isolated particle process.

To further the understanding of physical adsorption as it applies to the hydrogen dynamics of MOFs, computer models are valuable to study the interactions of hydrogen and metal-organic structures. With this knowledge, new avenues of material optimization may become apparent when we present some studies including papers that include the use of hydrogen storage in metal-organic frameworks, adsorption sites of hydrogen in Zeolitic Imidazolate Frameworks, diffusion of hydrogen guest molecules, and other areas of interest to the research in this area.

First, Murray and his group, (2009) investigated hydrogen storage in metal-organic frameworks. The author recognized a need for hydrogen storage to be used in objects that use hydrogen as a clean energy, particularly automobiles[43].

Therefore, the authors point out the need to look for new materials that can store hydrogen safely and adequately. In this regard, microporous metal-organic frameworks can serve as an appropriate potential for hydrogen storage. To investigate this potential, the scholars in this project investigate the relationships between enthalpy of hydrogen adsorption and structural features, and come up with proposals to improve storage capacity.

According to the study, metal-organic frameworks can exhibit great performance for hydrogen storage to a temperature reaching 77K and 100 bar pressures, which are great for such practical applications, though may have challenges at temperature of 298K. This project also pointed out that the cost aspects as it is important to produce affordable storage material and hence the proposal of using metal-organic frameworks.

Another paper by Suh and her colleagues (2012) reviewed various versatile structures with high surface areas and pore volumes. The ultimate goal is to find the most adequate and usable material for use as hydrogen storage[42].
The focus was on those MOFs that exhibit great adsorption capacities at temperatures between 77 and 87K. The H_2 adsorption in MOFs however, contributed to lower capacities with increase in room temperature.

The scholars indicated that in order to increase the surface area of MOFs, which is important for storage applications, poly-carboxylate ligands could be used as organic building blocks. Also, the used framework should exhibit high H_2 adsorption enthalpy to allow for sufficient storage ability. Also, to create improved H_2 storage at room temperature, for instance, occurs with increase isostatic heat of H_2 adsorption, which can be achieved through developing open metal sites.

Another study (Prakash and coworkers, 2013) investigated H_2, N_2 and CH_4 adsorption in Zeolitic Imidazolate Frameworks (both ZIF-95 and ZIF-100)[96]. To proceed with this research, Monte Carlo simulations and ab initio were used to analyze the uptake behaviors of the above listed molecules. Some of the findings regarding these gases include the fact that uptake capacity for ZIF-100 outperforms that of ZIF-95 for H_2 molecules though the contrary is true for CH_4.

Also, the research results highlight the gravimetric adsorption density for the gases at temperature 77 and 300K, which could be significant for use in certain applications. Also, the isosteric heat of adsorption for the gases was determined for the same temperature range.

Zhou and coworkers (2009) studied adsorption sites of hydrogen in Zeolitic Imidazolate Frameworks as well. Monte Carlo simulations were also employed to investigate the hydrogen adsorption in imidazolate frameworks, thoroughly[97].

The simulations were carried out at 77K and pressure ranging from 10 to 8000 kPa. To achieve the simulation results, an OPLS-AA force field was employed and a Buch model was applied for hydrogen molecules.

Furthermore, computer tomography for materials was used to explore the ZIF-8 materials. The results indicated that the first adsorption sites occurred at the imidazolate C=C bond and that hydrogen molecules were adsorbed through a pore channel. The scholars recommend the development of a new ZIF material for better adsorption competence upon analytically comparing MOFs and ZIF materials. Additionally, in 2007 Hirscher and Panella wrote a viewpoint paper about hydrogen storage in metal-organic frameworks and they pointed out that the physisorption is the base in the

hydrogen storage and that the adsorption process possesses really fast kinetics and is adjustable. This advantage makes physisorption a promising mechanism for hydrogen storage[98]. They also noted that at low temperature physisorption is a possible alternative to chemical hydrogen storage in complex hydrides or metal hydrides where the nanoporous metal-benzenedicarboxylate $M(OH)(O_2C\text{-}C_6H_4\text{-}CO_2)$ ($M = Al^{3+}$, Cr^{3+}) absorbs one molecule of water at room temperature, and can be promptly removed by heating until above of 100 °C. The sample handle was plunged into liquid nitrogen so as to cool it down to 77 K. At this temperature, they observed that there is significant adsorption. It is worth mentioning that a small amount of heat is created in molecular hydrogen adsorption[99].

1.6 The interaction between H_2 and Zeolitic Imidazole Frameworks (ZIF)

The nature of the interaction between H_2 and ZIFs is a kind of adsorption that is based on van der Waals forces (London dispersion force)[71], so called physical adsorption (physisorption). On the other hand, there are another type of adsorption and it is completely different from physical adsorption and namely chemisorption where the strong force between adsorbate (small molecule) and adsorbent (solid surface) leads to chemisorption adsorption and a weak force (van der Waals) leads to physical adsorption as shown in Figure 1-7.

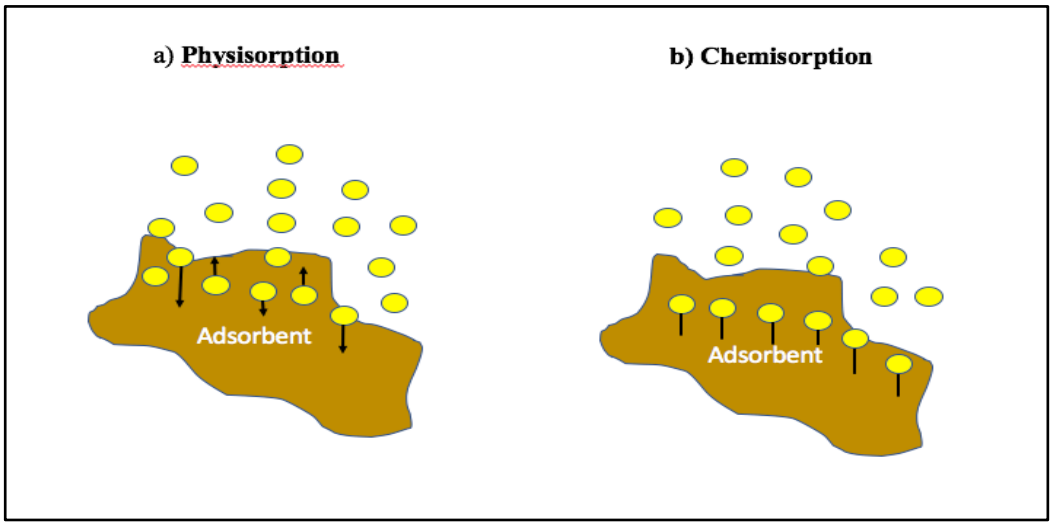


Fig.1-7: a) physisorption adsorption. b) chemisorption adsorption, where the strong and the weak bonds are formed.

Indeed, both types of adsorption are described by the Lennard-Jones potential [100] (as we shall explain in chapter 2 in this thesis), where the interaction or adsorption is presented by a potential curve for the small molecule and the solid surface.

In our research, we shall focus on where physisorption is an exothermic procedure, which means that heat is released by the formation of new bonds. Also, the entropy of the system is declined because of the loss of at least one degree of translational freedom of the molecules. Physical adsorption is usually described by low adsorption enthalpy values, which mean that it is straightforward to release H_2 when needed.

1.7 References

1. Balat M, *Energy Exploration & Exploitation*, 2005, **23**, 141.
2. Demirbas AF, *Energy Exploration & Exploitation*, 2004, **22**, 231.
3. Kaygusuz K, *Energy Exploration & Exploitation*, 2004, **22**, 145.
4. Ni M, *Energy Exploration & Exploitation*, 2006, **24**, 197.
5. Ni M, Leung MKH, Sumathy K and Leung DYC, *International Journal of Hydrogen Energy*, 2006, **31**, 1401.
6. Ni M, Sumathy K, Leung MKH and Leung DYC, *Proceedings of the International Conference on Alternative Energy*, 2005, 171.
7. Kaygusuz K, *Energy Exploration & Exploitation*, 2001, **19**, 603.
8. Vamvuka D and Tsoutsos TD, *Energy Exploration & Exploitation*, 2002, **20**, 113.
9. Demirbas A, *Energy Exploration & Exploitation*, 2002, **20**, 105.
10. Demirbas A, *Energy Exploration & Exploitation*, 2002, **20**, 325.
11. Demirbas A, *Energy Exploration & Exploitation*, 2003, **21**, 269.
12. Balat H, *Energy Exploration & Exploitation*, 2005, **23**, 61.
13. Balat M, *Energy Exploration & Exploitation*, 2005, **23**, 41.
14. Goldberg I, *Chemical Communications*, 2005, **10**, 1243.
15. Cheetham AK, Rao CNR and Feller RK, *Chemical Communications*, 2006, **46**, 4780.
16. Alhamami M, Doan H and Cheng CH, *Materials*, 2014, **7**, 3198.
17. Ferey G, *Chemical Society Reviews*, 2008, **37**, 191.
18. Kitagawa S, Kitaura R and Noro S, *Angewandte Chemie-International Edition*, 2004, **43**, 2334.
19. Perry JJ, Perman JA and Zaworotko MJ, *Chemical Society Reviews*, 2009, **38**, 1400.
20. Zhou HC, Long JR and Yaghi OM, *Chemical Reviews*, 2012, **112**, 673.
21. Tranchemontagne DJ, Mendoza-Cortes JL, O'Keeffe M and Yaghi OM, *Chemical Society Reviews*, 2009, **38**, 1257.
22. Li JR, Kuppler RJ and Zhou HC, *Chemical Society Reviews*, 2009, **38**, 1477.
23. Hu YH and Zhang L, *Advanced Materials*, 2010, **22**, E117.

24. Dhakshinamoorthy A and Garcia H, *Chemical Society Reviews*, 2014, **43**, 5750.
25. Mendes RF and Paz FAA, *Inorganic Chemistry Frontiers*, 2015, **2**, 495.
26. Long JR and Yaghi OM, *Chemical Society Reviews*, 2009, **38**, 1213.
27. Farah OK, Shultz AM, Sarjeant AA, Nguyen ST and Hupp JT, *Journal of the American Chemical Society*, 2011, **133**, 5652.
28. Lee J, Farha OK, Roberts J, Scheidt KA, Nguyen ST and Hupp JT, *Chemical Society Reviews*, 2009, **38**, 1450.
29. Shultz AM, Farha OK, Hupp JT and Nguyen ST, *Journal of the American Chemical Society*, 2009, **131**, 4204.
30. Wood CD, Tan B, Trewin A, Niu HJ, Bradshaw D, Rosseinsky MJ, Khimyak YZ, Campbell NL, Kirk R, Stockel E and Cooper AI, *Chemistry of Materials*, 2007, **19**, 2034.
31. Wood CD, Tan B, Trewin A, Su F, Rosseinsky MJ, Bradshaw D, Sun Y, Zhou L and Cooper AI, *Advanced Materials*, 2008, **20**, 1916.
32. Horike S, Dinca M, Tamaki K and Long JR, *Journal of the American Chemical Society*, 2008, **130**, 5854.
33. Lin WB, *Journal of Solid State Chemistry*, 2005, **178**, 2486.
34. Zou RQ, Sakurai H, Han S, Zhong RQ and Xu Q, *Journal of the American Chemical Society*, 2007, **129**, 8402.
35. Alaerts L, Seguin E, Poelman H, Thibault-Starzyk F, Jacobs PA and De Vos DE, *Chemistry-a European Journal*, 2006, **12**, 7353.
36. Yaghi OM, Davis CE, Li GM and Li HL, *Journal of the American Chemical Society*, 1997, **119**, 2861.
37. Ward MD, *Science*, 2003, **300**, 1104.
38. Rowsell JLC and Yaghi OM, *Angewandte Chemie-International Edition*, 2005, **44**, 4670.
39. Collins DJ and Zhou HC, *Journal of Materials Chemistry*, 2007, **17**, 3154.
40. Morris RE and Wheatley PS, *Angewandte Chemie-International Edition*, 2008, **47**, 4966.
41. Dinca M and Long JR, *Angewandte Chemie-International Edition*, 2008, **47**, 6766.
42. Such MP, Park HJ, Prasad TK and Lim DW, *Chemical Reviews*, 2012, **112**, 782.

43. Murray LJ, Dinca M and Long JR, *Chemical Society Reviews*, 2009, **38**, 1294.
44. An J and Rosi NL, *Journal of the American Chemical Society*, 2010, **132**, 5578.
45. Wang QM, Shen DM, Bulow M, Lau ML, Deng SG, Fitch FR, Lemcoff NO and Semanscin J, *Microporous and Mesoporous Materials*, 2002, **55**, 217.
46. Custelcean R and Moyer BA, *European Journal of Inorganic Chemistry*, 2007, 1321.
47. Furukawa H, Miller MA and Yaghi OM, *Journal of Materials Chemistry*, 2007, **17**, 3197.
48. Liu J, Thallapally PK, McGrail BP and Brown DR, *Chemical Society Reviews*, 2012, **41**, 2308.
49. Mallick A, Saha S, Pachfule P, Roy S and Banerjee R, *Journal of Materials Chemistry*, 2010, **20**, 9073.
50. Grajciar L, Wiersum AD, Llewellyn PL, Chang JS and Nachtigall P, *Journal of Physical Chemistry C*, 2011, **115**, 17925.
51. Pera-Titus M, *Chemical Reviews*, 2014, **114**, 1413.
52. Sava DF, Kravtsov VC, Nouar F, Wojtas L, Eubank JF and Eddaoudi M, *Journal of the American Chemical Society*, 2008, **130**, 3768.
53. Hoskins BF and Robson R, *Journal of the American Chemical Society*, 1990, **112**, 1546.
54. Chuck CJ, Davidson MG, Jones MD, Kociok-Kohn G, Lunn MD and Wu S, *Inorganic Chemistry*, 2006, **45**, 6595.
55. Horcajada P, Serre C, Maurin G, Ramsahye NA, Balas F, Vallet-Regi M, Sebban M, Taulelle F and Ferey G, *Journal of the American Chemical Society*, 2008, **130**, 6774.
56. Zhao B, Chen XY, Cheng P, Liao DZ, Yan SP and Jiang ZH, *Journal of the American Chemical Society*, 2004, **126**, 15394.
57. Chen BL, Yang Y, Zapata F, Lin GN, Qian GD and Lobkovsky EB, *Advanced Materials*, 2007, **19**, 1693.
58. Chen BL, Wang LB, Zapata F, Qian GD and Lobkovsky EB, *Journal of the American Chemical Society*, 2008, **130**, 6718.

59. Harbuzaru BV, Corma A, Rey F, Atienzar P, Jorda JL, Garcia H, Ananias D, Carlos LD and Rocha J, *Angewandte Chemie-International Edition*, 2008, **47**, 1080.
60. Zuluaga S, Canepa P, Tan K, Chabal YJ and Thonhauser T, *Journal of Physics-Condensed Matter*, 2014, **26**, 15.
61. O'Keeffe M, *Chemical Society Reviews*, 2009, **38**, 1215.
62. Zhu QL and Xu Q, *Chemical Society Reviews*, 2014, **43**, 5468.
63. Hydrogen ,Fuel Cells & Infrastructure Technologies Program: Multi-yrearresearch, Development and Demonstration Plan, U. S. Department of Energy, February 2005, http://www.eere.energy.gov/hydrogenandfuelcell/mypp/ . 21 May 2018.
64. Basic Reearch Needs for the Hydrogen Economy,report of the Basic Energy Sciences,workshop on Hydrogen,Production,Storage and Use, U. S. Department of Energy,May 13–15, 2005, http://www.sc.doe.gov/bes/. Accessed 21 May 2018.
65. Lee K, Howe JD, Lin LC, Smit B and Neaton JB, *Chemistry of Materials*, 2015, **27**, 668.
66. Soares CV, Borges DD, Wiersum A, Martineau C, Nouar F, Llewellyn PL, Ramsahye NA, Serre C, Maurin G and Leitao AA, *Journal of Physical Chemistry C*, 2016, **120**, 7192.
67. Mavrandonakis A, Vogiatzis KD, Boese AD, Fink K, Heine T and Klopper W, *Inorganic Chemistry*, 2015, **54**, 8251.
68. Nagaoka M, Ohta Y and Hitomi H, *Coordination Chemistry Reviews*, 2007, **251**, 2522.
69. Fischer M, Gomes JRB and Jorge M, *Molecular Simulation*, 2014, **40**, 537.
70. Park KS, Ni Z, Cote AP, Choi JY, Huang RD, Uribe-Romo FJ, Chae HK, O'Keeffe M and Yaghi OM, *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**, 10186.
71. Assfour B, Leoni S, Yurchenko S and Seifert G, *International Journal of Hydrogen Energy*, 2011, **36**, 6005.

72. Lewis DW, Ruiz-Salvador AR, Gomez A, Rodriguez-Albelo LM, Coudert FX, Slater B, Cheetham AK and Mellot-Draznieks C, *Crystengcomm*, 2009, **11**, 2272.
73. http://www.freegrab.net/zeolite.htm. Accessed 11 May 2018.
74. Bustamante EL, Fernandez JL and Zamaro JM, *Journal of Colloid and Interface Science*, 2014, **424**, 37.
75. Li R, Ren XQ, Feng X, Li XG, Hu CW and Wang B, *Chemical Communications*, 2014, **50**, 6894.
76. Bachovchin WW, *Biochemistry*, 1986, **25**, 7751.
77. Noguchi T, Inoue Y and Tang XS, *Biochemistry*, 1999, **38**, 399.
78. Sorlie M, Christiansen J, Lemon BJ, Peters JW, Dean DR and Hales BJ, *Biochemistry*, 2001, **40**, 1540.
79. Bender ML, Bergeron RJ and Komiyama M, *Nature*, 1985, **314**, 478.
80. Stewart R, Academic Press, 1st edn., 1985, ch. 324.
81. Blow DM, *Accounts of Chemical Research*, 1976, **9**, 145.
82. McMullan RK, Epstein J, Ruble JR and Craven BM, *Acta Crystallographica Section B-Structural Science*, 1979, **35**, 688.
83. Craven BM, McMullan RK, Bell JD and Freeman HC, *Acta Crystallographica Section B-Structural Science*, 1977, **33**, 2585.
84. Pekmez NO, Can M and Yildiz A, *Acta Chimica Slovenica*, 2007, **54**, 131.
85. Veyisoglu F, Sahin M, Pekmez NO, Can M and Yildiz A, *Acta Chimica Slovenica*, 2004, **51**, 483.
86. Bredas JL, Poskin MP, Delhalle J, Andre JM and Chojnacki H, *Journal of Physical Chemistry*, 1984, **88**, 5882.
87. Tafazzoli M and Amini SK, *Chemical Physics Letters*, 2006, **431**, 421.
88. Tugsuz T, *Journal of Physical Chemistry B*, 2010, **114**, 17092.
89. Park SJ and Lee SY, *International Journal of Hydrogen Energy*, 2010, **35**, 13048.
90. Bae YS and Snurr RQ, *Microporous and Mesoporous Materials*, 2010, **135**, 178.
91. Li J, Furuta T, Goto H, Ohashi T, Fujiwara Y and Yip S, *Journal of Chemical Physics*, 2003, **119**, 2376.

92. Bhatia SK and Myers AL, *Langmuir*, 2006, **22**, 1688.
93. Bae YS and Snurr RQ, *Microporous and Mesoporous Materials*, 2010, **132**, 300.
94. Thornton AW, Nairn KM, Hill JM, Hill AJ and Hill MR, *Journal of the American Chemical Society*, 2009, **131**, 10662.
95. Rao DW, Lu RF, Meng ZS, Xu GJ, Kan EJ, Liu YZ, Xiao CY and Deng KM, *Molecular Simulation*, 2013, **39**, 968.
96. Prakash M, Sakhavand N and Shahsavari R, *Journal of Physical Chemistry C*, 2013, **117**, 24407.
97. Zhou M, Wang Q, Zhang L, Liu YC and Kang Y, *Journal of Physical Chemistry B*, 2009, **113**, 11049.
98. Hirscher M and Panella B, *Scripta Materialia*, 2007, **56**, 809.
99. Ferey G, Latroche M, Serre C, Millange F, Loiseau T and Percheron-Guegan A, *Chemical Communications*, 2003, 2976.
100. Roquerol F, Roquerol J and Sing K, *Adsorption by Powders and Porous Solids*, Academic Press, San Diego 1998.

Chapter 2
Methodology

In this chapter, we shall review the theoretical background of the methodologies that used in this project.

2 Theoretical Overviews

2.1 Schrödinger equation

The main equation of quantum mechanics is the Schrödinger equation. In fact, there are two kinds of Schrödinger equation namely the time-independent equation (TISE) and the time-dependent equation (TDSE)[1]. The time-independent form of Schrödinger equation is as the following:

$$\hat{\mathrm{H}}\Psi = \mathrm{E}\Psi \tag{1}$$

$\hat{\mathrm{H}}$ is the Hamiltonian operator of the system, Ψ is the wave function which describes the positions and motions of nuclei and electrons, and E is the energy[2]. To solve this equation for molecules, we shall define the Hamiltonian operator as the following:

$$\hat{\mathrm{H}} = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{m_A}\nabla_A^2 - \sum_{i=1}^{n}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{n}\sum_{j=1}^{i-1}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B=1}^{A-1}\frac{Z_A Z_B}{R_{AB}} \tag{2}$$

where in this equation the atomic units are used, i and j are electrons, A and B are nuclei, m_A is the mass of nucleus A, r_{ij} is the distance between i and j electrons, R_{AB} is the distance between A and B nuclei, M number of nuclei, n number of electrons, Z_A , Z_B charges of A and B, and ∇^2 is the Laplacian operator:

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \tag{3}$$

where x, y, and z are Cartesian coordinates[3].

Generally, solution of this equation Eq. 1, gives the total energy of the molecule and the molecular wavefunction Ψ. The Hamilton operator consists of the kinetic energy and potential energy, where the first two components in Eq. 2 are the kinetic energy while the last three components are the potential energy.

Indeed, the solution of the Schrödinger equation for three or more particles is impossible, so approximations are required, because of the correlated motions of particles. One of these approximations is the Born-Oppenheimer approximation that is used to solve the Schrödinger equation for many-body systems[4, 5].

2.2 Born-Oppenheimer Approximation

The Hamiltonian for a molecule has been written as a sum of five terms:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \qquad (4)$$

n = nuclear
e = electronic
ne = nucleus-electron
ee = electron-electron
nn = nucleus-nucleus

The principle of the Born-Oppenheimer approximation is that, the nuclei of molecular systems are much heavier than electrons by about 1800 times (moving very slowly compared to electrons): therefore; their velocities are much smaller. This implies that the nuclei may be considered stationary with respect to electron motion and any changes in the positions of the nuclei will affect the electronic wavefunction.

In this approximation, the molecular Schrödinger equation separates into two parts, one for electronic wavefunction and another one for the nuclear motions. So, the Hamiltonian of the nuclear Schrödinger equation consists of the nuclear-nuclear repulsion that is considered a constant (coulomb repulsion) and the nuclear kinetic energy that is ignored (equals zero). The electronic Hamiltonian includes the first, third and fourth terms of equation (2) and in equation (4) Hamiltonian includes the second, third and fourth terms:

$$\hat{\mathrm{H}}_{e-} = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{n}\sum_{j=1}^{i-1}\frac{1}{r_{ij}} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \qquad (5)$$

Thus, we obtain the electronic Schrödinger equation:

$$\hat{H}_{e^-}\ \Psi_{e^-} = E_{e^-}\ \Psi_{e^-} \qquad (6)$$

This implies that Schrödinger equation is solved in the electrostatic field of the nuclei for the electrons alone[6]. The total energy of the system is consists of two terms namely electronic energy and nuclear repulsion term.

$$E_{Tot} = E_{e^-} + \sum_{A=1}^{M}\sum_{B=1}^{A-1}\frac{Z_A Z_B}{R_{AB}} \qquad (7)$$

where E_{Tot} is called potential energy surface (PES). Furthermore, it is also called potential due to the potential energy in the dynamical equation of nuclear motion is one of the most important concepts in physical chemistry[7].

On the other hand, there are limitations of the Born-Oppenheimer approximation. The total wavefunction is limited to a particular electronic state, i.e. one electronic surface. Also, the BO approximation is usually great, but breaks down when the energies of the electronic Schrödinger equation are close at particular nuclear geometries[4, 5, 8].

2.3 Ab initio methods

The term ab initio, also known as wavefunction theory (WFT), does not mean exact solution of the Schrödinger equation. It means from first principles. This means it should be select a method that in principle can lead to a reasonable approximation to the solution of the Schrödinger equation.

Ab initio quantum mechanics uses the laws of physics to predict the properties of molecules, specifically by solving the Schrödinger equation for the system see Eq. 1. Indeed, among all the approximation methods available, *ab initio* methods present complementary advantages. The *ab initio* method is accurate but time consuming and is limited to small systems[9].

An extensive range of ab initio methods has been used, but we will focus on the methods that are used in the majority of all calculations carried out nowadays. This is the method that uses the molecular orbital method, probably followed by a post

molecular orbital method that uses the molecular orbital wave function as the reference function. The molecular orbital method is mostly referred to as the Hartree-Fock method and the post molecular orbital methods are generally referred to as the electronic correlation methods.

2.3.1 Slater Determinants

It might be useful to start with a wavefunction of the general form

$$\Psi_{HP}(r_1, r_2, \cdots, r_N) = \varphi_1(r_1)\,\varphi_2(r_2)\cdots\varphi_N(r_N) \qquad (8)$$

where Ψ_{HP} is a Hartree Product. N is the number of electrons and molecular orbitals. r_i, φ are the coordinates of electron i and the molecular orbital respectively.

In this case, the wavefunction does not satisfy the antisymmetry principle (interchange of electrons results in a change in the sign of the wavefunction). To understand that and know why, consider the wavefunction for only two electrons:

$$\Psi_{HP}(r_1, r_2) = \varphi_1(r_1)\,\varphi_2(r_2) \qquad (9)$$

If we swap the coordinates of electron 1 with those of electron 2, we will find

$$\Psi_{HP}(r_2, r_1) = \varphi_1(r_2)\,\varphi_2(r_1) \qquad (10)$$

To get the negative of the original wavefunction, there is only one way to get that

$$\varphi_1(r_2)\,\varphi_2(r_1) = -\,\varphi_1(r_1)\,\varphi_2(r_2) \qquad (11)$$

but that will not be true in general, because the Hartree Product actually does not have the properties we require. So, we can satisfy the antisymmetry principle by a wavefunction like:

$$\Psi_{HP}(r_2, r_1) = \frac{1}{\sqrt{2}}\left[\varphi_1(r_1)\,\varphi_2(r_2) - \varphi_1(r_2)\,\varphi_2(r_1)\right] \qquad (12)$$

But in case of more than two electrons, it is useful to generalize the above equation to N electrons by using determinants. For two electrons we can rewrite the above equation as the following:

$$\Psi_{\mathrm{HP}}(\mathrm{r}_2, \mathrm{r}_1) = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_1(\mathrm{r}_1) & \varphi_2(\mathrm{r}_1) \\ \varphi_1(\mathrm{r}_2) & \varphi_2(\mathrm{r}_2) \end{vmatrix} \quad (13)$$

For N electrons, wavefunction can be written as a determinant like

$$\Psi_{\mathrm{HP}}(\mathrm{r}_1, \mathrm{r}_2, \ldots\ldots, \mathrm{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathrm{r}_1) & \varphi_2(\mathrm{r}_1) & \ldots & \varphi_N(\mathrm{r}_1) \\ \varphi_1(\mathrm{r}_2) & \varphi_2(\mathrm{r}_2) & \ldots & \varphi_N(\mathrm{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathrm{r}_N) & \varphi_2(\mathrm{r}_N) & \ldots & \varphi_N(\mathrm{r}_N) \end{vmatrix} \quad (14)$$

This determinant of spin orbitals is called a Slater Determinant, and the interchange of two rows or columns of a determinant changes the sign of the determinant from linear algebra, which makes them perfect for expressing electronic wavefunctions and satisfy the antisymmetry principle[10].

2.3.2 Hartree-Fock method

The standard method of ab initio electronic structure calculation is Hartree-Fock (HF) that is able to solve the Schrödinger equation for many-body systems and accounts for about 99% of the total energy. Although the 1% of the total energy is fundamental for describing chemical phenomena, HF fails to account for it because it does not take into consideration electronic correlation, which in turn makes it inappropriate for some purposes such as calculations of energies in cases of reaction and bond dissociation. So, it does not provide the exact energy E_{exact}:

$$E_{cor} = E_{exact} - E_{HF} \quad (15)$$

where E_{cor} is the correlation energy, which is the difference between the exact energy E_{exact} and HF energy E_{HF}. It is an important term for calculating chemical properties.
In fact, the HF method depends on the one-electron Fock operator *h(i),* in which the interelectronic repulsion is presented in a non-local operator (an average potential) $U^{HF}(i)$[11], where the single electron *(i)* has been affected by the other electrons occupying orbitals.

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_k \frac{Z_k}{r_{ik}} + U^{HF}(i) \quad (16)$$

$$U^{HF} = \sum_{i<j} v\,(i,j) + \sum_{i=1} k(i,j) \qquad (17)$$

The solution of one electron affects alternate electrons by the exchange potential term $k(i,j)$, and $v\,(i,j)$ is the operator for two-electron (their coulomb repulsion).

Now we can write the electronic Hamiltonian as the following form

$$\widehat{H} = \sum_{i} h(i) + \sum_{i<j} v\,(i,j) + \sum_{i=1} k(i,j) + V_{NN} \qquad (18)$$

where V_{NN} is a constant (for the fixed set of nuclear coordinates {R}), it does not change the eigenfunctions, and just shifts the eigenvalues so it will be ignored. In Hartree-Fock theory[12], the Slater Determinant is used to introduce the independent particle model and it is used as the basis to describe N interacting electrons. The energy of the (Hartree-Fock) wavefunction $|\Psi\rangle$ that is associated with the Slater determinant can be obtained by minimization of the standard quantum mechanical energy.

$$E_{HF} = \frac{\langle \Psi_0 | \widehat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \qquad (19)$$

The main aim of Hartree-Fock theory is to construct the optimal set of 1-electron spinorbitals $\{\varphi_i\}$ that define the Slater determinant, to order to minimize the standard quantum mechanical energy. The Hartree – Fock equation (20) need to be solved by the Self-Consistent Field (SCF) method.

$$h(i)\varphi_i = \mathcal{E}_i \varphi_i \qquad (20)$$

Generally, E_{cor} in Eq.15 is an essential part for calculating chemical properties, so many types of calculations begin with the Hartree-Fock calculation and subsequently correct for electron-electron correlation, also referred to as the electronic correlation[8]. To describe the correlation energy, we can divide it into dynamic correlation and non-dynamic (static) correlation.

2.3.2.1 Non-dynamic (static) correlation and dynamic correlation

The effects of static correlation are most observable when studying bond stretching and reaction pathways. It is also known as "long range" correlation because when atoms are placed at long bond lengths, states become closely degenerate, and that leads to multi-reference behavior, and it has been called "near degeneracy" correlation[13].

The molecular orbitals are formed when a molecule forms a bond and the atomic valence orbitals overlap. Therefore there will be a gap between the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). In this case, there are two possibilities. When the gap is large, the valence electrons seem to move in a mean-field potential of the other electrons. Therefore, then the system can be described qualitatively by a theory based on a single reference determinant, and sometimes by a single reference mean field approximation as Hartree-Fock.

When the gap between HOMO and LUMO is small, near to degeneracy. This correlation of electrons is called static correlation where the wavefunction is no longer dominated by a single configuration and is instead a superposition of several configurations similar in energy.

In terms of dynamic correlation Hartree-Fock assumes that the possibility of existence of electron 1 at a certain position and electron 2 at another position is simply product of the two 1-electron probability densities. From the Pauli Exclusion Principle, electrons of the same spin are less likely to be close to each other, hence dynamic correlation effects are very essential for electrons of opposite spin. So, dynamic correlation is greatest in doubly occupied orbitals, and that is because molecular orbitals are spatially small. For example in F_2 where the atomic orbitals are greater than the molecular ones, that causes increasing the dynamic correlation energy. Since the probability of finding 2 electrons near together is overvalued in Hartree-Fock as a mean field theory, the repulsion energy is overvalued, and the resulting total energy is greater than that of the exact answer. A lack of dynamic correlation causes Hartree-Fock to underestimate binding and overestimate bond lengths[14]. The wave function in the HF model is a single Slater determinant, which might be a rather poor representation of a many-electron system's state: in certain cases, an electronic state can be well described only by a linear combination of more than one (nearly-)degenerate Slater determinant.

2.3.2.2 Basis sets and Classification of basis sets

To calculate the energy of atoms, it was necessary to define mathematical functions for orbitals, and it was simply by using solutions of the Schrödinger equation for the H atom as a starting point and use the variational principle to find the best exponent for each function. But in case of molecular orbitals (MO) what functions should we use? To answer this question, we shall introduce the basis set[15].

The simple definition of basis set is a set of functions that is centred on the atoms to describe the atomic orbitals for a range of atoms where they are combined in linear combinations of atomic orbitals (the LCAO approximation) to create basis functions for the spatial part of the molecular orbitals in Slater determinant.

$$\psi_i = \sum_{\mu=1}^{n} c_{\mu i}\ \chi_\mu \qquad (21)$$

Where the ψ_i is the i[th] molecular orbital, n is the total number of the atomic orbitals, $c_{\mu i}$ are the coefficients of the linear combination (sometimes called an MO coefficient), χ_μ is the μth atomic basis set orbital. There are a lot of basis sets that are designed to provide a description for the lowest cost because large basis sets give a better or more accurate result but they cause a higher computational cost. Therefore, there are many kinds of basis sets namely a minimal basis set, an expanded basis set (double zeta / triple zeta / etc.), a split valence, a diffuse function, a polarization function, Dunning basis sets (correlation consistent basis sets)[16] [17] and augmented correlation consistent basis sets. For more illustration,

- **Minimal basis set** means one basis function for each atomic orbital in the atom as Slater types orbital (STO) and Gaussian type orbital GTO). In fact, STO is not used very much and GTO is considered better than STO, since the GTO have better integration performance where, computer evaluation is much faster. Sometimes used as linear combination of Gaussian (STO-3G) is used to approximate STOs[18]. For example:

H: 1 (1s)

C: 5 (1s, 2s, 2px, 2py, 2pz)

- **Expanded basis sets,** where there is more than just single function of each angular moment on each atom, are normally required, in order to allow the molecular orbitals to assume a different size and shape to the atoms orbitals. The degree of flexibility is sometimes expressed through terminology such as "double zeta", including two functions corresponding to each atomic orbital. More explicitly,

Double Zeta: 2 basis functions for each orbital

H: 2 basis functions

C: 10 basis functions

CH_4: 18 basis functions

Triple Zeta: 3 basis functions for each orbital

H: 3 basis functions

C: 15 basis functions

CH_4: 27 basis functions

For example: 6-311G

- **A split-valence basis** means a larger basis for the valence AOs and using only one basis function for each core AO (minimal basis sets). The reason for that is the core electrons of an atom are less affected than the valence electrons by the chemical environment. Also split valence basis sets are used for large molecules because they decline the amount of time of a central processing unit (CPU) time which is required for the calculation.

Split Valence: 1 basis function for each core orbital, but 2 or more for valence

H: 2 basis functions

C: 9 basis functions

CH_4: 17 basis functions

For example: 3-21G

- **Polarization Functions** means any orbital has a higher angular momentum used in a basis set that is not usually occupied in the separated atom. Adding polarization functions in the basis set is another method to increase the size of the basis set in order to obtain closer to the exact wavefunction and electronic energy is to include polarization functions in the basis set. For example, for the hydrogen atom, the orbital that is occupied is s-type only. Therefore, if p-type or d-type basis functions were added to the hydrogen atom they would be known as polarization functions. Also, for first row

elements as carbon atom, d-type and f-type basis functions would be considered to be polarization functions. For transition metals the orbitals that are occupied are d-type orbitals, so only f-type or higher functions would be considered polarization basis functions. In fact, there is need to add polarization basis functions to improve the flexibility of the basis set, especially to better represent electron density in bonding areas. In more details, in the isolated hydrogen atom, the electron density is spherical as the s-type orbital is occupied, while it is shifted or polarized when the hydrogen atom makes a bond with another atom like the C-H bond in methane.

For p orbitals, add in d functions (6 of them).

Example: 6-31G* = 6-31G(d)

For s orbitals, add in p functions (3 of them).

Example: 6-31G** = 6-31G(d, p)

- **Diffuse Functions** are necessary for accurate polarizabilities or binding energies of van der Waals complexes (bound by dispersion). Also, they are useful to do computations on anions, excited states, transition states and molecules with lone pairs where electrons can move far from the nucleus. So there is a need to introduce diffuse functions.

Example:

6-31+G

To understand the meaning of **6-311+G**** basis set, we shall indicate the meaning of every symbol in this basis set as the following:

6: Each inner shell (core) basis function composed of 6 primitives

311: Triple-zeta split valence basis: One is contracted function of 3 primitives, and the other two are single Gaussians

+: Add diffuse Functions

*: Polarization of p-orbitals with d functions

*: Polarization of s-orbitals with p functions

- **Dunning's Correlation-Consistent Basis Sets** as cc-pVXZ is a Dunning correlation-consistent, polarized valence, X-zeta basis; X=D, T, Q, 5, 6, 7. These have been designed to over-come the high cost, to reduce valence flexibility of the atomic orbitals

bases [19] and to recover consistently the correlation in the valence electrons, by adding polarization functions. For example, cc-pVDZ for C atom consists of 3s2p1d. cc-pVTZ would be 4s3p2d1f. cc-pVQZ would be 5s4p3d2f1g. Where diffuse functions are needed, standard augmenting sets, denoted with the prefix "aug" (e.g., aug-cc-pVTZ), have been obtained from calculation on the atomic negative ions as mention in diffusion functions section. For example: aug-cc-pVDZ for C atom has diffuse s, p, d functions. Moreover, to describe core correlation, the letter "C" in the cc-pCVXZ or aug-cc-pCVXZ basis sets are needed and these basis sets should be used if we do not freeze core [4, 8]. For more details, aug-cc-pVTZ as the following:

aug: diffuse functions.

cc: correlation consistent.

pVTZ: polarization Valence Triple Zeta.

2.3.2.2.1 Basis set extrapolation

Generally, when we want to obtain the intermolecular interaction energy, we face two problems, basis set superposition error (BSSE) will be explained in the next section, and a slow convergence of the intermolecular interaction energy[20]. The interaction energy can be calculated using the following equation only when an infinite basis set is used:

$$\Delta E = E_{A\text{-}B} - (E_A + E_B) \qquad (22)$$

To overcome these problems, we can use the counterpoise method as will be explained in the next section, and basis set extrapolation (Extrapolations between two adjacent basis sets (*e.g.,* cc-pVTZ & cc-pVQZ or aug-cc-pVDZ & aug-cc-pVTZ) to complete the basis set limit. When the finite basis is expanded towards an (infinite) complete set of functions, calculations using this a basis set are said to approach the complete basis set (CBS) limit.

Basis set extrapolation to the CBS limit has been studied for post Hartree-Fock correlated methods, with the purpose of reducing as much as possible the basis set truncation error without using large basis sets[21]. Indeed, to obtain high accuracy values of energies through extrapolation of the results obtained from a series of correlation consistent basis sets with increasing sequential cardinal numbers x. From literature

review, there are many formulas of extrapolation that can be used. One of them is the x^{-3} expression with consecutive x-tuple-ζ basis sets[22-27] as the following:

$$E_{x,x+1}^{corr} = \frac{(x+1)^3 E_{x+1}^{corr} - x^3 E_x^{corr}}{(x+1)^3 - x^3} \quad (23)$$

Extrapolated total energies are calculated by adding to the extrapolated correlation energy the Hartree-Fock energy, E_x^{corr}, from the calculation with the larger of the two basis sets used in the extrapolation

$$E_{x,x+1} = E_{x+1}^{HF} + E_{x,x+1}^{corr} \quad (24)$$

Overall, for the intermolecular interactions energies, CP-correction and basis set extrapolation are yield very reliable results that are very close to the complete basis set limit[28].

2.3.2.2.2 Basis Set Superposition Error and counterpoise method

The BSSE is due to an imbalance between the approximations used for the free monomer and for the supermolecule (complex), where the complex has more basis functions used in the calculations than in either of monomers. That means that, in the complex, each monomer is able to use, at least in part, the basis functions of its interaction partners. When the binding or interaction energy is calculated, the computed energy of the whole system is artificially low in comparison to the separated monomers, which do not benefit from the basis functions of their interaction partners. This causes grave problems for respect of calculating intermolecular interaction energies such as deformations of shape and depth of the calculated potential surfaces, particularly if the basis set applied is small[29 30]. We can obtain the interaction energy of the complex AB from the following equation,

$$\Delta E(R) = E^{AB}(R) - E^A - E^B \quad (25)$$

Where R is the interfragment distance A-B. For large separations, E^{AB} increases to the sum E^A+E^B. In equation (25), E^A and E^B are assumed to be evaluated using the A basis set for E^A and the B basis set for E^B and when A and B not both infinite basis sets, there are more basis sets employed in the calculation of the complex. It means each monomer

uses basis set for other. A ΔE obtained in this case is too large. In other words, when one of the monomers borrows the basis set of the other to improve its own wave function and the ΔE obtained will be declining, in this case the basis set is called the ghost basis set. The counterpoise method is one of the methods proposed to correct this phenomenon and we can explain how we can obtain the correction of energy through the following equation,

$$\Delta E^{CP}(R) = E^{AB}(R) - E^{A\{AB\}}(R) - E^{B\{AB\}}(R) \qquad (26)$$

Where $E^{A\{AB\}}$ and $E^{B\{AB\}}$ are the energies of the monomer obtained using the full dimer basis {AB} at the particular AB geometry one is studying. The {A} basis in the $E^{B\{AB\}}$ calculation and the {B} basis in the $E^{A\{AB\}}$ calculation are called the ghost basis sets as previously mentioned[31], implying that $E^{A\{AB\}}(R)$ is the energy of a dimer consisting of an A atom and a B ghost atom (an atom without nucleus and electrons, but having its orbitals), and $E^{B\{AB\}}$ vice versa. Note that in Eq. (26) the energy of the separate atoms depends on the distance between the atom and the ghost atom (an internuclear distance). Furthermore, the BSSE can be reduced by extrapolating the ab initio energies to the complete basis set limit[32, 33] as mentioned above.

To understand how the calculation of dissociation energy with counterpoise correction is run, we can assume A-B dimer. In this case the **Molpro** package is used for all our calculations. **Molpro** program calculates the following energies to obtain dissociation energy D_o and counterpoise correction CP:

1- Compute the total energy of the unrelaxed monomers when the distance between the two monomers is infinite (large separation).The symbol of this energy is E_{inf}.
2- Compute the energies of the monomers in the dimer basis set. These energies are called E_a for monomer a and E_b for monomer b.
3- Compute the energy of the dimer or complex E_c.
4- Compute optimized monomer energies (energy of relaxed monomer) $E_{a`}$,$E_{b`}$.
5- Finally, compute:
    a. Counterpoise correction (CPC)= ($E_{inf} - E_a - E_b$)
    b. CPC corrected energy relative to unrelaxed monomers
    (de)= ($E_c - E_a - E_b$)

c. Relaxation energy (erelax)= ($E_{a`}$ + $E_{b`}$ - E_{inf})
d. CPC corrected energy relative to relaxed monomers (derel)=de-erelax.

2.3.3 Electronic structure (Electronic correlation) theory

We shall present an overview of Electronic Structure Theory, which is an ever-increasing field, which combines chemistry and theoretical physics with mathematics and computer science. Electronic Structure Theory describes the motions of electrons in atoms or molecules and it concentrates on the structure of molecules and their reactivity. Electronic Structure Theory comprises many types of calculations such as computing the energy of the molecule, performing geometric optimization and calculating the vibrational frequencies of molecules. All of the geometric optimization and vibrational frequencies depend on the first derivative and second derivative of energy respectively[2].

To start and present Electronic Structure Theory, we shall ask two questions. First, why do we need to apply the Electronic Structure Theory? Second, why is the Hartree-Fock method not efficient to give the correct solution to the Schrödinger equation if a very flexible and large basis set is chosen? In fact, the second question is the answer to the first question, so a wide number of methods have been used to improve the Hartree-Fock method. The "Hartree-Fock limit" is the name of the best Hartree-Fock wave function, obtained with a complete basis set. The problem is the two electrons have the same probability of being in the same region of space as being in separate symmetry equivalent regions of space. Also, the Hartree-Fock method evaluates the repulsion energy only as an average over the whole molecular orbital. In reality, the two electrons in a molecular orbital are moving in such a way that they keep more separately from each other than being close. This effect is called "correlation", and the correlation energy is defined as the difference of the exact energy and the Hartree-Fock limit energy. The classification of methods that deal with the correlation problem is divided into three methods: perturbation methods, variational methods or density functional methods.

2.3.3.1 Møller–Plesset Perturbation Theory (MPPT)

MPPT method is based on perturbation theory and it improves HF method by adding electron correlation effects. MPPT method is considered a special case of Rayleigh–Schrödinger perturbation theory (RSPT) and the difference between MPPT and other RSPT lies on the choice of the perturbation operator V[34, 35]. The Hamiltonian H of a system is divided into two parts; a zeroth order Hamiltonian H_0 and a perturbation λV, which is assumed to be small

$$H = H_0 + \lambda V \tag{27}$$

H equals to its true value if λ is 1. The starting points of perturbation expansion in this approach are the HF calculation, the eigenfunctions Ψ_i and the eigenvalues E_i of the Hamiltonian H, which are expanded as a power series in λ:

$$\Psi_\lambda = \Psi_0 + \lambda\Psi_1 + \lambda^2\Psi_2 + \ldots \tag{28}$$

$$\mathrm{E}_\lambda = \mathrm{E}_0 + \lambda\,\mathrm{E}_1 + \lambda^2\mathrm{E}_2 + \ldots \tag{29}$$

Ψ_0 is the HF wavefunction. $\mathrm{E}_1$and E_2 are the first-order correction and second-order correction, respectively. From the eigenfunction, the energies can be calculated as:

$$E_0 = \int \Psi_0\, H_0\, \Psi_0\, d\tau \tag{30}$$

$$E_1 = \int \Psi_0\, V\, \Psi_0\, d\tau \tag{31}$$

$$E_2 = \int \Psi_0\, V\, \Psi_1\, d\tau \tag{32}$$

$$E_3 = \int \Psi_0\, V\, \Psi_2\, d\tau - E_1 \int \Psi_0\ \Psi_2\, d\tau \tag{33}$$

In fact, the problem of electron correlation can be effectively addressed using MPPT, which at low order can carried out at low computational cost (depends on the order of the perturbation); even so except at second order, it can be prohibitively expensive for large systems. In addition, there is a limitation of this method, which is limited to small system. It is considered as a high level of theory and the accuracy is satisfactory compared to its relatively low computational cost. Another limitation is the convergence of the MPn series. The inclusion of new correlation effects at even orders n and a coupling of these correlation effects at the next higher odd order can lead to an

oscillatory behaviour of properties and molecular energies with increasing order *n* (Figure 2-1)[34, 36-40].

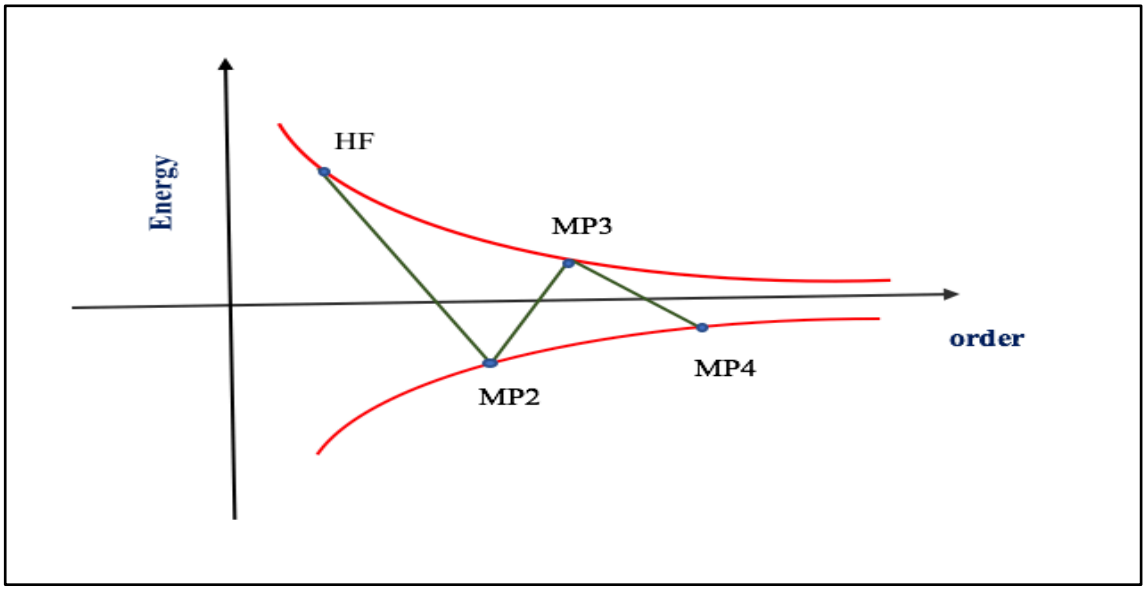


Fig.2-1: Typical oscillatory behaviour of calculated MPn response properties and molecular energy on the order n.

It is worth mentioning that MPPT theory is not variational. This implies that the calculated energy may be lower than the true ground state energy. Also, it is known that the interaction energy is usually overestimated in the case of MP2.

In terms of the orders of MPPT, many energy levels of MPPT are there namely MP0, MP1, MP2, MP3, MP4 and MP5, etc[41]. MP3, MP4 and MP5 are available but they are rarely used because of various important molecular properties calculated at MP2 is in better than their MP3 and MP4 level counterparts, even for small molecules.

The zeroth-order energy is given by the sum of occupied orbital energies

$$E_0^{(0)} = \sum_a^{occ} \varepsilon_a \qquad (34)$$

The first-order energy E_{MP1} is:

$$E_{MP1} = \langle \Psi_0 | \hat{V} | \Psi_0 \rangle \qquad (35)$$

The HF energy corresponds to the sum of zeroth-order energy and first-order energy

$$E_{HF} = E_0 + E_1 \qquad (36)$$

The lowest-order MPPT correlation energy appears in second order MP2[34]. This result is the Møller–Plesset perturbation: The correlation potential does not contribute in first-order to the exact electronic energy. For closed-shell molecules, through second-order in the correlation potential, the total electronic energy is given by the sum of Hartree–Fock energy and second-order MP2 correction:

$$E_{tot} = H_{HF} + E_{MP2} \tag{37}$$

From equations (30-37), we can understand that second-order Møller–Plesset perturbation (MP2) is required to make an improvement on the Hartree-Fock energy.

The MP2 total energy is defined as follows:

$$E_{MP2} = \frac{1}{4}\sum_{i<j}^{occ}\sum_{a<b}^{vir} \frac{(<ab\|ij>)^2}{\varepsilon_i+\varepsilon_j-\varepsilon_a-\varepsilon_b} \tag{38}$$

where $_i$, $_j$ and $_a$, $_b$ are the canonical occupied orbitals and virtual (unoccupied) orbitals, respectively and the double bar integral is an anti-symmetrized two-electron integral. The quantities ε_i, ε_j, ε_a, and ε_b are the corresponding orbital energies[34]. MP2 provides accurate binding energies for hydrogen bonds, although, it is known that MP2 usually overestimates the interaction energy[48-42] . Moreover, MP2 with small basis sets is suitable to apply to molecules with hundreds of atoms[48]. In addition, the simple two-electron systems cannot be described exactly, meaning if the perturbation series is summed to infinite order, the MP2 theory can only compute the exact electronic wavefunction for any system. Also, the perturbation series may not converge completely[49, 50].

Finally, to enhance the performance of MP2 and reduce its cost, many methods have been developed such as density-fitting MP2 (DF-MP2), local MP2 (LMP2), and explicitly correlated MP2 (MP2-F12).

Density fitting can be used with several methods like restricted Hartree-Fock, second-order Møller-Plesset perturbation theory, density functional theory, explicitly correlated MP2 (MP2-F12), also for all levels of closed-shell local correlation method ((LCC2, LMP2, LMP4, LQCISD, LCCSD(T)), as well as for CASSCF and CASPT2[51]. One enormous advantage of the density fitting scheme is that the storage requirements are incredibly decreased. Also, the error caused by the DF approximation is negligible[52, 53].

In LMP2, in this technique the local nature of electron correlation is used[54, 55]. The local MP2 (LMP2) method is virtually free from BSSE on the correlated level through the usage of a subset of the virtual orbitals for the perturbation calculation. Two important advantages of the local MP2 method (LMP2) are reduced basis set superposition error (BSSE), and reduced dependence of the computational cost on the size of molecule. Moreover, it has been found that LMP2- and CP-corrected MP2 equilibrium geometries of water and water clusters are justly close[56, 57]. So, recently it appeared as an alternative for the study of intermolecular interactions. Also, a comparison between LMP2 and MP2 shows that the most important advantage of LMP2 compared to MP2 is the significant time saving. Hence, LMP2 is very useful for large molecules.

In this context, the explicitly correlated MP2 (MP2-F12), overcomes the slow convergence of the interaction energy with respect to basis set[43, 47, 51]. The majority of ab initio methods represent the electronic wave function by a linear combination of products of one-electron functions, which do not describe accurately the Coulomb hole and which cannot represent the electron-correlation cusp. The R12 method introduces additional two-electron basis functions $r_{12}\,\psi_i(r1)\,\psi_j(r2)$ for each pair of occupied orbitals, where $r_{12} = |\overrightarrow{r_1} - \overrightarrow{r_2}|$; the F12 approach improves this with more general correlation factor $f(r_{12})$[58]. The results of interaction energies calculated by this method with the aug-cc-pVTZ basis set are typically equivalent to those obtained from the MP2/ aug-cc-pV5Z[59].

2.3.3.2 Configuration interaction methods (CI)

The configuration interaction (CI) method is considered perhaps the easiest method to understand and the oldest methods which contributed to solving the problem of electronic correlation[8]. The CI method includes the effects of electronic correlation and the exact wavefunction is represented as a linear combination of N-electron configurations[60]. The coefficients of the different configurations are optimized by using the variational method (The variational method depends on guess a ``trial'' wavefunction for the problem, which contains some adjustable parameters called ``variational parameters.'' These parameters are adjusted to obtain the minimized energy of the trial wavefunction. The resulting of the energy of the trial wavefunction is variational method approximation to the exact wavefunction and energy)[61].

$$\Psi = \sum_k c_k \Psi_k \qquad (39)$$

We can rewrite the wave function in equation (39) as follows:

$$\Psi = c_0\Psi_0 + \sum_{i=1}^{N}\sum_{a=N+1}^{M} c_i^a\Psi_i^a + \sum_{i>j=1}^{N}\sum_{a>b=N+1}^{M} c_{ij}^{ab}\Psi_{ij}^{ab} + \cdots \qquad (40)$$

Where

N is the number of electrons

M is the total number of HF orbitals

Ψ_0 is the ground-state HF wavefunction.

Ψ_i^a is a Slater determinant with an electron "excited" from the occupied orbital i to the unoccupied orbital a.

Ψ_{ij}^{ab} are "doubly-excited" Slater determinants.

When all electrons have been excited to all virtual orbitals the basis set of N-electron wavefunctions used is complete and an exact energy would be obtained. This is called full configuration interaction (FCI), for which the number of determinants ignoring spin is [8]

$$N_{tot} = \binom{M}{N} = \frac{M!}{N!(M-N)!} \qquad (41)$$

The FCI calculation will give an essentially exact result when M is large. On the other hand, there is a limitation in the FCI calculation, which is used for small molecules and with small basis sets because of its N! computational scaling. In addition, an FCI calculation is enormously expensive. Truncated forms of CI are performed where the excitation operator is truncated to include only specific excitation levels. The typical truncations are to include only doubles or to include single and double excitations, giving the methods CID and CISD[62, 63, 24, 64].

2.3.3.2.1 Size-extensivity and Size-consistency

Size extensivity and Size-consistency are properties of a computational method (scaling properties), and they are important concepts in many-body physics and quantum chemistry[65]. The definition of size extensivity is when the energy of the system and its size are properly proportional (the correlation energy scales with the number of electrons in the system). In another meaning, the number of particles scales properly with the method[8]. This property is of specific significance to obtain correctly dissociation curves. On the other hand, the definition of size-consistency is when the energy of the single particle is not half the energy of two infinitely separated particles. It is worth to mention that FCI method is size-extensive and size-consistent, whereas the CI method is not size-extensive and size-consistent.

2.3.3.3 Coupled cluster methods (CC)

One of the more accurate and elegant techniques for the treatment of electronic correlation effects is coupled- cluster (CC) theory. All types of corrections (S, D, T, Q, etc.) are added to the reference wave function[48]. The abbreviations for coupled-cluster methods begin with the letters CC followed by S - for single, D - for double excitations, etc. In addition, the CC method is a size- consistent method. In CCSD(T), the T in brackets means perturbative triple excitations. It is generally an improvement over CCSD, and CCSD(T) is now widely known as the ” gold-standard ” of quantum chemistry for single reference calculations[48].

The wave function of the coupled- cluster method can be written as

$$(\Psi) = \exp(T)\Psi_{HF} \tag{42}$$

T is the cluster operator and it is defined as

$$T = T_1 + T_2 + T_3 + \cdots + T_n = \sum T_i \tag{43}$$

Where n is the total number of electrons[8, 66], and i is the excitation level.

The coupled cluster is a computationally expensive method (computationally expensive means that finding the optimal solution takes a long time and enormous amounts of memory and disk space); so, in practice, it is limited to relatively small systems; and

therefore, economic schemes are needed in its place such as coupled cluster with double excitation operator CCD where T is approximated as ($T = T_2$), coupled cluster with single and double-excitation operator CCSD where T is approximated as ($T = T_1 + T_2$). Coupled cluster with single, double and triple-excitation operator CCSDT where T is approximated as ($T = T_1 + T_2 + T_3$)[67]. The CCSDT method is generally much more expensive, and is highly accurate. So, it is suitable for small systems[48].

The following order is observed in terms of computational accuracy with a medium-size basis set[8, 67].

HF << MP2 < CISD < CCSD < MP4 <CCSD(T).

Moreover, the (CC) method is non-variational which means that it does not provide an upper boundary to the true energy, but is size extensive[8]. In addition, the standard formulation of single reference coupled cluster unsuccessfully describes cases in which multiple bonds are broken[68].

2.4 Density function theory method (DFT)

DFT is a computational quantum mechanical method used in materials science, physics and chemistry and to investigate the electronic structure (essentially the ground state) of many-body systems, in particular atoms and molecules. DFT calculates the electronic density distribution instead of the wavefunction. This implies that it calculates the energy of a system (E) as a functional of the density and this is the main difference between ab initio methods and density function theory. Hence the name density functional theory comes from the use of functional of the electron density.

DFT can be used to calculate molecules with 100 atoms or above unlike ab initio methods that are used for small molecules or clusters because of their cost. The computational costs of DFT theory are relatively low when compared to correlation methods, such as MP2 theory[42 69, 70].

Moreover, in DFT method, the electronic correlation is taken into account and accurate results can often be obtained. All these factors have made DFT a very popular method. Indeed, despite these advantages of DFT theory, there are still difficulties in terms of using density functional theory to properly describe intermolecular interactions,

particularly charge transfer excitations; van der Waals forces (dispersion)[42, 45, 48,71-75]; transition states, global potential energy surfaces and some other strongly correlated systems[76].

In addition, the incomplete treatment of dispersion in DFT can negatively affect the accuracy of DFT (at least when used alone and uncorrected) in the treatment of systems which are dominated by dispersion (e.g. interacting noble gas atoms) or where dispersion competes significantly with other effects (e.g. in biomolecules). So, the development of new DFT methods designed to overcome this problem, by modifications to the functional and inclusion of additional terms to account for both core and valence electrons[77].

Many attempts have been done in order to involve the dispersion energy in HF and DFT calculations. One successful attempt has been reached by calculating a dispersion term separately by means of a modified C_6R^{-6} formula, where C_6 is a dispersion coefficient and R is the interatomic distance[78, 79] , after that, adding it to HF and DFT calculations according to the following equation[42, 45, 70]

$$E_{MF-D} = E_{MF} + E_{disp} \qquad (44)$$

where E_{disp} is dispersion correction, E_{MF} is mean field energy (HF or DFT). Early studies displayed that by adding a dispersion correction to the HF energy, calculation of the binding energy of larger complexes and rare-gases can be successful[80, 81]. This is successful encouraged many groups adding a dispersion correction term to the DFT energy[82, 83]. This approach is successful in studying non-covalent molecular interactions, particularly hydrogen-binding and dispersion energies[45, 72, 78].

Thomas and Fermi developed the first model of DFT that contains some basic elements[84]. However, Hohenberg and Kohn set out the main formulas that underpin modern DFT. The electron density in three-dimensional space is needed to calculate the energy and other property of the ground state of a system (*E*); the energy is a function of the density *p(r)*:

$$E[p(r)] = \int V_{ext}(r)\ p(r)dr + F[p(r)] \qquad (45)$$

It is clear from this equation that the energy of the system is the sum of two components, the interaction of the electrons with the external potential V_{ext} and the $F[P(r)]$ term that consists of the interelectronic interactions and the kinetic energy of the electrons[4, 85].

In the second formula, a variation principle gives for the density functionals:

$$\varepsilon_{el}[p\,(r)] \geq \varepsilon_{el}[p_0\,(r)] \qquad (46)$$

p_0 any other density and p is the true density for the system

$$\int p(r)dt = \int p_0\,(r)dt = N \qquad (47)$$

The negative of these formulas is that $p(r)$ is unknown and E depends on it. In fact, Kohn and Sham solved this problem[86 76]. On this approach, F[$p(r)$]is defined as the sum of three terms kinetic energy, electron-electron repulsive energy (Coulombic energy), and exchange and correlation as the following formula:

$$F\,[p(r)] = E_{KE}[p(r)] + E_H[p(r)] + E_{XC}[p(r)] \qquad (48)$$

Where E_{KE}[p(r)], E_H [p(r)] and E_{XC} [p(r)] are the kinetic energy, the electron-electron Coulombic energy, and the exchange and correlation, respectively. The full expression of the Kohn-Sham energy is:

$$E[p(r)] = \sum_{i=1}^{N} \int \varphi_i\,(r)dr + \frac{1}{2}\int\int \frac{p(r_1)\,p(r_2)}{|r_1 - r_2|} dr_1 dr_2 + E_{XC}[p(r)] - \sum_{A=1}^{M} \int \frac{Z_A}{|r - R_A|}\, p(r)dr \qquad (49)$$

Kohn and Sham presented the density $p(r)$of the system as "the sum of the square absolute values of a set of one-electron orthonormal orbitals"[87]:

$$p(r) = \sum_{i=1}^{N} |\varphi_i\,(r)|^2 \qquad (50)$$

From Eq. (49) we can obtain the one-electron Kohn-Sham equation:

$$\left\{-\frac{\nabla_1^2}{2} - \left[\sum_{A=1}^{M} \frac{Z_A}{r_{1A}}\right] + \int \frac{p(r_2)}{r_{12}}\, dr_2 + V_{XC}[r_1]\right\} \varphi_{i\,(r_1)} = \varepsilon_i \varphi_i(r_1) \qquad (51)$$

where:

$$V_{XC}[r_1] = \frac{\partial E_{XC}[n]}{\partial n(r)} \qquad (52)$$

$V_{XC}[r_1]$ is the exchange-correlation functional and ε_i is the energy of orbital. If we knew $E_{XC}[n]$ we could solve for the exact ground state energy and density.

2.5 Overview of increasing the accuracy of calculated intermolecular interaction energies (Composite CCSD(T)/CBS Schemes)

CCSD(T) is the one of the first methods that provides the required accuracy of the description of electronic correlation. To raise the level of the accuracy of electronic correlation calculations with a small error with respect to the complete basis set limit (CBS), the results should be extrapolated to the CBS. In fact, the first reliable extrapolation from triple- and quadruple-ζ basis sets can be done. So, for small systems, the direct extrapolation of CCSD(T) correlation energies is limited to them. On the other hand, for large systems, a compound approach can be used where the final result is built gradually from calculations in as large a basis set as possible, leading to the best estimate of the CCSD(T)/CBS energy[88].

The most common scheme used for noncovalent interactions is to divide the CCSD(T) energy into two terms; MP2 correction energy E^{MP2}; and a higher-order correction Δ CCSD(T) as the following formula[89-92]

$$E^{CCSD(T)} = E^{MP2} + \Delta\, E^{CCSD(T)} \qquad (53)$$

where every contribution of term can be calculated in a different basis set. The second term CCSD(T) correction is defined as

$$\Delta E^{CCSD(T)} = E^{CCSD(T)} - E^{MP2} \quad (54)$$

where each term $E^{CCSD(T)}$ and E^{MP2} are calculated in the same basis set. This equation corrects the error that came from that MP2 overestimates the dispersion energy where it is considered that this overestimation is the main source of the error[88].

Hobza and his colleagues[93] have established a scheme to compare the calculated energies at two different levels, where the first scheme presented the CCSD(T)/CBS interaction energy with extrapolated basis set using aug-cc-pVTZ and aug-cc-pVQZ basis sets as the two terms; MP2/CBS-interaction energies with extrapolated basis sets from aug-cc-pVTZ and aug-cc-pVQZ basis sets, and CCSD(T) correction term, which is the difference between CCSD(T) and MP2 interaction energies in a smaller basis set (aug-cc-pVDZ). A second scheme based on the same previous scheme but the CCSD (T) correction term with (aug-cc-pVTZ) basis set as the following equations

$$E^{CCSD\ (T)/CP/\ [34]} = E^{MP2/CP/\ [34]} + [E^{CCSD\ (T)/CP/A\text{-}2} - E^{MP2/CP/A\text{-}2}] \quad (55)$$

$$E^{CCSD\ (T)/CP/\ [34]} = E^{MP2/CP/\ [34]} + [E^{CCSD\ (T)/CP/A\text{-}3} - E^{MP2/CP/A\text{-}3}] \quad (56)$$

Where *A-2 and A-3* are the aug-cc-pVDZ and aug-cc-pVTZ respectively, CP is the counterpoise correction, and *[34]* denotes extrapolation using [aug-cc-pVTZ and aug-cc-pVQZ].

These schemes also used in the S66 database (database of interaction energies calculated using an accurate CCSD(T)/CBS scheme for 66 molecular complexes)[94], and they have recently tested. This study was applying on 11 H-bonded and 11 dispersion-bound complexes, and they found that using the larger basis sets for MP2 and ΔCCSD(T) energies changes the resulting CCSD(T)/CBS interaction energies by less than 1% in case H-bonded complexes and stabilization energies are much smaller than those of previously discussed H-bonded complexes. On the other hand, they were not able to compare the theoretical energies with experimental energies because it is not easy even in the simplest case represented by isolated gas-phase complex at very low temperature[93].

In another study, Hobza with Riley and Rezac[94] who presented a database of accurate interaction energies for 66 molecular complexes, where the S66 database, calculated at the estimated CCSD(T)/CBS level of theory. The complexes located within the database represent the inclusive range of interaction, involving dispersion dominated, hydrogen bonding (electrostatic dominated) interaction and mixed (dispersion/ electrostatic) interaction.

Moreover, they presented not only accurate interaction energies at potential energy minima, but also a set of 8 point along the dissociation curve, denoted to as the S66×8 data set. The importance of the accurate description of the potential energy surface for any method that is used in the calculations for nonequilibrium geometry, vibration analyses, molecular dynamics simulations and geometry minimizations. The former is particularly vital in the case of large systems where a given moiety may interact with a great number of other chemical groups, with quickly increasing for the number of interactions as a function of distance.

In this study, they were able to optimize the geometries of the complexes in S66 database, and equilibrium bond length at CCSD (T)/CBS level of theory. This is really significant advantage compared to the previous database, where the geometries of systems (very small complexes) had optimized at MP2 level of theory. Hobza and his peers had applied the equation 53 to increase the accuracy of calculated binding energies, and they found that the S66 benchmark method presents an average error value of 1.2% with the largest error being 2.5%.
These results are expected because the errors should be small; for S66 data set the errors should generally be below 3% as expected. Because the errors in this study are close to the errors of some of studied methods to which they are compared[94].

In 2009 Hobza and his colleague explored the intermolecular interaction energies for 24 different pairs of amino acid side chains in proteins at many computational methods as MP2, DFT and force field and they calculated the reference binding energies at CCSD(T)/CBS level, and the geometries of these pairs were derived from X- ray crystal structure data to a resolution of 2.0 Å or better. The estimated CCSD(T)/CBS method is considered the reference method by applying the following formula

$$E^{CCSD(T)/CP/[23]} = E^{MP2/CP/[23]} + [E^{CCSD(T)/CP/6\text{-}31G^*} - E^{MP2/CP/6\text{-}31G^*}] \quad (57)$$

[23] means the extrapolation basis set between [aug-cc-pVDZ and aug-cc-pVTZ], where this scheme presented the CCSD(T)/CBS interaction energy with extrapolated basis set between aug-cc-pVDZ and aug-cc-pVTZ basis sets as two terms; MP2/CBS interaction energies with extrapolated basis sets between aug-cc-pVDZ and aug-cc-pVTZ basis sets, and CCSD(T) correction term, which is as a difference between CCSD(T) and MP2 interaction energies in a smaller basis set (6-31G*).

In fact, they expected that the resulting interaction energies to be very close to the (still unknown) real interaction energies. The main important point regarding to the data obtained for these complexes is that all of the interactions were evaluated as attractive and the expectation was repulsive because the pairs of side chains have the same charge[72].

In addition to this study, there was another study in (2006), which addressed Benchmark database of accurate (MP2 and CCSD (T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs, carried out by Hobza, Jurecka, Sponer, and Cerny, where they calculated the interaction energies at MP2 and CCSD(T) complete basis set (CBS) limit interaction energies and geometries for more than 100 DNA base pairs, amino acid pairs and model complexes[44]. Extrapolation to the CBS limit is calculated by using two-point extrapolation methods and different basis sets (aug-cc-pVDZ - aug-cc-pVTZ, aug-cc-pVTZ - aug-cc-pVQZ, cc-pVTZ - cc-pVQZ) are applied as the following equations

$$E^{CCSD(T)/CP/[23]} = E^{MP2/CP/[23]} + [E^{CCSD(T)/CP/A\text{-}2} - E^{MP2/CP/A\text{-}2}] \quad (58)$$

$$E^{CCSD(T)/CP/[34]} = E^{MP2/CP/[34]} + [E^{CCSD(T)/CP/A\text{-}2} - E^{MP2/CP/A\text{-}2}] \quad (59)$$

$$E^{CCSD(T)/CP/[34]} = E^{MP2/CP/[34]} + [E^{CCSD(T)/CP/A\text{-}2} - E^{MP2/CP/A\text{-}2}] \quad (60)$$

The CCSD(T) correction term, defined as a difference between CCSD(T) and MP2 interaction energies, is estimated with smaller basis sets (6-31G** and cc-pVDZ) as the following formulas

$$E^{CCSD(T)/CP/[23]} = E^{MP2/CP/[23]} + [E^{CCSD(T)/CP/6\text{-}31G^*} - E^{MP2/CP/6\text{-}31G^*}] \quad (61)$$

$$E^{CCSD(T)/CP/[34]} = E^{MP2/CP/[34]} + [E^{CCSD(T)/CP/6\text{-}31G^*} - E^{MP2/CP/6\text{-}31G^*}] \quad (62)$$

$$E^{CCSD(T)/CP/[34]} = E^{MP2/CP/[34]} + [E^{CCSD(T)/CP/6\text{-}31G^*} - E^{MP2/CP/6\text{-}31G^*}] \quad (63)$$

They found that analysis of this data showed that very reasonable evaluations of the complete basis set interaction energies in proteins and DNA can be obtained employing a two-point extrapolation scheme with a pair of computationally available basis sets aug-cc-pVDZ and aug-cc-pVTZ. However, MP2 level of theory is insufficient and whenever important dispersion contribution is expected a correction for higher order correlation effects must be applied[44].

Besides of these studies, there was the study, which focuses on the accurate interaction energies of hydrogen-bonded nucleic acid base pairs. This study carried out by Hobza, Jurecka and Sponger in (2004). The summary of the interaction energy is defined as the following way

$$\Delta E^{A\ldots B} = E^{A\ldots B} - (E^A + E^B) + E_{Def} \quad (64)$$

In this study, they calculated the deformation energy built on the CBS extrapolation, where the interaction energy ΔE of a dimer A…B is defined as the previous equation, also it is important to add the deformation energy E_{Def} to interaction energy. To obtain the higher–order contributions to the correlation energy were taken into account by adding ΔCCSD (T) term to the RI-MP2/CBS energy, where ΔCCSD (T) as the following definition

$$E^{CCSD(T)/CP/[23]} = E^{RI\text{-}MP2/CBS/CP[23]} + [E^{CCSD(T)/CP/6\text{-}31G^*} - E^{MP2/CP/6\text{-}31G^*}] \quad (65)$$

Generally, they found that the interaction energies after applying this scheme are very close to the fully converged data[95].

2.6 Force field method (Molecular mechanics method)

In this computational method, the energy of system is calculated through the function of the nuclear positions particularly, in very large system. In another meaning, the energy of a system is depending on the coordinates of its particles[96].

$$E_{FF} = f(\text{nuclear positions}) \quad (66)$$

The most important in these methods the electronic motion has not been taken into account. So, this approach is useful for studying larger systems and for calculating a broad variety of dynamic and thermodynamic properties [97]. Moreover, the force field method consists of the two kinds of energies (bonded and non- bonded) energies.

$$E_{FF} = E_{bond} + E_{non-bond} \quad (67)$$

Where the bonded energy contains the bond (stretching, angle (bend), and rotation (torsion or dihedral)) energies. In contrast, the non-bonded energy contains the Van der Waals and electrostatic energies.

$$E_{FF} = E_{str} + E_{bend} + E_{tor} + E_{vdw} + E_{el} \quad (68)$$

$$E_{FF} = \sum_{str} \frac{k_b}{2}(r - r_0)^2 + \sum_{angles} \frac{k_b}{2}(\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2}[1 + \cos(n\phi - \delta)] + \sum_{i=1}^{N}\sum_{j=i+1}^{N}(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}) \quad (69)$$

In Eq. (69), first term is stretching energy, second term is bend energy, third term is torsion energy, fourth term is Van der Waals energy (this term is called Lennard- Jones equation) and the last term is electrostatic energy (Coulomb law). Where k_b and V_n are the force constants. r is the bond length and θ is the valence angle. r and θ deviate from the reference values r_0 and θ_0 respectively. δ is the phase angle. r_{ij} is the distance between atom i and j. ε_{ij} and σ_{ij} are Lennard-Jones parameters. q_i, q_j are the atomic charges on atom i and j[98, 99].

In terms of Van der Waals forces, it is describing the repulsive or attractive forces between molecular articles (or between groups within the same molecular article) other than those because of bond formation or to the electrostatic interaction of ions or of ionic groups with one another or with neutral molecules. Furthermore, the Van der Waals (attraction) includes: dipole–dipole, dipole-induced dipole and London dispersion (induced dipole-induced dipole) forces. Indeed, Van der Waals force becomes very repulsive for short intermolecular distances and it equals zero for a large distance. In quantum mechanics, the reason of repulsive term is Pauli principle, where the electronic clouds of two articles are overlapping and the negatively charged electrons repulse. While, at medium distances also there is a small attraction between two electron clouds. Implying that Van der Waals interaction energy is very positive at small distances before two particles touch each other[8, 100].

In terms of types of Force Field, there are many Force Fields, but the first one was its appearance in the 1960's, as a development of the molecular mechanics (MM) method and the aim of this type was to predict molecular structure, enthalpies of isolated molecules and vibrational spectra[101] and it was treating small organic molecules. Also, there are MM2, MM3 and MM4 all of these force fields developed by Allinger's group[96, 102-104]. Regarding to MM2, it established to study hydrocarbons, then they were improved it to able to treat various different types of organic (sulfides, ethers, amides, alcohols, etc.). since then the improvement has continued to treat with much more complex system and it became there many applicable force field such as Dreiding and Universal (UFF) force fields, where they have parameters for all the atoms in the periodic table[99, 105]. Also, there are other types of force fields such as CHARMM, AMBER and GROMOS, they often used in simulation of biomolecules[106-108]. In addition, all these three types are general as these types OPLS and COMPASS, but these two types were improved to simulate condensed matter[109, 110]. Furthermore, there are some classification of force field depends on using of energy as equation (67), [111]and it is called class I (or first generation), and class II force field (or second generation) where it includes the corrections of the intramolecular energy. For example, about class II are COMPASS, UFF, MM2, MM3 and MM4. Moreover, CFF (consistent force field) and MMFF (Merck molecular force field) are considered class II force fields[112, 113]. It is worth to mention, that the majority force field have deferent versions

and they are continuously improving (e.g. CHARMM19, CHARMM22, CHARMM27; GROMOS96, GROMOS45A3, GROMOS53A5, GROMOS53A6; AMBER91, AMBER94, AMBER96, AMBER99, AMBER02; etc.)[114-116]. Also, there are many types of force field some of them polarizable force fields and other one describes just a particular system or a class of compounds[111, 117].

2.6.1 Lennard-Jones parameters and formula

The Lennard-Jones potential (LJ) is used to describe the nonbonding interaction of two particles i and j (uncharged particules), it is given by the formula

$$E_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (70)$$

Where, interaction parameters ε (is the depth of the potential well and a measure of how strongly the two particles attract each other), σ is the finite separation at which the inter-particle potential is zero and a measure of how close two nonbonding (uncharged) particles can get, is thus referred to as the Van der Waals radius and equals to one-half of the distance between nonbonding particles and r is the separation between the particles and measured from the center of mass of one particle to the center of mass of the other particle, as shown in figure 2-2. Also, r^{-12} and r^{-6} have the following physical meaning; r^{-12} is the repulsive term, it describes Pauli repulsion at small distances; while the r^{-6} is the attraction term, it represents molecular attraction at long distances[118-121]. This formula can be transformed into equation 71,

$$E_{LJ}(r_{ij}) = A/r_{ij}^{12} - B/r_{ij}^{6} \quad (71)$$

where $A = 4\,\varepsilon\sigma^{12}$ (A is for non-polar interaction and it is the same as the C_{12} Pauli repulsion coefficient), and $B = 4\,\varepsilon\sigma^{6}$ (B is for non-polar interaction and it is the same as the C_6 dipole-dipole dispersion coefficient). Also, 12-6 potential is another name of The Lennard-Jones potential and this name comes from the exponents in the equation. The σ and ε have different value for different atoms and that represent physical properties of the showed system[122, 123]. In the case of separated particulars interacting, combination

rules can be applied to create new values of σ and ε. These values have been obtained by applying standard Lorentz-Berthelot combining rules[124-126],

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad , \qquad \varepsilon_{ij} = \sqrt{\varepsilon_{ii} * \varepsilon_{jj}} \tag{72}$$

when ***i*** and *j* refer to dissimilar atoms, in this case, the Lennard-Jones parameters for these parameters are determined using a set of combining rules. It is usual to use (the Berthelot rule) for *ε* and (the Lorentz rule) for σ as above in Eq.(72) where the unlike interactions, $i \neq j$, between different pairs of particulars, are not easily obtained[127]. For this reason, combining (mixing) rules are appropriate. The most common combining rules are Lorentz- Berthelot (LB) combining rules[128-130].

Regarding to how obtain the r_{ij}, it just the distance obtainable from the Cartesian coordinates of the two atoms:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{73}$$

where essentially all force field calculations use atomic Cartesian coordinates as the variables in the energy expression[8].

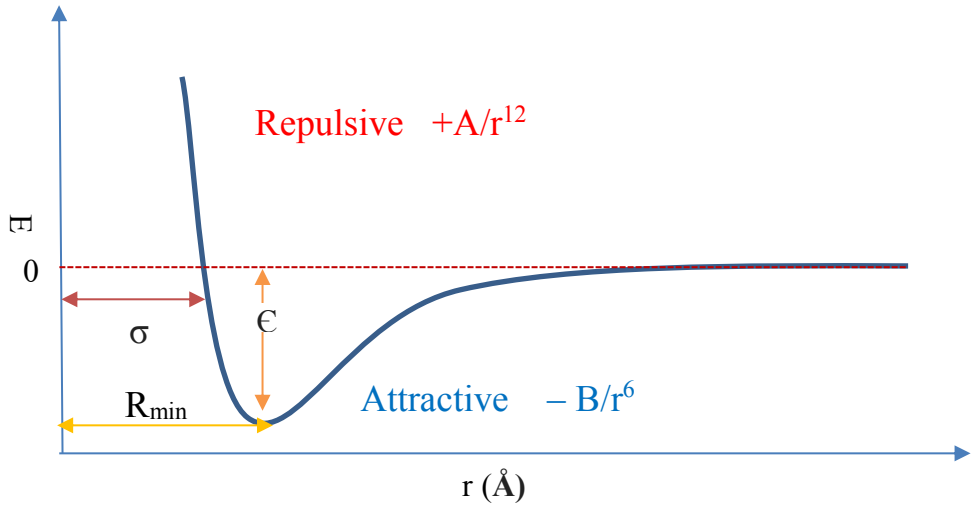


Fig. 2-2: Lennard- Jones potential (the intermolecular interactions of two particles).

Lennard-Jones potential as the bonding potential energy, the stability of an arrangement of atoms is a function of Lennard-Jones separation distance. The potential energy becomes positive (repulsive force). The overlapping of atomic orbitals causes a large potential energy. However, the potential energy is negative and equals zero at infinity separation distances, as the separation distance increases (attractive force). Indeed, this denotes that at a large distance, the pair of particles experiences a small stabilizing force. Finally, the potential energy reaches a minimum value (zero force), when the separation between the two particles reaches a distance slightly greater than σ. In this case, the pair of particles is most stable and will stay in that direction until an outer drive is applied upon it[123].

Also, there is a formula describes Van der Waals and Pauli repulsion energies $\Psi_{12}(r)$, it is called the Buckingham potential proposed by Richard Buckingham. The Buckingham potential is a function of interatomic distance r for the interaction of two atoms that are not directly bonded. Where it is known from electronic structure theory, that the repulsion causes overlap of the electronic wave functions, Moreover, the electron density falls off roughly exponentially with the distance from the nucleus. The general form of the "*Exponential –R* $^{-6}$ " *is* E_{vdW} function or Buckingham or Hill potential[8, 131], and it is written as:

$$\Psi_{12}(r) = A\exp(-Br) - \frac{C}{r^6} \qquad (74)$$

Here, A, B and C are suitable constants, $A\exp(-Br)$ is a repulsion term and $-\frac{C}{r^6}$ is an attractive term. This formula is considered as simplification of lennard Jones potential as Buckingham proposed. It is known from electronic structure theory the repulsion is because of overlap of electron clouds around two nonbonded particulars, where wavefunction $\Psi_{12}(r)$ is very positive at short separations (very repulsive), is attractive at intermediate separations (negative value), and is zero at large separations. Moreover, the attraction is because of electronic correlation that leads to dispersion or London force. Generally, the performance of the Buckingham potential is significantly better than a Lennard-Jones 12-6 potential[8, 27, 132].

2.7 Data fitting and error estimation

Fitting a mathematical function *f (x, a)* to a set of observed or given data y_1, y_2,…, y_n is an essential activity that has been the concentration of scientific interest and application for many fields. The major goal is to find that vector of parameters of *f*, which yields the "best" fit of function *f (x, a)* to the given data points, possibly subject to constraints. This typically includes some numeric or analytical procedure, which attempts to optimise a goodness-of-fit metric. In fact, the fit procedures need high skill to select the best analytical functions and adjust the parameters[133]. The fit procedures should supply; parameters for the functional forms, error estimation of the parameters and a statistical measure of the quality of fit. Indeed, in our project, there is a need to fit the data (ab initio potential energy surface PESs), and there are several methods to fit the data, where fitting ab initio PESs is still an art more than science and one of the most important methods is nonlinear least squares method. So, we shall apply this technique to fit ab initio potential energy surface to Lennard-Jones potential.

2.7.1 Nonlinear least squares method

The least square method (LSM) is probably the most common technique in statistics science and it provides the general logical for the position of the line (typically curve) of best fit among the data points being studied. In these methods, there is a dependence on the iteration to obtain the solution[134]. Also, there are two types of the least squares method namely; linear least squares method and nonlinear least square method[135]. In his section, we shall provide details about nonlinear square method. To explain this method, we suppose that there are M data points (x_i, y_i) $i = 1, 2, .. M$, need to fit to a model that has N adjustable parameters a_j, $j = 1, 2, \dots, N$, the description of the functional relationship between the measured dependent and independent variables is,

$$y = y(x_i; a_1 \dots . a_N) \qquad (75)$$

In the least-squares method, we adjust the parameters in the model function $f(x)$ in order to minimize the sum of the squares of its error at each calibration point *x(i)*. Where, in a large number of particles, the function $f(x)$ is a sum of squares of nonlinear functions and that needs to be minimized as the following:

$$\min f(x) = \sum_{i=1}^{M} (f(x_i))^2 \ , f(x_i) = \ y_i - y(x_i; a_1 \ldots . a_N) \qquad (76)$$

The minimum value of $f(x)$ occurs when the gradient (slop) is zero, where there are n gradient equations because the model has n parameters:

$$\frac{\partial f(x)}{\partial a_j} = 2 \sum_i f\ (x_i) \frac{\partial f(x_i)}{\partial a_j} \ = 0 \qquad (j = 1, \ldots, n) \qquad (77)$$

In a nonlinear system, these gradient equations do not have a closed solution because the derivatives $\frac{\partial f(x_i)}{\partial a_j}$ are functions of both the parameters and the independent variable. So initial values must be chosen for the parameters. Then, the parameters are refined iteratively Substituting Eq.76 into the gradient equations, it becomes

$$-2\ (y_i - y(x_i; a_1 \ldots . a_N)) = 0 \qquad (78)$$

In our project our data (Ab initio interaction potential energies) need to fit to Lennard-Jones equation

$$y(x_i; a_1, a_2) \ = 4a_1 \left[\left(\frac{x_i}{a_2}\right)^{12} - \left(\frac{x_i}{a_2}\right)^{6} \right] \qquad (79)$$

where Eq.79 nonlinear equation, the nonlinear least squares method is used in this situation.

In fact, the studies in this field are extensive. One of these studies that carried out by Palmer and Anchell in 1995, where they applied the calculations of ab initio to get the intermolecular potential energy surfaces for fluorine- substituted methanes using 6-31+G* basis set and include correlation using second-order Møller–Plesset perturbation theory. Also, the least squares fit of the ab initio surface to a molecular mechanics potential function involving Lennard-Jones interaction plus partial charges is performed. Moreover, the molecular mechanics potential for the thermodynamics properties are calculated using conventional molecular dynamics simulations and compared to experimental result. They found that the analytic formula used in the fit to the Fluorine methane surface reproduces the global features of the ab initio surface

reasonably well, but it is not exactly. This may be the main reason of uncertainly in the fitted parameters[136].

Other study was 2012, Wongsinlatam and coworkers, where they calculated the intermolecular interaction between CO_2 and benzimidazolate. This intermolecular interaction has been derived by Levenberg-Marquardt algorithm and least square method to ab initio single point energies. The energies are obtained at HF/6-31G(d). They found that the new parameters of Levenberg-Marquardt algorithm able to represent QM data better than the least square method[123].

Additional, the importance of LSM comes from many reasons. First, it used in many different areas, where it used in classical research areas such as economics, physics and chemistry etc. Second, using squares makes LSM mathematically very easy to handle because the Pythagorean theorem denotes that, when the error is independent of an evaluated amount, one can add the squared error and the squared evaluated amount. Third, algorithms and the mathematical tools involved in least square method have been well studied for a relatively long period[123, 137, 138].

2.9 References

1. Briggs JS and Rost JM, *Foundations of Physics* 2001, **31**.
2. Foresman JB and Frisch A, *Exploring Chemistry with Electronic Structure Methods*, Gaussian, second edn., 1996.
3. Barde NP, Patil SD, Kokne PM and Bardapurkar PP, *Leonardo Electronic Journal of Practices and Technologies* 2015, **14**, 31.
4. Cramer CJ, *Essential Computational Chemistry Theories and Models*, Wiley, 2004.
5. Atkins P and Friedman R, *Molecular Quantum Mechanics*, Oxford University Press, 2005.
6. Diestler DJ, *Journal of Physical Chemistry A*, 2013, **117**, 4698.
7. Yarkony DR, *Reviews of Modern Physics*, 1996, **68**, 985.
8. Jensen F, *Introduction to Computational Chemistry*, Chichester: John Wiley, 1999.
9. Jeannot E and Monard G, *ITCC 2005: International Conference on Information Technology: Coding and Computing, Vol 1*, 2005, 286.
10. http://vergil.chemistry.gatech.edu/courses/chem6485/pdf/hf-intro.pdf.
11. Flambaum VV, *Physical Review A*, 2009, **80**, 055401.
12. Hartree DR, *Proceedings of the Cambridge Philosophical Society*, 1928, **24**, 89.
13. Handy NC and Cohen AJ, *Molecular Physics*, 2001, **99**, 403.
14. Kutzelnigg W and Herigonte PV, *Electron correlation at the dawn of the 21s^t century*, Academic Press, 2000.
15. http://molecularmodelingbasics.blogspot.co.uk/2015/06/a-brief-introduction-to-basis-sets.html.
16. Hashimoto T, Hirao K and Tatewaki H, *Chemical Physics Letters*, 1995, **243**, 190.
17. Dunning TH, *Journal of Chemical Physics*, 1989, **90**, 1007.
18. http://www.computationalscience.org/ccce/Lesson2/Notebook 2 Lecture.pdf.
19. Almlof J and Taylor P R, *Journal of Chemical Physics*, 1987, **86**, 4070.
20. Mezei PD, Csonka GI and Ruzsinszky A, *Journal of Chemical Theory and Computation*, 2015, **11**, 3961.
21. Fabiano E and Della Sala F, *Theoretical Chemistry Accounts*, 2012, **131**, 1278.
22. Martin JML, *Chemical Physics Letters*, 1996, **259**, 669.

23. Klopper W, Bak KL, Jorgensen P, Olsen J and Helgaker T, *Journal of Physics B-Atomic Molecular and Optical Physics*, 1999, **32**, R103.
24. Bak KL, Jorgensen P, Olsen J, Helgaker T and Klopper W, *Journal of Chemical Physics*, 2000, **112**, 9229.
25. Helgaker T, Klopper W, Koch H and Noga J, *Journal of Chemical Physics*, 1997, **106**, 9639.
26. Halkier A, Helgaker T, Jorgensen P, Klopper W, Koch H, Olsen J and Wilson AK, *Chemical Physics Letters*, 1998, **286**, 243.
27. Kutzelnigg W and Morgan JD, *Journal of Chemical Physics*, 1992, **96**, 4484.
28. Boese AD, Jansen G, Torheyden M, Hofener S and Klopper W, *Physical Chemistry Chemical Physics*, 2011, **13**, 1230.
29. Daza MC, Dobado JA, Molina JM, Salvador P, Duran M and Villaveces JL, *Journal of Chemical Physics*, 1999, **110**, 11806.
30. S. MM and Scheiner S, *Journal of Chemical Physics*, 1986, **84**, 6328.
31. Alagona G, Ghio C, Latajka Z and Tomasi J, *Journal of Physical Chemistry*, 1990, **94**, 2267.
32. Varandas AJC, *Theoretical Chemistry Accounts*, 2008, **119**, 511.
33. Varandas AJC, *International Journal of Quantum Chemistry*, 2011, **111**, 416.
34. Cremer D, *Wiley Interdisciplinary Reviews-Computational Molecular Science*, 2011, **1**, 509.
35. Kutzelnigg W, *International Journal of Quantum Chemistry*, 2009, **109**, 3858.
36. Cremer D and He Z, *Journal of Physical Chemistry*, 1996, **100**, 6173.
37. Kraka E, Gauss J and Cremer D, *Journal of Molecular Structure-Theochem*, 1991, **80**, 95.
38. Gauss J and Cremer D, *Advances in Quantum Chemistry*, 1992, **23**, 205.
39. He Z and Cremer D, *International Journal of Quantum Chemistry*, 1996, **59**, 71.
40. Cremer D and He Z, *Journal of Molecular Structure-Theochem*, 1997, **398**, 7.
41. Lewars EG, *Computational ChemistryIntroduction to the Theory and Applications of Molecular and Quantum Mechanics*, Springer, Dordrecht, 2003.
42. Cerny J and Hobza P, *Physical Chemistry Chemical Physics*, 2007, **9**, 5291.
43. Marchetti O and Werner HJ, *Journal of Physical Chemistry A*, 2009, **113**, 11580.

44. Jurecka P, Sponer J, Cerny J and Hobza P, *Physical Chemistry Chemical Physics*, 2006, **8**, 1985.
45. Grimme S, *Journal of Computational Chemistry*, 2004, **25**, 1463.
46. Gkionis K, Hill JG, Oldfield SP and Platts JA, *Journal of Molecular Modeling*, 2009, **15**, 1051.
47. Marchetti O and Werner HJ, *Physical Chemistry Chemical Physics*, 2008, **10**, 3400.
48. Riley KE, Pitonak M, Jurecka P and Hobza P, *Chemical Reviews*, 2010, **110**, 5023.
49. Olsen J, Jorgensen P, Helgaker T and Christiansen O, *Journal of Chemical Physics*, 2000, **112**, 9736.
50. Handy NC, Knowles PJ and Somasundram K, *Theoretica Chimica Acta*, 1985, **68**, 87.
51. Werner HJ, Knowles P J, Knizia G, Manby FR, Schütz M, Celani P, Korona T, Lindh R, Mitrushenkov A, Rauhut G, Shamasundar KR, Adler TB, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F, Goll E, Hampel C, Hesselmann A, Hetzer G, Hrenar T, Jansen G, Köppl C, Liu Y, Lloyd AW, Mata RA, May AJ, McNicholas SJ, Meyer W, Mura ME, Nicklaß AO, Neill DP, Palmieri P, Peng D, Pflüger K, Pitzer R, Reiher M, Shiozaki T, Still H, Stone AJ, Tarroni R and Thorsteinsson TMW, ***Molpro*** *quantum chemistry package version 2012.1* http://www.molpro.net.
52. Burow AM, Sierka M and Mohamed F, *Journal of Chemical Physics*, 2009, **131**, 214101.
53. Lazarski R, Burow AM and Sierka M, *Journal of Chemical Theory and Computation*, 2015, **11**, 3029.
54. Usvyat D, Maschio L, Manby FR, Casassa S, Schutz M and Pisani C, *Physical Review B*, 2007, **76**, 075101.
55. Werner HJ, Manby FR and Knowles PJ, *Journal of Chemical Physics*, 2003, **118**, 8149.
56. Saebo S, Tong W and Pulay P, *Journal of Chemical Physics*, 1993, **98**, 2170.
57. Reyes A, Tlenkopatchev MA, Fomina L, Guadarrama P and Fomine S, *Journal of Physical Chemistry A*, 2003, **107**, 7027.

58. Klopper W, Manby FR, Ten-No S and Valeev EF, *International Reviews in Physical Chemistry*, 2006, **25**, 427.
59. Ten-no S and Noga J, *Wiley Interdisciplinary Reviews-Computational Molecular Science*, 2012, **2**, 114.
60. Jackson JL and Wyatt RE, *Chemical Physics Letters*, 1970, **4**, 643.
61. http://vergil.chemistry.gatech.edu/notes/quantrev/node28.html.
62. Roothaan CCJ, *Reviews of Modern Physics*, 1951, **23**, 69.
63. Pople JA and Beveridge DL, *J. Chem. Educ.*, 1970, **48**, A116.
64. Jolicard G, Leforestier C and Austin EJ, *Journal of Chemical Physics*, 1988, **88**, 1026.
65. Primas H, *Modern Quantum Chemistry*, O. Sinanoglu Academic Press New York, 1965.
66. Raghavachari K and Anderson JB, *Journal of Physical Chemistry*, 1996, **100**, 12960.
67. Bartlett RJ, *Wiley Interdisciplinary Reviews-Computational Molecular Science*, 2012, **2**, 126.
68. Robinson JB and Knowles PJ, *Journal of Chemical Theory and Computation*, 2012, **8**, 2653.
69. McNamara JP and Hillier IH, *Physical Chemistry Chemical Physics*, 2007, **9**, 2362.
70. Sharma R, McNamara JP, Raju RK, Vincent MA, Hillier IH and Morgado CA, *Physical Chemistry Chemical Physics*, 2008, **10**, 2767.
71. Sponer J, Riley KE and Hobza P, *Physical Chemistry Chemical Physics*, 2008, **10**, 2595.
72. Berka K, Laskowski R, Riley KE, Hobza P and Vondrasek J, *Journal of Chemical Theory and Computation*, 2009, **5**, 982.
73. Kristyan S and Pulay P, *Chemical Physics Letters*, 1994, **229**, 175.
74. Janowski T and Pulay P, *Chemical Physics Letters*, 2007, **447**, 27.
75. Vydrov OA, Wu Q and Van Voorhis T, *Journal of Chemical Physics*, 2008, **129**, 014106.
76. Cohen AJ, Mori-Sanchez P and Yang WT, *Chemical Reviews*, 2012, **112**, 289.
77. Harrison NA, *Computational Materials Science*, 2003, **187**, 45.
78. Grimme S, *Journal of Computational Chemistry*, 2006, **27**, 1787.

79. Zimmerli U, Parrinello M and Koumoutsakos P, *Journal of Chemical Physics*, 2004, **120**, 2693.
80. Hepburn J, Scoles G and Penco R, *Chemical Physics Letters*, 1975, **36**, 451.
81. Hobza P, Mulder F and Sandorfy C, *Journal of the American Chemical Society*, 1981, **103**, 1360.
82. Meijer EJ and Sprik M, *Journal of Chemical Physics*, 1996, **105**, 8684.
83. W. T. M. Mooij, F. B. van Duijneveldt, J. van Duijneveldt-van de Rijdt and B. P. van Eijck, *Journal of Physical Chemistry A*, 1999, **103**, 9872.
84. Jones RO, *Reviews of Modern Physics*, 2015, **87**, 897.
85. Geerlings P, De Proft F and Langenaeker W, *Chemical Reviews*, 2003, **103**, 1793.
86. Kohn W and Sham LJ, *Physical Review*, 1965, **140**, 1133.
87. Leach A R, *Molecular Modelling Principles and Applications*, Prentice Hall, 2001.
88. Rezac J and Hobza P, *Chemical Reviews*, 2016, **116**, 5038.
89. Sinnokrot MO, Valeev EF and Sherrill CD, *Journal of the American Chemical Society*, 2002, **124**, 10887.
90. Koch H, Fernandez B and Christiansen O, *Journal of Chemical Physics*, 1998, **108**, 2784.
91. Hobza P and Sponer J, *Journal of the American Chemical Society*, 2002, **124**, 11802.
92. Tsuzuki S, Honda K, Uchimaru T, Mikami M and Tanabe K, *Journal of the American Chemical Society*, 2002, **124**, 104.
93. Haldar S, Gnanasekaran R and Hobza P, *Physical Chemistry Chemical Physics*, 2015, **17**, 26645.
94. Rezac J, Riley KE and Hobza P, *Journal of Chemical Theory and Computation*, 2011, **7**, 2427.
95. Sponer J, Jurecka P and Hobza P, *Journal of the American Chemical Society*, 2004, **126**, 10142.
96. Gonzalez MA, *Neutrons Et Simulations, Jdn 18*, 2010, **12**, 169.
97. Dubbeldam D, Calero S, Ellis DE and Snurr RQ, *Molecular Simulation*, 2016, **42**, 81.

98. Good fellow JM, *Molecular Dynamics Applications in Molecular Biology*, Macmillan Press Scientific and Medical, 1991.
99. Rappe AK, Casewit CJ, Colwell KS, Goddard WA and Skiff WM, *Journal of the American Chemical Society*, 1992, **114**, 10024.
100. Halgren TA, *Journal of the American Chemical Society*, 1992, **114**, 7827.
101. Gavezzotti A, *Molecular Aggregation:Structure analysis and molecular simulation of crystals and liquids*, Oxford University Press, 2007.
102. Allinger NL, *Journal of the American Chemical Society*, 1977, **99**, 8127.
103. Allinger NL, Yuh YH and Lii JH, *Journal of the American Chemical Society*, 1989, **111**, 8551.
104. Allinger NL, K. Chen and Lii JH, *Journal of Computational Chemistry*, 1996, **17**, 642.
105. Mayo SL, Olafson BD and Goddard WA, *Journal of Physical Chemistry*, 1990, **94**, 8897.
106. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D and Karplus M, *Journal of Physical Chemistry B*, 1998, **102**, 3586.
107. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, T. Fox, J. W. Caldwell and P. A. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179.
108. Oostenbrink C, Villa A, Mark AE and Van Gunsteren WF, *Journal of Computational Chemistry*, 2004, **25**, 1656.
109. Jorgensen WL, Maxwell DS and TiradoRives J, *Journal of the American Chemical Society*, 1996, **118**, 11225.
110. Sun H, *Journal of Physical Chemistry B*, 1998, **102**, 7338.
111. Cieplak P, Dupradeau FY, Duan Y and Wang JM, *Journal of Physics-Condensed Matter*, 2009, **21**, 21.
112. Halgren TA, *Journal of Computational Chemistry*, 1996, **17**, 490.
113. Maple JR, Hwang MJ, Stockfisch TP, Dinur U, Waldman M, Ewig CS and Hagler AT, *Journal of Computational Chemistry*, 1994, **15**, 162.

114. AMBER web site. http://ambermd.org.
115. W. F. van Gunsteren et al. GROMOS web site. http://www.igc.ethz.ch/GROMOS/index.
116. A. D. Mackerell. CHARMM FF parameters. http://mackerell.umaryland.edu/CHARMM_ff_params.html.
117. Guillot B, *Journal of Molecular Liquids*, 2002, **101**, 219.
118. Frenkel D and Smit B, *Understanding Molecular Simulation : From Algorithms to Applications*, Elsevier 2004.
119. Ramachandran KI, Deepa G and Namboori K, *Computational Chemistry and Molecular Modeling: Principles and Applications*, Springer, 2008.
120. Leach A, *Molecular Modelling : Principles and Applications*, Prentice Hall, 2001.
121. Hinchliffe A, *Molecular Modelling for Beginners*, Wiley, 2008.
122. Atkins P and Paula JD, *Physical Chemistry for the Life Sciences*, W.H. Freeman and Company, 2006.
123. Wongsinlatam W, Keawkri S and Remsungnen T, *applied Mathematical Sciences*, 2012, **6**, 4261
124. Lorentz HA, *Ann. Phys*, 1881, **12**, 127.
125. Berthelot D C R, *Hebd. Seances Acad.Sci*, 1898, **126**, 1703.
126. Stone AJ and Tong CS, *Journal of Computational Chemistry*, 1994, **15**, 1377.
127. Delhommelle J and Millie P, *Molecular Physics*, 2001, **99**, 619.
128. Boda D and Henderson D, *Molecular Physics*, 2008, **106**, 2367.
129. Duh DM, Henderson D and Rowley RL, *Molecular Physics*, 1997, **91**, 1143.
130. Rouha M and Nezbeda I, *Molecular Physics*, 2017, **115**, 1191.
131. Hill T L, *Journal of Chemical Physics*, 1948, **16**, 394.
132. Hayes JM, Greer JC and Morton-Blake DA, *Journal of Computational Chemistry*, 2004, **25**, 1953.
133. Maisuradze GG and Thompson DL, *Journal of Physical Chemistry A*, 2003, **107**, 7118.
134. http://www.investopedia.com/terms/l/least-squares-method.asp.
135. Vogt F and Byrd RK, *Analytical Letters*, 2017, **50**, 945.
136. Palmer B J and Anchell J L, *Journal of Physical Chemistry*, 1995, **99**, 12239.

137. Kittrell JR, Mezaki R and Watson CC, *Industrial and Engineering Chemistry*, 1965, **57**, 19.

138. Levenberg K, *Quarterly Journal of Applied Mathmatics*, 1944, **2**, 164.

Chapter 3:

Calculation of adsorption of H_2 with imidazole and another non-covalent interaction

3 Introductions

3.1 introduction of non-covalent interactions and introduction of adsorption of H_2 with imidazole

Noncovalent interactions are a main source of stability for many molecular complexes in nanoscience, biochemistry and materials chemistry[1-3]. Also, they are widespread in chemistry. Generally, non-covalent interactions are forces that are inter- or intra-molecular (i.e. between molecules, or between atoms within the same molecule) in nature, and happen when the separation between the subsystems is bigger than typical range for covalent bonds that equals or slightly more than 2 Å[4], they may be attractive or repulsive forces and they are generally weak forces[5]. So, the adsorbed molecules can be easily removed since the interactions are noncovalent[6].

To understand the nature of non-covalent interactions, we first must return to the definition of covalent bonds. Many of the covalent bonds include the approximately equal contribution of electrons between the two atoms in the bond. Implying that the electronegativity is nearly the same for two atoms, where the electronegativity is the force of an atom in a molecule to attract electrons to itself. Unlike the non-covalent interaction that do not include the sharing of electrons, but it includes more dispersed variations of electromagnetic interaction within a molecule or between molecules[7].

Moreover, there are several classifications of non-covalent interactions, the most popular classification involves two kinds of forces namely; a long-range force and a short-range force[8-10]. Hence the long-range interaction includes three kinds of interaction; London forces (dispersion interaction or van der Waals interaction), Electrostatic interaction, and induction interaction. Additionally, some classifications classify the hydrogen bond as a non-covalent interaction and considers it to be the fourth kind of the long-range interaction, while another classification classifies hydrogen bond under an electrostatic interaction.

Regarding to the short-range interaction, the most important kind of it is an exchange-repulsion. It is a consequence of the Pauli principle that, due to the exclusion of putting

two electrons in an orbital with the same set of quantum numbers, reduces the electrostatic repulsion between electron pairs resulting in attractive term[11-13].

In dispersion interaction, a force caused by attraction of polarized electron clouds. The electron cloud polarization is induced (caused when the electron clouds repel each another), creating neighbouring regions of electron shortage (δ^+) and electron overload (δ^-). Dispersion interactions occur between all compounds and are especially significant in compounds with large polarizability[5, 14-16] It is considered a weak force. While ionic force is a strong force and the attractive force caused by electrostatic (opposite charge) attraction between a cation and an anion.

In induction interaction, the electron cloud of a molecule distorts in reaction to the electric field of another molecule in its neighbourhood, (also known as polarisation) where on one molecule, the induction interaction is the interaction between a permanent multipole with an induced multipole on another[17, 18].

In terms of hydrogen bond, it is considered one of the most important bonds in all chemical cases[14], where in the IUPAC, there is a re-definition of it as "an attractive interaction between a hydrogen atom in a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a functional group in a different molecule or even the same molecule, in which there is evidence of bond construction"[19]. The interaction energies of hydrogen bonds are between -2.4 and -12 kcal/mol[14] (1kcal /mol=4.184 kJ mol^{-1}), and C—H...O or C—H...π interactions may be weaker than that. Regarding to π stacking, it is the interaction between aromatic groups without overlap of π-orbitals[15, 16].

In this work, we have built many systems of noncovalently bound complexes [H_2…benzene, H_2…. imidazole, CO…. imidazole, N_2… imidazole, NH_3…imidazole and H_2O …imidazole], to try to investigate the adsorption of small molecules (H_2, CO, N_2, NH_3, and H_2O) on organic fragments (imidazole and benzene) through high-accuracy electronic structure calculations, with a view to understanding how to carry out calculations of the properties of larger systems, such as metal-organic frameworks (Zeolitic Imidazolate Frameworks (ZIFs)), with controlled errors. We shall also

establish and calibrate a computational protocol for accurately predicting the binding energy and structure of weakly bound complexes.

Firstly, we studied the adsorption of H_2 with imidazole by applying the ab initio methods on the H_2-imidazole system to obtain intermolecular interaction energy.

The H_2-Imidazole system has two main conformations of interest on its potential energy surface (PES): the parallel (P) structure and the perpendicular structure (T)[20, 21] as in Figures (3-1), and we shall establish a protocol for defining the geometry of an approximate reaction coordinate for the absorption of H_2 (parallel and perpendicular) to imidazole.

(a) (b)

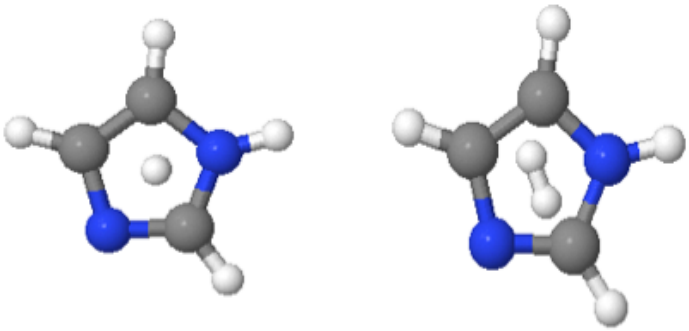
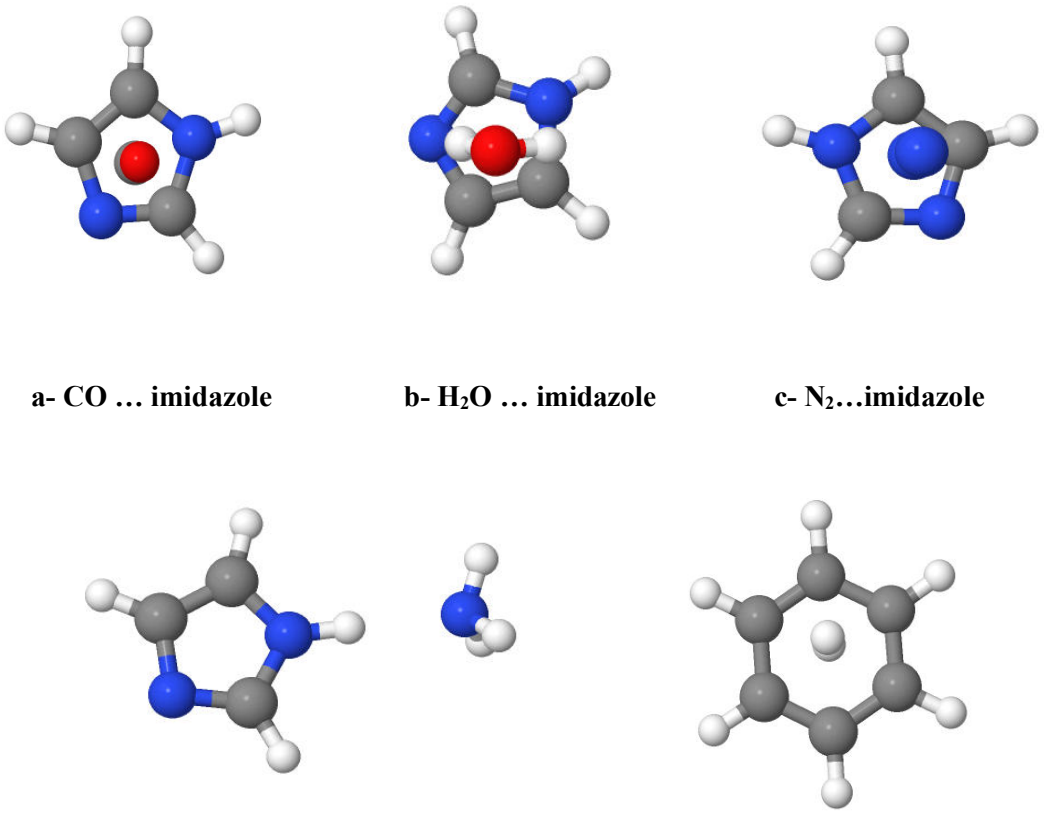


Fig. 3-1: The perpendicular structure (a) and the parallel structure (b) geometry of the H_2-imidazole, where [blue balls are (N) atoms, dark grey balls are (C) atoms and light grey balls are (H) atoms].

Indeed, the interactions between H_2 and ZIFs frameworks are governed by long-range London dispersion terms[20]. So, the nature of the interaction between H_2 and imidazole based on van der Waals forces (London dispersion force) and this, so called physical adsorption (physisorption), and these forces are weak forces and this point as mentioned before in first chapter, was a great point to find a perfect process to store and transport H_2 safely.

Then, we studied the intermolecular interaction energy in several other noncovalently bound complexes. Namely, CO… imidazole, N_2… imidazole, NH_3…imidazole, H_2O…imidazole and H_2…benzene (see Fig 3-2).



a- CO … imidazole **b- H_2O … imidazole** **c- N_2…imidazole**

d- NH_3…imidazole **e- H_2 … benzene**

Fig. 3-2: Structure geometries of noncovalently bound complexes where [blue balls are (N) atoms, dark grey balls are (C) atoms and light grey balls are (H) atoms, red balls are (O)].

In fact, the interactions between the two fragments are governed by van der Waals forces. Where London dispersion forces control H_2…benzene as H_2 imidazole, while dipole-dipole forces control others[22-27]. Furthermore, to understand the nature of interactions between these fragments we shall present some of the properties of them, for example imidazole, it is a highly polar compound where its electric dipole moment is 3.67 D[28] and it has six π-electrons, consisting of a pair of electrons from the lone pair of nitrogen atom and one from each of the residual four atoms of the ring. So, it is classified as aromatic compound. In terms of CO (both O down and C down structures were tested and we found out that when the C down is more stable than when O is down by comparing the energies), H_2O and NH_3, there are differences in the electronegativity between the atoms that form every molecule, and every molecule will have negative and positive partial charges. Whilst H_2 and N_2 are covalent molecules and they show some

degree of local charge separation. When an electrical field (e.g. polar molecule) is present near these molecules where the induction effects will occur there and this effect will due to a separation of these internal charges[29]. Regarding the benzene molecule, every carbon atom in benzene has the same electronegativity. So, the partial charge distribution is identical among the carbon atoms. Moreover the total partial charge of the hydrogen atoms has the same magnitude with opposite sign and the π partial charge is zero for every atoms[30].

3.2 Ab initio calculations

In the calculation, we used ab initio molecular orbital theory, all the ab initio calculations in this work have been carried out using Molpro 2012.1[31]. Potential energy surfaces (PESs) for the parallel and perpendicular configuration of H_2 …imidazole were computed via second order perturbation theory (MP2) (The vdW interaction is usually evaluated by using the second-order perturbation theory[32, 33], coupled cluster with singles, doubles (CCSD), and coupled cluster with singles, doubles, and perturbative triples CCSD(T) methods of theory (triple excitations play a critical role in non-covalent interactions, and when they are not included (e.g.,CCSD method) the accuracy of scheme powerfully declines[34])). Both the correlation-consistent basis sets (cc-pVXZ) and augmented correlation-consistent basis sets (aug-cc-pVXZ) basis sets were used to show the interactions between the H_2 molecule and ZIF frameworks and some properties with controlled errors[35, 36] where X=(D, T, Q, 5 and 6). The geometries of our system were first optimised at the B3LYP/6-311G** level except the centre-to-centre distance (R). For both configurations of H_2…imidazole, the centre-to-centre distance, R, was systematically varied, where the potential energy curves between the two molecules were obtained by changing the intermolecular distance. The centre of mass of imidazole was determined and its position marked using a dummy atom and all calculations were performed with the R=2.5 Å and this geometry was allowed to vary in the calculations as mentioned. In addition, also, the isolated molecular geometries of a H_2 molecule and imidazole (as monomers) were optimized by the LMP2/aug-cc-pV5Z.

The basis set superposition error (BSSE)[37-39] that comes from the use of an incomplete basis set was corrected by the counterpoise (CP) method of Boys and Bernardi[20, 37, 40].

Also, the slow convergence of the intermolecular interaction energy, with the size of the basis set has to be taken into consideration. Where it can be handled in two ways: First one, it can be overcome by extrapolation of the energies calculated from a series of augmented correlation-consistent basis sets with increasing *X*-tuple-*z* quality [41, 42], we have selected the aug-cc-pVXZ (X = T and Q) basis sets (MP2/[34] and CCSD(T)/[34]), the two consecutive members of the standard sequence of basis sets that allow for approaching the complete basis set limit by going to higher levels in the sequence as explained that in chapter 2 and as we shall see in the result and discussion section. Second, by using explicitly correlated wave functions[43] ,where we have used the explicitly correlated second-order closed-shell Møller-Plesset perturbation theory (MP2-F12), the explicitly correlated coupled-cluster with singles, doubles, and perturbative triples (CCSD(T))-F12 and the explicitly correlated coupled-cluster with singles, doubles (CCSD-F12) with aug-cc-pVTZ as a basis set.

To account for the effect of triple excitations on the intermolecular interaction energies of H_2… imidazole, CCSD(T) potential energy curves were computed using aug-cc-pVQZ basis. Because of the expensive computational cost, it was not possible to obtain CCSD(T) PEC's using the aug-cc-pVQZ basis set. The ΔCCSD(T) correction is computed in an aug-cc-pVQZ basis as

$$\Delta CCSD(T) = E_{CCSD(T)}^{aug-cc-pVQZ} - E_{MP2}^{aug-cc-pVQZ} \qquad (1)$$

This correction is combined with the MP2/aug-cc-pV5Z curves to estimate high-quality CCSD(T)/aug-cc-pV5Z potential energy curves for H_2… imidazole according to the equation

$$E_{CCSD(T)}^{aug-cc-pV5Z} = E_{MP2}^{aug-cc-pV5Z} + \Delta CCSD(T) \qquad (2)$$

Generally, to evaluate calculated energies, we applied Hobza's scheme [44-48] as the following:

$$E_{CCSD(T)}^{big\ basis\ set} = E_{MP2}^{the\ same\ big\ basis\ set} + \Delta CCSD(T) \qquad (3)$$

$$\Delta CCSD(T) = E_{CCSD(T)}^{small\ basis\ set} - E_{MP2}^{the\ same\ basis\ set} \qquad (4)$$

After all these calculations, we have built several systems of noncovalently bound complexes (dispersion-bound systems) [H_2…benzene, CO…. imidazole, N_2… imidazole, NH_3…imidazole and H_2O …imidazole] and optimized geometries of these systems through calculate numerical gradients at MP2/CP level and LMP2 level with different augmented basis sets, then applied all established computational protocol for accurately predicting the binding energy on these weakly bound complexes.

The binding energies of all the complexes were defined as the following,

$$De = E(complex) - E(mo_1) - E\ (mo_2) \qquad (5)$$

where $E(complex)$ is the total electronic energy of the complex and E(mo_1) and E(mo_2) denote the electronic energies of the corresponding subsystems monomer 1 and monomer 2, respectively and chapter 2 gives more details about how $E(complex)$ is obtained in a counterpoise calculation.

Finally, the accurate interaction energies were determined for all complexes at the CCSD(T)/ complete basis set (CBS) limit level was used, and the CCSD(T)/CBS interaction energy was defined as follows[34],

$$\Delta E_{CCSD(T)/CBS} = \Delta E_{MP2/CBS} + (\Delta E_{CCSD(T)}^{\text{small basis set}} - \Delta E_{MP2}^{\text{the same basis set}}) \qquad (6)$$

The MP2/CBS interaction energy was determined by extrapolation (aug-cc-pVTZ - aug-cc-pVQZ basis sets) called MP2[34] as Helgaker and co-workers used[49-51]. The second term in Eq. (2), named the ΔCCSD(T) correction term, calculated as the difference between CCSD(T) and MP2 interaction energies, was calculated in the smaller basis set (aug-cc-pVDZ and aug-cc-pVTZ) and The MP2/CBS (aug-cc-pVDZ - aug-cc-pVTZ basis sets) called MP2/ [23] and the CCSD(T)/CBS (aug-cc-pVDZ - aug-cc-pVTZ basis sets) called CCSD(T)/ [23]as the following equation:

$$\text{CCSD (T)/ [34]} = \text{MP2/ [34]} + \text{CCSD (T)/ [23]} - \text{MP2/ [23]} \qquad (7)$$

In terms of, how we applied the extrapolation to complete the basis set limit for our system to obtain high accuracy values of binding energies, it was through extrapolating

between the energies obtained for two sequential cardinal numbers x= 3 and x+1= 4. The two-point formula gives the extrapolated correction energy[49, 50, 52].

$$E_{x,x+1}^{corr} = \frac{(x+1)^3 E_{x+1}^{corr} - x^3 E_x^{corr}}{(x+1)^3 - x^3} \quad (8)$$

Extrapolated total energies are calculated by adding to the extrapolated correlation energy the Hartree-Fock energy, E_x^{corr}, from the calculation with the larger of the two basis sets used in the extrapolation

$$E_{x,x+1} = E_{x+1}^{HF} + E_{x,x+1}^{corr} \; E_{X,X+1} = E_{x+1}^{HF} + E_{x,x+1}^{corr} \quad (9)$$

Because the extrapolation formula is linear in E_x^{corr} and E_{x+1}^{corr} it can be applied equaly well to the correlation part of energy different across a potential energy surface.

3.3 Results and discussions

3.3.1 H_2 … imidazole

Basis Set Superposition Error. Counterpoise correction in the weakly interacting systems, such as the H_2-imidazole system, leads to a more rapid convergence of interaction energy with respect to the size of basis set, but it can also lead to a larger error for the small basis set. In fact, this behaviour is observed in hydrogen-bonded complexes[53-55]. On the other hand, our study with the H_2-imidazole system shows that the counterpoise-corrected binding energies converge more promptly with respect to the basis set than uncorrected energies[56]. This is exhibited in Figure 3-3, which shows the MP2 potential energy curves for a series of basis sets, both with and without counterpoise correction.

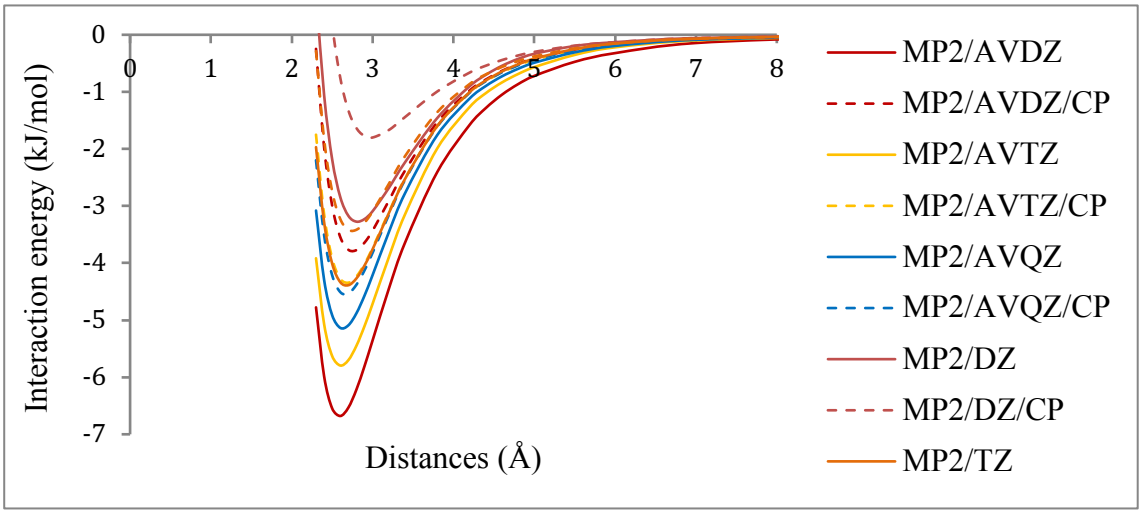


Fig. 3-3: Effect of counterpoise (CP) correction on MP2 potential energy curves for the perpendicular configuration of the H_2-imidazole system.

Additionally, we have found that the interaction energies have decreased significantly when the basis set has been improved[57] (see Tables 3-1 and 3-2).

Perpendicular configuration. The potential energy curves for the perpendicular configuration of the H_2-imidazole system are plotted in Figure 3-3 along with the ΔCCSD(T) correction, where ΔCCSD(T) denotes the difference between CCSD(T) and MP2 at aug-cc-pVQZ basis set.

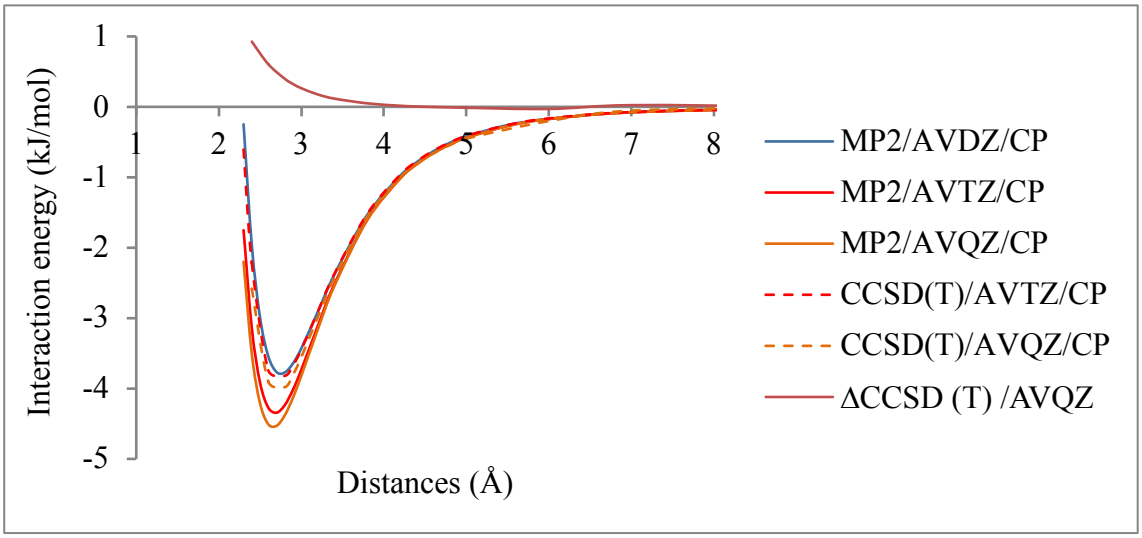


Fig. 3-4: MP2 and CCSD(T) potential energy curves for the perpendicular configuration of the H_2-imidazole system. ΔCCSD(T) denotes the difference between CCSD(T) and MP2 at aug-cc-pVQZ basis set. All results reflect counterpoise correction.

For the MP2 method, the aug-cc-pVTZ and aug-cc-pVQZ curves are close, and they give nearly the same equilibrium distances of (2.69 Å), (2.66 Å) respectively. The aug-cc-pVDZ curve is parallel to the aforementioned curves and gives a slightly larger equilibrium distance of 2.75 Å (see Table 3-1). The aug-cc-pVTZ basis stabilises the system by 0.45 kJ mol^{-1} relative to the much smaller aug-cc-pVDZ basis set, with the difference in interaction energies being larger than 1.0 kJ mol^{-1} at shorter distances (2.3 Å or less). The aug-cc-pVQZ basis stabilises the system by only an additional 0.123 kJ mol^{-1} compared to the aug-cc-pVTZ basis at the corresponding minima and by about 0.15 kJ mol^{-1} at shorter distances.

On the other hand, to better account for electron correlation, the CCSD(T) potential energy curve was computed using the aug-cc-pVTZ basis set, in order to obtain the ΔCCSD(T) correction. It is clear from Figure 3-4 that ΔCCSD(T) is very large at smaller *R (*e.g., ΔCCSD(T)= 0.93 kJ mol^{-1} at 2.4 Å).

Table 3-1: Basis set dependence of binding energies of the H_2-imidazole system (perpendicular).

Basis sets	HF		MP2		CCSD		CCSD(T)	
Without CP	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}
A-2	3.58	-0.6601	2.63	-6.1846	2.48	-5.0076	2.73	-5.8473
A-3	3.62	-0.4563	2.65	-5.4266	2.65	-4.1386	2.42	-4.9493
A-4	3.76	-0.3891	2.59	-5.1163	2.78	-3.5190	2.69	-4.3352
A-5	—	—	2.12	-4.3203	—	—	—	—
2	3.58	-0.5889	2.8	-3.2614	2.93	-2.3955	2.87	-2.83108
3	3.56	-0.5319	2.47	-4.2528	2.79	-3.1120	2.65	-3.75105
4	3.63	-0.3820	2.68	-3.4016	2.89	-2.3467	2.81	-2.8450
5	3.76	-0.3875	2.29	-4.4584	—	—	—	—
6	—	—	2.31	-4.4718	—	—	—	—
With CP	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}
A-2	3.76	-0.3757	275	-3.7563	2.87	-2.7544	2.79	-3.3557
A-3	3.79	-0.3733	2.69	-4.2176	2.80	-3.0671	2.65	-3.8043
A-4	3.65	-0.3652	2.66	-4.3644	2.76	-3.1619	2.59	-3.9277
[34]	3.65	-0.3652	2.70	-4.4860	2.81	-3.2274	2.75	-4.0142
2	3.77	-0.3508	2.99	-1.7872	3.14	-1.2340	3.09	-1.4025
3	3.63	-0.3820	2.68	-3.4016	2.89	-2.3467	2.81	-2.8450
4	3.64	-0.3823	2.49	-3.9952	2.81	-2.8295	2.67	-3.5037
5	3.79	-0.3665	2.39	-4.25305	—	—	—	—

R: the minimum of the potential energy curve along the one-dimensional cut. CP: Counterpoise correction /A-x: denotes the aug-cc-pVXZ basis set where X= D, T, Q, 5, 6. The numbers 2,3,4,5, denote the cc-pVXZ basis set where x=D, T, Q, 5, 6.

Table 3-1 illustrates that the binding energy at CCSD(T) with aug-cc-pVQZ/CP is smaller than the binding energy at CCSD(T) level of theory with extrapolated basis sets[31] by about 0.09 kJ mol^{-1} where they are -3.92, - 4.014 kJ mol^{-1} respectively. Also, the binding energy at CCSD with aug-cc-pVQZ/CP is larger than the binding energy at CCSD(T) level of theory with extrapolated basis sets by about 0.85 kJ mol^{-1}, and that displays the effect of the extrapolation basis sets to converge the correlation energy[52]. In addition, Table 3-1 shows the basis set errors of binding energies, where we can observe that the error is decreasing when we improve size of basis set relative to the extrapolated aug-cc-pVXZ basis set where x=3 and 4. For example, for MP2/CP the error is declined from (0.73 kJ mol^{-1}) at aug-cc-pVDZ to (0.12 kJ mol^{-1}) at aug-cc-pVQZ. Also, for CCSD(T)/CP the error is declined from (0.66 kJ mol^{-1}) at aug-cc-pVDZ to (0.09 kJ mol^{-1}) aug-cc-pVQZ. On the other side, it can be seen that MP2 overestimates the binding energy but when we extrapolate the basis set to complete the basis set limit, we can converge the binding energy.

where, ΔCCSD (T) denotes the difference between CCSD(T) and MP2 at aug-cc-pVTZ

Parallel configuration: The potential energy curves for the parallel configuration of the H_2-imidazole system are plotted in Figure 3-5 along with the ΔCCSD(T) correction basis set.

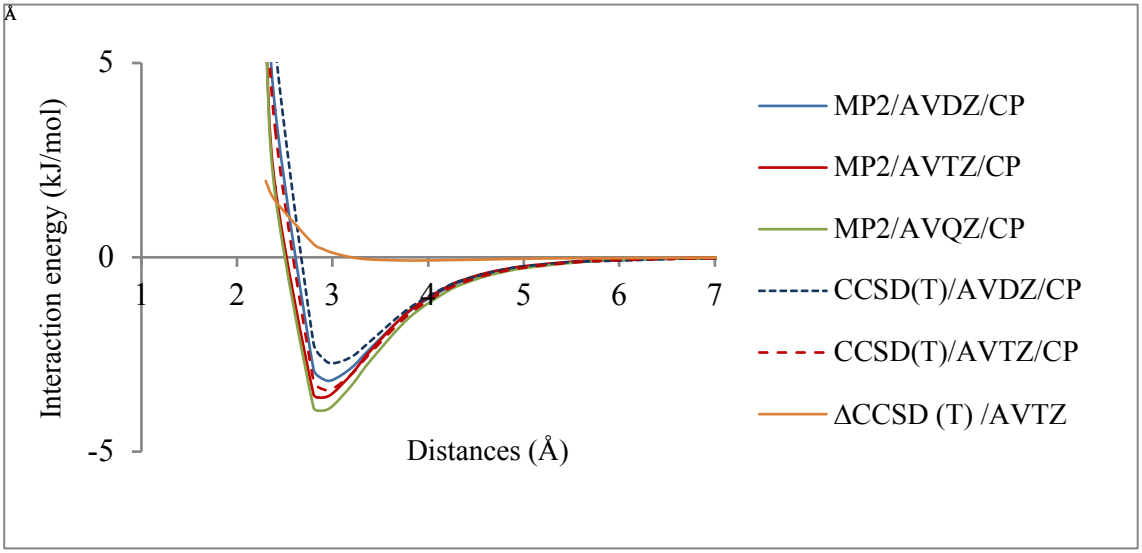


Fig. 3-5: MP2 and CCSD (T) potential energy curves for the parallel configuration of the H_2-imidazole system. ΔCCSD (T) denotes the difference between CCSD (T) and MP2 at aug-cc-pVTZ basis set. All results reflect counterpoise correction.

At the MP2 level, we see that the aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ curves are nearly parallel and give similar equilibrium distances of 2.96, 2.89, and 2.87 Å, respectively (see Table 3-2); in fact, the aug-cc-pVTZ and aug-cc-pVQZ curves are nearly identical and are hard to distinguish in the figure. By examining Figure 3-5 and Table 3-2, we see that at the MP2 level the aug-cc-pVTZ basis stabilises our system by 0.79 kJ mol^{-1} relative to the aug-cc-pVDZ basis at their corresponding minima; the difference in interaction energies is larger at shorter *R*. The aug- cc-pVQZ basis stabilises our system by 0.21 kJ mol^{-1} compared to the aug-cc-pVTZ basis set.

Table 3-2: Basis set dependence of binding energies of the H_2-imidazole system (Parallel).

Basis sets	MP2		CCSD		CCSD(T)	
Without CP	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}
A-2	2.79	-6.1576	2.92	-4.6015	2.85	-5.59993
A-3	2.83	-5.0822	2.96	-3.4260	2.90	-4.4093
A-4	2.83	-4.64451	2.99	—	2.91	—
A-5	2.84	-4.43841	—	—	—	—
2	3.07	-2.0269	3.21	-1.3406	3.13	-1.6488
3	2.90	-3.6066	3.05	-2.2784	2.96	-2.94397
4	2.86	-4.0511	—	—	—	—
5	2.85	-4.2528	—	—	—	—
With CP	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}	R/ Å	E/kJ mol^{-1}
A-2	2.96	-3.1564	3.10	-2.0573	3.04	-2.72238
A-3	2.89	-3.9440	3.04	-2.5651	2.95	-3.38978
A-4	2.87	-4.1541	—	—	—	—
2	3.26	-0.6569	3.56	-0.2838	3.49	-0.42297
3	2.95	-2.7289	3.13	-1.5897	3.05	-2.1474
4	2.90	-3.6080	—	—	—	—
5	2.87	-4.0207	—	—	—	—

CP: Counterpoise correction /A-x: denotes the aug-cc-pVXZ basis set where x= D, T, Q, 5,6, the numbers 2,3,4,5, denote the cc-pVXZ basis set where x=D, T, Q, 5, 6.

The equilibrium distances are 3.04 and 2.95 Å at the CCSD(T)/aug-cc-pVDZ and the CCSD(T)/ aug-cc-pVTZ levels of theory, respectively. The difference between the CCSD(T) and the MP2 equilibrium geometries is in agreement with the trend observed with the perpendicular structure, where the CCSD(T) equilibrium distances were found to be 0.1-0.2 Å larger than the MP2 ones. The ΔCCSD(T) correction is large for *R* smaller than the equilibrium distance (e.g., ΔCCSD (T) is about 2 kJ mol^{-1} at *R* = 2.3 Å). ΔCCSD(T) is 2 kJ mol^{-1}, compared with a much smaller value of 0.9 kJ mol^{-1} for the perpendicular configuration of our system.

The majority of our results agree with the previous study in terms of the effect of counterpoise (CP) correction on MP2 potential energy curves for the (perpendicular and parallel) configuration of the H_2-imidazole system and the effect of improved basis sets but that study was for Benzene dimer[53].

On the other hand, we have applied the explicitly correlated MP2-F12 to calculate interaction energies for our system to obtain high accuracy values of energies through reducing the basis set super position error (BBSE) (see Fig. 3-6).

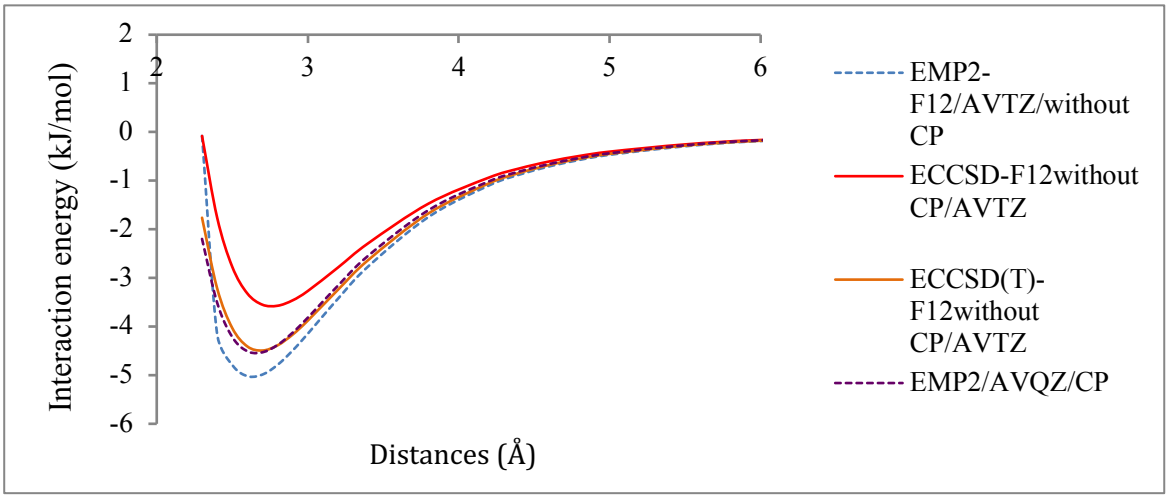


Fig. 3-6: MP2 and MP2-F12 interaction energies (kJ mol^{-1}) for the H_2-imidazole system. All results reflect counterpoise correction.

It is clear from Figure 3-6 that the MP2 results converge very slowly with increasing basis sets size. In contrast, the MP2-F12 calculation gets good and accurate values with a small basis set. We used the aug-cc-pVTZ as a basis set with MP2-F12 and we have found that the interaction energy is much more accurate than the MP2/aug-cc-pVQZ results where the MP2-F12/AVTZ stabilises the system by 0.64 kJ mol^{-1} relative to the MP2/AVQZ at their corresponding minima (R=2.6 Å)[58]. While we obtained nearby energies in case of using MP2/AVQZ/CP and CCDS(T)-F12/AVTZ/without CP, where the difference between them about 0.099 kJ mol^{-1} at equilibrium distance (R=2.6 Å).

Generally, parallel hydrogen position has a higher energy than the perpendicular orientation that leads to an underestimation of the amount of H_2 adsorbed and that low level of description of the complex adsorption process. On the other hand, using the perpendicular position of hydrogen has a lower energy that leads to a high description of the amount of H_2 adsorbed as in the work of Assfour and co-workers[20]. Additionally, we have confirmed that by using a high level of basis sets at MP2 as cc-pVXZ (x= Q, 5, 6) and aug- cc-pVXZ (x=D, T, Q, 5, 6) and by using the same basis sets at CCSD and CCSD(T) as the high level of theory.

3.3.2 Other noncovalent interaction

Optimization of geometries. after build, the species as mentioned above H_2… imidazole, CO… imidazole, N_2… imidazole, NH_3…imidazole, H_2O…imidazole and H_2…Benzene we optimized their geometries using numerical gradients at MP2/CP and local MP2 (LMP2) levels. Where the latter method is particularly useful for the calculation of weak intermolecular interactions because the basis set superposition error (BSSE) is basically reduced and counterpoise corrections are usually not required, for more details about that, the correlation basis is limited to the Atomic Orbitals (AO) in the spatial vicinity of the correlated pairs and the occupied orbitals are localized. Hence, distant pairs are either treated at a lower level pair correlation or neglected, so this approach reduces or eliminates the Basis Set Superposition Error (BSSE)[59, 60]. Also, larger basis sets, at least of triple-zeta quality, are usually required to obtain sufficiently accurate results[61, 62]. Moreover, we tried to site every small molecule (H_2, CO, N_2, NH_3 and H_2O) on the top of imidazole or benzene in an attempt to initialize adsorption. The results of our calculation at LMP2/aug-cc-pVDZ showed that it is difficult to site NH_3 and H_2O fragments above of imidazole since they prefer to make hydrogen bonds between them and N_{imi} (see Fig. 3-7 and 3-8) and (see Table 3-3). In terms of NH_3, we observed that structure B is more stable than structure A by about 15.7 kJ mol^{-1}, where the $N\text{-}H_{imi}$…$N\text{-}H_{ammonia}$ interaction is formed, in structure B, NH_3 acts as Lewis base (donate a pair of nonbonding electrons), while in structure A, NH_3 acts as Lewis acid (accept a pair of nonbonding electrons) and N_{imi}…$H\text{-}N_{ammonia}$ interaction is formed. The interaction energies of both complexes A and B are −16.5 kJ $mol^{-1}$and −32.2 kJ mol^{-1} respectively, and the distances between dimers in every complex are 2.3 Å for structure A and 2 Å for B structure.

On the other hand, in case of H_2O … imi LMP2/aug-cc-pVDZ showed that two isomers N_{imi} …$H\text{-}O_{water}$ and $N\text{-}H_{imi}$ … $O\text{-}H_{water}$ are equally stable and the difference in energy is small, about 4.2 kJ mol^{-1}. Where the interaction energies of both complexes are −24.5 kJ mol^{-1} for **A** and −28.7 kJ mol^{-1} for **B** and the distances between dimers in every complex are 2 Å in both cases. Indeed, these results are in agreement with the results obtained by Gonzalez et al and Choi et al[25, 63]. Also, the higher energy (form A) has the

water donating a proton to the nitrogen atom in the imidazole ring, while the lowest energy (form B) matches to the water acting as a proton acceptor[25, 63, 64].

To confirm our result regarding the position of NH_3 and H_2O relative to imidazole we attempted to site these small molecules above the imidazole plane, we have applied M06 method with 6-311G** as a basis set. M06 is considered as a new hybrid meta exchange-correlation functional of density function of theory and it is recommend for application in organometallic and inorganometallic chemistry and for noncovalent interactions[65]. We found the same results that NH_3 and H_2O with imidazole prefer to form hydrogen bonds rather than physical adsorption (London dispersion force).

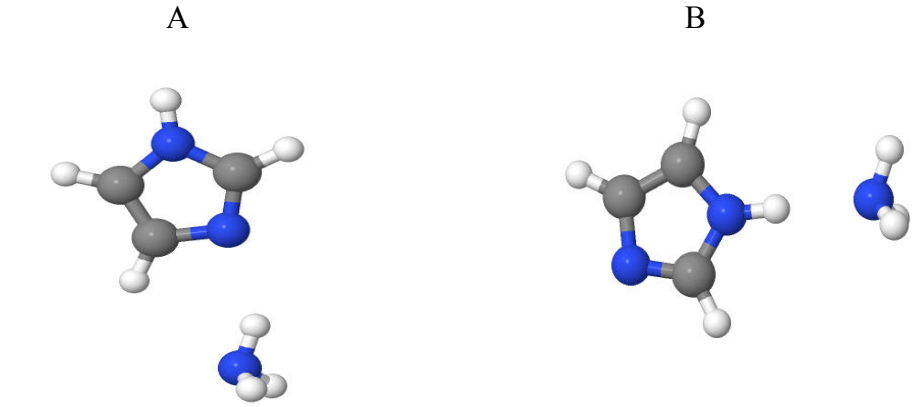


Fig. 3-7: Optimized geometries of NH_3-imidazole at LMP2/aug-cc-pVDZ and M06/ 6-311G**

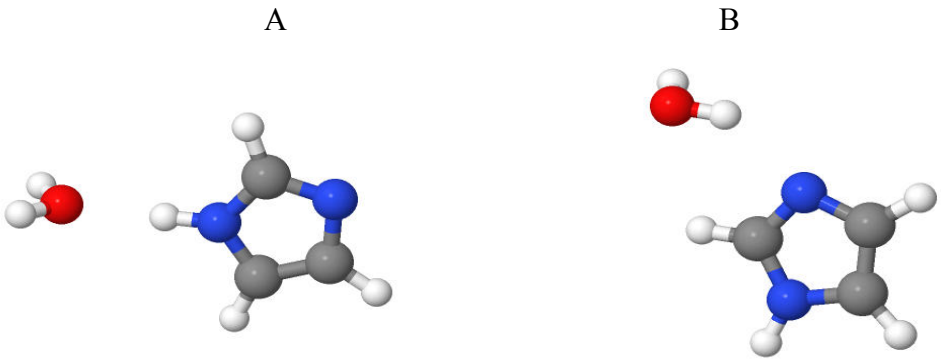


Fig. 3-8: Optimized geometries of H_2O-imidazole at LMP2/aug-cc-pVDZ and M06/ 6-311G**

Table 3-3: Hydrogen bond energies of different isomers for H_2O-imidazole and NH_3-imidazole system.

Orientation of hydrogen bond in complexes	LMP2/aug-cc-pVDZ	M06/6-311G**
A/ N-H_{imi} … O-H_{water}	–24.5 kJ mol^{-1}	-29.4 kJ mol^{-1}
B/ N_{imi} …H-O_{water}	-28.7 kJ mol^{-1}	-38.94 kJ mol^{-1}
A/ N_{imi}…H-$N_{ammonia}$	-16.5 kJ mol^{-1}	-37.39 kJ mol^{-1}
B/ N-H_{imi}…N-$H_{ammonia}$	-32.2 kJ mol^{-1}	-41.21 kJ mol^{-1}

Therefore, these two systems NH_3…imi and H_2O…imi are excluded and we focused on remain system CO…imidazole, N_2…imidazole and H_2…benzene. In addition to our basic system H_2…imidazole. Now we performed MP2/CP and LMP2 level at different size of augmented basis sets to optimize the geometries of previous species (see Table 3-4).

Table 3-4: The binding energy of systems using numerical gradients at MP2/CP level and LMP2 levels. The values of energy are given in hartree.

Monomers/ Complexes	A-2		A-3		A-4	
	MP2/CP	LMP2	MP2/CP	LMP2	MP2/CP	LMP2
H_2	-1.1562	-1.1562	-1.1650	-1.1650	-1.1667	-1.1667
N_2	-109.2805	-109.2797	-109.3646	-109.3639	-109.3936	-109.3932
NH_3	-56.4049	-56.4031	-56.4605	-56.4588	-56.4778	-56.4766
H_2O	-76.2609	-76.2599	-76.3289	-76.3277	-76.3519	-76.3509
CO	-113.0548	-113.0539	-113.1422	-113.1416	-113.1729	-113.1724
Benzene	-231.5396	-230.7279	-231.7443	-230.7822	-231.8098	-230.7955
Imi	-225.5993	-225.5883	-225.7936	-225.7830	-225.8572	-225.8509
H_2…imi	-0.0014	-0.0016	-0.0016	-0.0023	-0.0017	-50.003
N_2…imi	-0.0016	-0.0022	-0.0021	-0.0032	-0.0022	-0.0033
CO…imi	-0.0013	-0.0023	-0.0019	-0.0028	-0.002	-0.0032
H_2…benzene	-0.0014	-0.7989	-0.0018	-0.9540	-0.0019	-1.0111

CP: Counterpoise correction /A-x: denotes the aug-cc-pVXZ basis set where X= D, T, Q, 5, 6. The numbers 2, 3, 4, 5, denote the cc-pVXZ basis set where x=D, T, Q, 5, 6.

Table 3-4 shows that to solve the basis set superposition error (BSSE) for weak intermolecular interaction (van der Waals force) as our systems, LMP2 method is mostly useful, where we found that the MP2/CP and LMP2 methods yield very similar results at the basis set limit. Also, the convergence of MP2 and LMP2 with increasing size of basis sets is different since the BSSE in the LMP2 is reduced[66].

Our results regarding intermolecular interaction energies of previous systems are shown in Table 3-5, where the rows 2-9 show the interaction energies obtained from MP2 and CCSD(T) methods at different size of augmented basis sets and with and without counterpoise (CP). The CCSD(T)/CP interaction energies are mostly larger than the MP2 ones and the differences are not negligible. The MP2 method significantly overbinds compared to CCSD(T)[55, 56, 67, 68].

At CCSD(T)/aug-cc-pVTZ level with counterpoise correction, the most strongly bound complex is the H_2 …Benzene with a binding energy of -3.8708 kJ mol^{-1} and the H_2…imi and N_2…imi complexes have similar values of binding energy -3.8228 kJ mol^{-1} and -3.5276 kJ mol^{-1}, respectively. While, the strength of interaction of CO…imi complex is lower than the other complexes -2.9798 kJ mol^{-1}. Contrary to the previous case, the CCSD(T)/CBS/ [23] interaction energies are smaller than the CCSD(T)/ aug-cc-pVTZ values (on average by -0.575 kJ mol^{-1}), with the difference not exceeding -0.589 kJ mol^{-1} and the difference between lower and higher –level calculations is not large. Moreover, the use of the CCSD(T)/ aug-cc-pVTZ level for dispersion-bound complexes are recommended here.

Regarding to the question can extrapolation to the basis set limit be an alternative to the counterpoise correction, the results in table 3-5 have shown that extrapolation to the CBS limit cannot offer an alternative to the counterpoise correction where the differences in the values of binding energies are large so we need to use both techniques together to overcome the BSSE problem. Although extrapolation to the basis set limit is more economical it may help in overcoming the difficulties with BSSE, particularly when more than two fragments are present.

In fact, these observations disagree with the study carried out by Varandas on the helium dimer, where it was shown that extrapolation to the CBS limit can offer an alternative to the counterpoise correction that yields a more accurate potential energy; however the anomalously low binding energy of helium dimer may mean that it does not behave like most other intermolecular systems[69].

In terms of, how accurate are these values, or, in other words, are they already converged? For example, for H_2 …imi system, the accuracy of the present composite scheme was shown that passing to the extrapolation basis sets for MP2/CBS [23]/CP energies changes the resulting CCSD(T)/aug-cc-pVTZ/CP interaction energies by 77.74%, whiles 77.82% and 86.63% are the changes of energies at MP2/CBS [34]/CP and CCSD(T)/CBS [23]/CP respectively. Moreover, MP2/CBS [34] as well as ΔCCSD(T) [CCSD (T)/ [23] – MP2 [23]] energies changes the resulting CCSD(T)/aug-cc-pVTZ/CP interaction energies by 86.74%.

Also, when we used Δ CCSD(T) [CCSD(T)/aug-cc-pVDZ–MP2/ aug-cc-pVDZ] alternative ΔCCSD(T) [CCSD (T)/ [23] – MP2 [23]] the resulting of CCSD(T)/CBS [23] interaction energies are changed by 86.17% and by 86.57% when used ΔCCSD(T) [CCSD(T)/aug-cc-pVTZ–MP2/ aug-cc-pVTZ]. We have investigated the convergence of the present CCSD(T)/CBS [34] composite scheme, specifically of the ΔCCSD(T) correction term, for one of the strongest dispersion-bound complexes, the passing from the aug-cc-pVDZ basis set to the much larger aug-cc-pVTZ basis set, its absolute value increased from 0.475 to 0.495 kJ mol^{-1}.

Table 3-5: Binding energies of noncovalently bound complexes, and evaluation of binding energies. Binding energy is given in kJ mol-1.

Systems	E/H_2…imi		E/N_2…imi		E/CO…imi		E/H_2…benzene	
	CP	CP*	CP	CP*	CP	CP*	CP	CP*
MP2/A-2	-3.8340	-6.3542	-4.8697	-9.4339	-4.4888	-8.8708	-4.0853	-7.7466
MP2/A-3	-4.3186	-5.6783	-5.4325	-7.9469	-4.9947	-7.4399	-4.8528	-6.7869
MP2/A-4	-4.5305	-5.2100	-5.6892	-6.9712	-4.8917	-6.2668	-5.0379	-5.6239
CCSD(T)/A-2	-3.3586	-5.9339	-3.0532	-7.8208	-2.3559	-8.4357	-3.4438	-7.2713
CCSD(T)/A-3	-3.8228	-5.0953	-3.5276	-5.9662	-2.9798	-6.4545	-3.8708	-5.7076
MP2/ [23]	-4.9171	-5.9755	-6.0557	-7.9797	-5.0215	-7.6883	-5.1493	-6.4653
MP2/ [34]	-4.9119	-4.8933	-6.6816	-7.0009	-5.1715	-6.2059	-5.2122	-12.1245
CCSD(T)/ [23]	-4.4123	-21.7532	-4.1136	-30.3736	-3.5310	-31.1967	-4.3826	-8.8899
Estimated energy CCSD (T)/AVQZ = MP2/ [34] + [CCSD (T)/AVDZ– MP2/AVDZ]	-4.4364	-4.4731	-4.8651	-5.3878	-3.0386	-5.7708	-4.5707	-11.6492
Estimated energy CCSD (T)/AVQZ = MP2/ [34] + CCSD (T)/AVTZ– MP2 AVTZ]	-4.4160	-4.3103	-4.7768	-5.0202	-3.1566	-5.2205	-4.2302	-11.0452
Estimated energy CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23])	-4.4071	-20.6709	-4.7396	-29.3948	-3.681	-29.7143	-4.4455	-14.5491

CP: Counterpoise correction/ CP*: without Counterpoise correction. A-x: denotes the aug-cc-pVXZ basis set where X= D, T, Q. The numbers 2,3,4 denote the cc-pVXZ basis set where x=D, T, Q. [23], [34] extrapolation to complete basis set at [aug-cc-pVDZ: aug-cc-pVTZ] and [aug-cc-pVTZ: aug-cc-pVQZ] respectively.

Overall, this scheme CCSD(T)/ [34] = MP2/ [34] + (CCSD(T)/ [23] – MP2 [23]) achieved a most high accurate of interaction energy for CO ...imi. On another hand, this scheme CCSD(T)/ [34] = MP2/ [34] + [CCSD(T)/ aug-cc-pVDZ – MP2/ aug-cc-pVDZ] produced a most high accurate of interaction energy for H_2...imi, N_2 …imi and H_2 …benzene.

3.4 Conclusion

In this chapter, we carried out calculations through high-accuracy electronic structure calculations (MP2, CCSD and CCSD(T)) levels of theory, with controlled errors to investigate the adsorption of small molecules on organic fragments. Also, we established and calibrated a computational protocol for accurately predicting the binding energy and structure of weakly bound complexes. For example, we have built many systems of noncovalently bound complexes [H_2…benzene, H_2…imidazole, CO…imidazole, N_2…imidazole, NH_3…Imidazole and H_2O…imidazole] and we have optimized geometries of these systems through calculated numerical gradients at MP2/CP level and LMP2 level of theory and extrapolated from aug-cc-pVTZ and aug-cc-pVQZ basis sets to evaluate binding energy by using Hobza's scheme to obtain correct interaction energies. The overall of our results were as the following:

i) The parallel hydrogen position has the highest potential energy surface. On the other hand, using the perpendicular position of hydrogen has a lowest potential energy surface so this position very useful to absorb H_2 on imidazole. Additionally, we have confirmed that by using a high level of basis sets at MP2 as cc-pVXZ (x= Q, 5, 6) and aug- cc-pVXZ (x=D, T, Q, 5, 6) and by using the same basis sets at CCSD and CCSD(T) as the high level of theory. Also, it is clear that the binding energies are sensitive to improvement the size of basis sets[55].

ii) The MP2-F12 calculation gets good and accurate values with a small basis set. We used the aug-cc-pVTZ as a basis set with MP2-F12 and we have found that the interaction energy is much more accurate than the MP2/AVQZ results.

iii) LMP2 method is mostly useful, where we found that the MP2/CP and LMP2 methods yield very similar results at the basis set limit. Also, the convergence of MP2 and LMP2 with increasing size of basis sets is different since the BSSE in the LMP2 is reduced.

iv) The extrapolation to the CBS limit cannot offer an alternative to the counterpoise correction where the differences in the values of binding energies are large so we need to use both techniques together to overcome the BSSE problem. Although extrapolation

to the basis set limit is more economical and may help in overcoming the difficulties with BSSE, particularly when more than two fragments are present

v) This scheme CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23]) achieved a most high accurate interaction energy for CO ...imi. On another hand, this scheme CCSD (T)/ [34] = MP2/ [34] + [CCSD (T)/AVDZ– MP2/AVDZ] produced a most high accurate of interaction energy for H_2...imi, N_2 …imi and H_2 …Benzene.

3.5 References

1. Claessens CG and Stoddart JF, *Journal of Physical Organic Chemistry*, 1997, **10**, 254.
2. Fyfe MCT and Stoddart JF, *Accounts of Chemical Research*, 1997, **30**, 393.
3. Kannan N and Vishveshwara S, *Protein Engineering*, 2000, **13**, 753.
4. Cerny J and Hobza P, *Physical Chemistry Chemical Physics*, 2007, **9**, 5291.
5. http://web.chem.ucla.edu/~harding/ec_tutorials/tutorial76.pdf.
6. Wong BM, *Journal of Computational Chemistry*, 2009, **30**, 51.
7. https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_with_a_Biological_Emphasis_(Soderberg)/Chapter_02%3A_Introduction_to_organic_structure_and_bonding_II/2.4%3A_Non-covalent_interactions.
8. Buckingham AD, Claverie P, Rein R, Schuster P and Pullman B, *Intermolecular Interactions: From Diatomics to Biopolymers*, John Wiley & Sons: ed., Chichester, UK, 1978.
9. Maitland GC, Rigby M, Smith EB and Wakeham WA, *Intermolecular Forces: Their Origin and Determination* Clarendon Press ed., Oxford, UK, 1987.
10. Stone AJ, *The Theory of Intermolecular Forces* Clarendon Press ed., Oxford, UK 2000.
11. Mulliken RS, *Journal of the American Chemical Society*, 1950, **72**, 600.
12. Mulliken RS, *Journal of the American Chemical Society*, 1952, **74**, 811.
13. Mulliken RS, *Journal of Physical Chemistry*, 1952, **56**, 801.
14. Alkorta I, Rozas I and Elguero J, *Chemical Society Reviews*, 1998, **27**, 163.
15. Waller MP, Robertazzi A, Platts JA, Hibbs DE and Williams PA, *Journal of Computational Chemistry*, 2006, **27**, 491.
16. Grimme S, *Angewandte Chemie-International Edition*, 2008, **47**, 3430.
17. Stone AJ, *The Theory of Intermolecular Forces*, Oxford University Press, Oxford 2013.
18. *Buckingham AD. In Intermolecular Forces;Hirschfelder JO., Ed.; Advances in Chemical Physics; John Wiley & Sons, Inc.,1967; Vol. 12; Chapter Permanent*

and Induced Molecular Moments and Long-Range Intermolecular Forces, pp 107–142, DOI: 10.1002/9780470143582.ch2.

19. Arunan E, Desiraju GR, Klein RA, Sadlej J, Scheiner S, Alkorta I, Clary DC, Crabtree RH, Dannenberg JJ, Hobza P, Kjaergaard HG, Legon AC, Mennucci B and N. DJ, http://media.iupac.org/reports/provisional/abstract11/arunan_prs.pdf.
20. Assfour B, Leoni S, Yurchenko S and Seifert G, *International Journal of Hydrogen Energy*, 2011, **36**, 6005.
21. Buckingham AD and Fowler PW, *Canadian Journal of Chemistry-Revue Canadienne De Chimie*, 1985, **63**, 2018.
22. Sirjoosingh A, Alavi S and Woo TK, *Journal of Physical Chemistry C*, 2010, **114**, 2171.
23. Perez-Pellitero J, Amrouche H, Siperstein FR, Pirngruber G, Nieto-Draghi C, Chaplais G, Simon-Masseron A, Bazer-Bachi D, Peralta D and Bats N, *Chemistry-a European Journal*, 2010, **16**, 1560.
24. Jagoda-Cwiklik B, Slavicek P, Cwiklik L, Nolting D, Winter B and Jungwirth P, *Journal of Physical Chemistry A*, 2008, **112**, 3499.
25. Gonzalez MM, Bravo-Rodriguez K, Suardiaz R, de la Vega JMG, Montero LA, Sanchez-Garcia E and Crespo-Otero R, *Theoretical Chemistry Accounts*, 2015, **134**.
26. Adesokan AA, Chaban GM, Dopfer O and Gerber RB, *Journal of Physical Chemistry A*, 2007, **111**, 7374.
27. Roszak R and Roszak S, *Journal of Molecular Modeling*, 2015, **21**, 28.
28. Christen D, Griffiths JH and Sheridan J, *Zeitschrift Fur Naturforschung Section a-a Journal of Physical Sciences*, 1981, **36**, 1378.
29. https://www.boundless.com/users/235424/textbooks/virtual-textbook-of-organic-chemistry/structure-bonding-1/structure-bonding-31/charge-distribution-in-molecules-117-16002/.
30. https://www.chemaxon.com/marvin-archive/4.1.3/marvin/chemaxon/marvin/help/Charge.html.
31. Werner HJ, Knowles P J, Knizia G, Manby FR, Schütz M, Celani P, Korona T, Lindh R, Mitrushenkov A, Rauhut G, Shamasundar KR, Adler TB, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F,

Goll E, Hampel C, Hesselmann A, Hetzer G, Hrenar T, Jansen G, Köppl C, Liu Y, Lloyd AW, Mata RA, May AJ, McNicholas SJ, Meyer W, Mura ME, Nicklaß AO, Neill DP, Palmieri P, Peng D, Pflüger K, Pitzer R, Reiher M, Shiozaki T, Still H, Stone AJ, Tarroni R and Thorsteinsson TMW, ***Molpro*** *quantum chemistry package version 2012.1* http://www.molpro.net.

32. Okamoto Y and Miyamoto Y, *Journal of Physical Chemistry B*, 2001, **105**, 3470.
33. Zaremba E and Kohn W, *Physical Review B*, 1976, **13**, 2270.
34. Rezac J, Simova L and Hobza P, *Journal of Chemical Theory and Computation*, 2013, **9**, 364.
35. Dhakshinamoorthy A and Garcia H, *Chemical Society Reviews*, 2014, **43**, 5750.
36. Fortenberry RC, Huang XC, Francisco JS, Crawford TD and Lee TJ, 2012, **136**, 234309.
37. van Duijneveldt FB, van Duijneveldt-van de Rijdt JGCM and van Lenthe JH, *Chemical Reviews*, 1994, **94**, 1873.
38. Xantheas SS, *Journal of Chemical Physics*, 1996, **104**, 8821.
39. Simon S, Duran M and Dannenberg JJ, *Journal of Chemical Physics*, 1996, **105**, 11024.
40. Boys SF and Bernardi F, *Molecular Physics*, 1970, **19**, 553.
41. Martin JML, *Chemical Physics Letters*, 1996, **259**, 669.
42. Klopper W, Bak KL, Jorgensen P, Olsen J and Helgaker T, *Journal of Physics B-Atomic Molecular and Optical Physics*, 1999, **32**, R103.
43. Rychlewski J, Kluwer Academic Publishers, Dordrecht, 2003.
44. Haldar S, Gnanasekaran R and Hobza P, *Physical Chemistry Chemical Physics*, 2015, **17**, 26645.
45. Rezac J, Riley KE and Hobza P, *Journal of Chemical Theory and Computation*, 2011, **7**, 2427.
46. Berka K, Laskowski R, Riley KE, Hobza P and Vondrasek J, *Journal of Chemical Theory and Computation*, 2009, **5**, 982.
47. Jurecka P, Sponer J, Cerny J and Hobza P, *Physical Chemistry Chemical Physics*, 2006, **8**, 1985.

48. Sponer J, Jurecka P and Hobza P, *Journal of the American Chemical Society*, 2004, **126**, 10142.
49. Halkier A, Helgaker T, Jorgensen P, Klopper W, Koch H, Olsen J and Wilson AK, *Chemical Physics Letters*, 1998, **286**, 243.
50. Helgaker T, Klopper W, Koch H and Noga J, *Journal of Chemical Physics*, 1997, **106**, 9639.
51. Halkier A, Helgaker T, Jorgensen P, Klopper W and Olsen J, *Chemical Physics Letters*, 1999, **302**, 437.
52. Bak KL, Jorgensen P, Olsen J, Helgaker T and Klopper W, *Journal of Chemical Physics*, 2000, **112**, 9229.
53. Dunning TH, *Journal of Physical Chemistry A*, 2000, **104**, 9062.
54. Halkier A, Klopper W, Helgaker T, Jorgensen P and Taylor PR, *Journal of Chemical Physics*, 1999, **111**, 9157.
55. Sinnokrot MO and Sherrill CD, *Journal of Physical Chemistry A*, 2004, **108**, 10200.
56. Sinnokrot MO, Valeev EF and Sherrill CD, *Journal of the American Chemical Society*, 2002, **124**, 10887.
57. Mardirossian N and H.-G. M, *Journal of Chemical Theory and Computation*, 2013, **9**, 4453.
58. Cársky P, Paldus J and P. J.(Eds.), *Recent Progress in Coupled Cluster Methods:Theory and Applications,Springer*.
59. Gresh N, Kafafi SA, Truchon JF and Salahub DR, *Journal of Computational Chemistry*, 2004, **25**, 823.
60. Saebo S, Tong W and Pulay P, *Journal of Chemical Physics*, 1993, **98**, 2170.
61. Werner HJ, Manby FR and Knowles PJ, *Journal of Chemical Physics*, 2003, **118**, 8149.
62. Schutz M, Hetzer G and Werner HJ, *Journal of Chemical Physics*, 1999, **111**, 5691.
63. Choi MY and Miller RE, *Journal of Physical Chemistry A*, 2006, **110**, 9344.
64. VanBael MK, Smets J, Schone K, Houben L, McCarthy W, Adamowicz L, Nowak MJ and Maes G, *Journal of Physical Chemistry A*, 1997, **101**, 2397.
65. Zhao Y and Truhlar DG, *Theoretical Chemistry Accounts*, 2008, **120**, 215.

66. Kaminsky J, Mata RA, Werner HJ and Jensen F, *Molecular Physics*, 2008, **106**, 1899.
67. Hobza P, Selzle HL and Schlag EW, *Journal of Physical Chemistry*, 1996, **100**, 18790.
68. Jaffe RL and Smith GD, *Journal of Chemical Physics*, 1996, **105**, 2780.
69. Varandas AJC, *Theoretical Chemistry Accounts*, 2008, **119**, 511.

Chapter 4

Application of potential energy surface

4 Introductions

Potential energy surfaces of non-bonded systems have been of great attention during the past few years. Potential energy surfaces (PES) have a very important role in analysis of molecular structures studies and chemical reaction dynamics[1, 2]. The Born-Oppenheimer approximation is used in molecular systems to construct PES, where PES can be constructed based on Born-Oppenheimer approximation which respects that electrons move much faster than nuclei and rests on the fact the nuclei are much more enormous than electrons, so can say that the nucleus are nearly fixed with respect to electron motion. In addition, there are 3 coordinates for each atom which together fully specify the geometry of the molecule, If the coordinate is just one, the surface is called a potential energy curve e.g. Morse potential. The Potential Energy Surface represents a unique potential energy for each geometry of the atoms of the molecules in a chemical reaction. Indeed, to describe the position of the atom in 3-dimensional form requires three coordinates, these coordinates may be x, y and z Cartesian coordinates or r, θ and ϕ in Spherical coordinates or degrees of freedom. Furthermore, to describe the position of the atom by degrees of freedom 3N, both translation and rotation of every part of system can be removed (each with 3 degree of freedom). So, the dimensionality of a PES is

$$3N - 6$$

where N is the number of atoms in the system (assuming non-linear geometries and 3N-5 for linear geometries), and in case of many degrees of freedom the potential energy surface (PES) is called a hypersurface[3, 4].

In terms of application of Potential Energy Surfaces, they are assisting the analysis of chemical reaction dynamics and molecular geometry as mentioned above. The PES can be used to theoretically study properties of structures of collection of atoms, for example, computing the rates of a chemical reaction or finding the minimum energy shape of a molecule[3].

4.1 3-dimensional cuts of the potential surface

In order to evaluate the potential energy surface of H_2 … imi complex, we choose 294 grid points (see fig 4-1) to calculate interaction energy between H_2 and imidazole for all three degrees of freedom. To build this grid we fixed the imidazole molecule and we let hydrogen molecule take 294 positions above the imidazole in perpendicular case, so we have selected 3-dimensional cuts of the potential energy surface, where we have chosen four dummy atoms X_1, X_2, X_3 and X_4 (see fig 4-2). The X_1 is an origin of a coordinate system and the distance between X_1 and N_1 of imidazole where a position of lone pair in the same plane of imidazole equals 7 Angstrom. The angle between $N_1X_1C_2$ =306 degree and the dihedral angle $N_1X_1C_2N_3$ = 0°. In the distance between X_1X_2 = 8.8 Å, it is considered as a Y-axis and it is divided into five distances [0.8,2.8,4.8,5.8,6.8,8.8] Å, while the angle $X_1X_2C_2$ = 45° and the dihedral angle $X_1X_2C_2N_3$ = 0°.

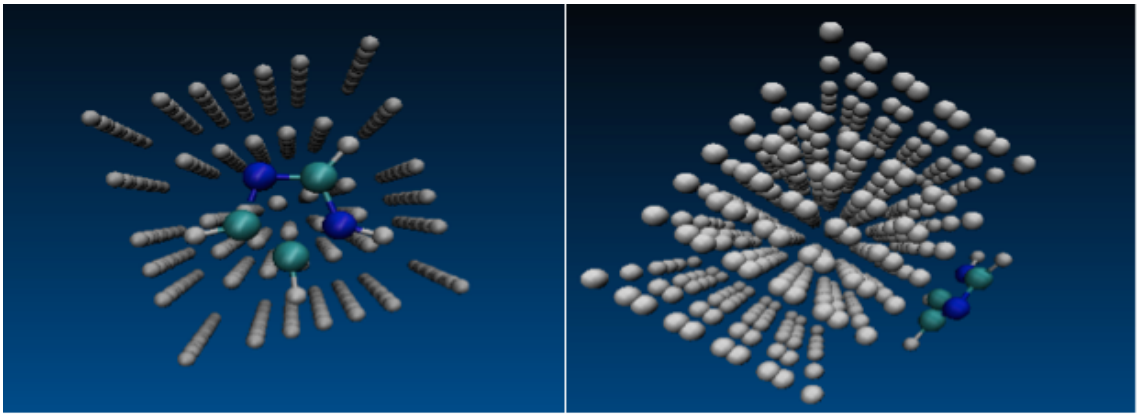


Fig. 4-1: The grid of 294 positions of H_2 in H_2…imidazole system

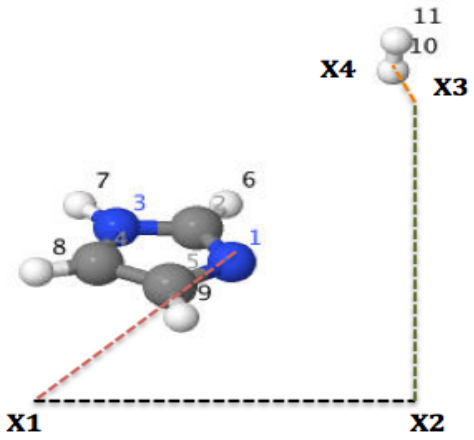


Fig. 4-2: How to build 3-dimensional cuts of the potential energy surface.

The distance between X_2X_3 = 9.4 Å, is considered as an X-axis and it is divided into six distances [1.8,3.7,4.7,5.6,6.5,7.5,9.4] Å and the angle $X_2X_3X_1$ = 90°, and also the dihedral angle $X_2X_3X_1C_2$ = 0°. The Z-axis in this system is the distance between X_3 and X_4 and it equals 6.9 Å. Where X_4 is located in the centre of mass of hydrogen molecule, it has these value [2.5,3,3.6,4,4.7,5.8,6.9] Å. In addition, 90° are the value of both angle $X_4X_3X_2$ and dihedral angle $X_4X_3X_2X_1$.

The intermolecular potential energy between every position of H_2 in the grid and imidazole, were calculated using Møller-Plesset second order perturbation theory (MP2) and single -and double- excitation coupled-cluster theory, with perturbative inclusion of the effected of connected triple excitations, (CCSD(T)). The basis sets used were the augmented correlation-consistent (doublet-triplet) -zeta (aug-cc-pV(D/T)Z) basis set. Also, the extrapolated basis set to complete basis set limit used the following details [MP2/aug-cc-pVDZ:aug-cc-pVTZ], [MP2/aug-cc-pVTZ:aug-cc-pVQZ] and [CCSD(T)/aug-cc-pVDZ:aug-cc-pVTZ] to apply the following scheme of Hobza[5-13] in frame of increasing the accuracy of energy:

$$\text{CCSD (T)/ [34]} = \text{MP2/ [34]} + (\text{CCSD (T)/ [23]} - \text{MP2 [23]}) \quad (1)$$

$$\text{CCSD (T)/ [34]} = \text{MP2/ [34]} + (\text{CCSD (T)/AVTZ} - \text{MP2 AVTZ}) \quad (2)$$

$$\text{CCSD (T)/ [34]} = \text{MP2/ [34]} + (\text{CCSD (T)/AVDZ} - \text{MP2 AVDZ}) \quad (3)$$

Where, [34] and [23] denote the basis set extrapolation using [aug-cc-pVTZ and aug-cc-pVQZ] and [aug-cc-pVDZ and aug-cc-pVTZ] basis sets, respectively. In addition, AVDZ and AVTZ donate the aug-cc-pVDZ and aug-cc-pVTZ basis sets, respectively. The full Counterpoise procedure [14] was employed to correct for basis set superposition error (BSSE). All calculations were carried out using the MOLPRO package[15].

4.2 Force field calculation

In this part of calculation, the intermolecular potential energy between every position (294 points) of H_2 in the grid and imidazole and least square fit of the ab initio potential energy to a standard molecular mechanics potential function, including Lennard-Jones potential were calculated using the Lennard-Jones potential (12-6 LJ) equation

$$E_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \qquad (4)$$

and for pair interaction potential just needs to apply the following formulas

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad , \quad \varepsilon_{ij} = \sqrt{\varepsilon_{ii} * \varepsilon_{jj}} \qquad (5)$$

where, i and j are atoms in our system i = H atoms in hydrogen molecule and j = N, C and H atoms in imidazole molecule, also as mentioned before, ε_{ij}(is the depth of the potential well and a measure of how strongly the two particles attract each other), σ_{ij} (is the finite distance at which the inter-particle potential is zero and a measure of how close two nonbonding particles can get, is thus referred to as the Van der Waals radius and equals to on-half of the distance between nonbonding particles) and r_{ij}(is the distance between the particles and measured from the centre of mass of one particle to the centre of mass of the other particle), and we can obtain r_{ij} by applying this formula

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \qquad (6)$$

where (x_i, y_i, z_i) is the Cartesian coordinates for i (H) atom and (x_j, y_j, z_j) is the Cartesian coordinates for j atoms (N, C and H in imidazole).

The values of these parameters obtain them from Dreiding force field[16, 17], it is worth mention that Dreiding has parameters for all the atoms in the periodic table and for some atoms may be they have more than value of parameters e.g. in our system there are four kinds of atom: nitrogen atom connected to hydrogen atom, carbon, nitrogen atom not connected to hydrogen atom (nitrogen atom with lone pair), and hydrogen. After that, when we obtain the intermolecular interaction energies from equation.4[18, 19], we can fit the energies that obtained from the difference between the energies of ab initio calculation and potential energies of force field for 294 interaction pair potential using a least square fit procedure, followed by calculation of the root-mean square (RMS) deviation as we will explain below in more details.

4.3 parameters estimation and fitting potential procedure

The calculation of molecular properties from the ab initio points requires a fit of the points to a suitable functional form; 294 points were fitted to a very flexible function to take full advantage of the high accuracy of the calculated points. As mentioned above the fitting was by the nonlinear least squares to Lennard-Jones potential function. This is the highest-level ab initio potential available for H_2…imidazole and that in an effort to improve and analyze the ab initio calculated energies, so also, we have fitted the 294 potential energies surface to 12-6 LJ equation. Furthermore, there is need of estimate the parameters and, to do that, we have chosen the nonlinear least squares method, where it is necessary to have realistic initial trial values of the fitting parameters to commence any non-linear least-squares fit[20]. In fact, the root-mean square value of a quantity (e.g. potential energy) is the square root of the mean value of the squared values of the quantity, the formula of RMS as the following:

$$E_{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} E_i^2} \qquad (7)$$

where n is the number of positions of H_2 in the grid and E_i is the potential energy for every position of H_2 in the grid. Hence, in our case, we can calculate the root-mean square error (RMSE) or it is also called the root-mean square deviation (RMSD) using the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_{ab}(r_i) - E_{FF}(r_i))^2} \qquad (8)$$

where, $E_{ab}(r_i)$ and $E_{FF}(r_i)$ are the potential energy of ab initio calculation and force field energy, respectively for every position of H_2 in the grid that is above of the imidazole. However, the better performance of the model when the RMSE value is small[20-22]

To calculate the intermolecular interaction energy by Lennard-Jones equation and fit the potential energy surface and estimate the parameters, we have built the program with python language. Python is a high-level programming language, and it is a widely used general-purpose language[23-26].

This program consists of four parts. The first part reads the molecular properties from the ab initio calculation of 294 position of H_2 in H_2…Imi system. These properties relate to the energies and coordinates and the input file of these properties as shown in the following:

```
 11
* CCSD(T)/AVDZ Energy:   -226.818842128224
N1      0.0000000000      0.0000000000      0.0000000000
C2      0.0000000000      0.0000000000      1.3184930800
N3      0.0000000000      1.2702573738      1.8011036373
C4      0.0000000000      2.1267893489      0.7321879102
C5      0.0000000000      1.3129258556     -0.3726930205
H6      0.0000000000     -0.8695816788      1.9519994063
H7      0.0000000000      1.5353167628      2.7698792978
H8      0.0000000000      3.1927943550      0.8665436326
H9      0.0000000000      1.6060243266     -1.4063543909
H10    -2.6320413650      0.0705626780      2.1511081477
H11    -3.3679586350      0.0705626780      2.1511081477
```

Where, the first line is the number of atoms in the system, and the second line is the kind of method and the kind of basis set and the value of energy. The third line is the name of atom in first column and the Cartesian Coordinates XYZ in the other column, where the second column is the value of X; the third column is the value of Y and the fourth column the value of Z [27-30].

Another property of this system that python program read it is charges and the input file of charges as shown in the following:

```
constant: -226.95839113
N1     -0.39183473
C2      0.23615684
N3     -0.41087212
C4     -0.00080572
C5     -0.03035661
H6      0.12422222
H7      0.24411590
H8      0.12099775
H9      0.11181569
H10     0.00409136
H11    -0.00753058
```

In this file, the second line consists of two columns, the first one is the numbered chemical symbols of atoms (see Fig: 4-3) and another one is about charges that we obtained through applied intrinsic basis bonding analysis (ibba) program in Molpro[15] (see Table 4-1), although we don't need these value of charges because the hydrogen molecule is neutral (no charge). Using ibba can be computed the intrinsic atomic orbital charges (IAO charges) and the intrinsic bond orbitals (IBOs), where IAO charges can be directly explained as the chemical AOs and IBOs provide a reliable method to produce localized orbitals and analyse wave functions[31].

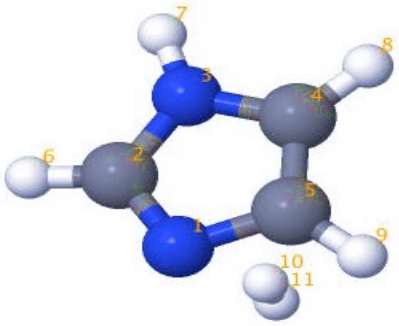


Fig. 4-3: the numbered chemical symbols of atoms, where [blue balls are (N) atoms, dark grey balls are (C) atoms and light grey balls are (H) atoms].

Table 4-1. Basis Set Convergence of calculated the intrinsic atomic orbital charges (IAO charges) by using ibba program. charge is given in e (charge of electron).

Atom	HF/6-311G**	B3LYP/ 6-311G**	DF-HF/AVDZ	B3LPY/ AVDZ	DF-HF/AVTZ	B3LYP /AVTZ	DF-HF /AVQZ	B3LYP /AVQZ
N1	-0.372	-0.325	-0.382	-0.337	-0.385	-0.340	-0.385	-0.340
C2	0.095	0.027	0.097	0.028	0.102	0.031	0.102	0.032
N3	-0.294	-0.236	-0.294	-0.238	-0.298	-0.240	-0.298	-0.241
C4	-0.116	-0.140	-0.121	-0.144	-0.119	-0.142	-0.119	-0.142
C5	-0.076	-0.096	-0.076	-0.096	-0.075	-0.095	-0.074	-0.095
H6	0.160	0.162	0.163	0.166	0.162	0.166	0.162	0.166
H7	0.295	0.290	0.299	0.295	0.299	0.295	0.298	0.295
H8	0.159	0.163	0.161	0.167	0.161	0.166	0.161	0.166
H9	0.152	0.158	0.155	0.162	0.154	0.161	0.154	0.161
H10	0.009	0.005	0.008	0.003	0.007	0.003	0.007	0.003
H11	-0.010	-0.008	-0.009	-0.005	-0.009	-0.005	-0.009	-0.005

HF: Hartree-Fock method. B3LYP: Hybrid functional of density functional theory method (DFT). AVXZ: aug-cc-pVXZ basis sets, where X=D, T, Q.

Table 4-1, shows that IAO charges are insensitive to the employed basis set, and all charges for specific atom are nearly the same at HF/6-311G**, HF/AVXZ, B3LYP/6-311G**, and B3LYP/AVXZ levels of theory, where (X=D, T, Q). For example, the average of charges for N1 is -0.356 and the difference about (0 - 0.01). In addition, the average of charges for C2 is 0.0624 and the difference about (0 - 0.08) and so on. Furthermore, it is clear from columns 6 and 8 that both HF/AVTZ and HF/AVQZ provide the same charges, and also the same result for B3LYP with AVTZ and AVQZ we observed from columns 7 and 9, thus, for the ibba program the AVTZ basis set is recommended and sufficient[15, 31].

With the last properties of the system, the python program reads it and the input file looks as the following:

```
N1   3.6950   0.0002310722
C    3.9800   0.0002310722
N2   3.7950   0.0003426242
H    3.2000   0.00001593601
```

This file consists of three columns first one indicated to the chemical symbols of atoms, second one denoted to Dreiding sigma values of atoms and third one represented Dreiding epsilon values of atoms. The initial values of parameters of these atoms were [ε (N_1: 0.1450, C: 0.1450, N_2: 0.2150, H: 0.010)] kcal mol^{-1} and [σ ((N_1: 3.6950, C: 3.9800, N_2: 3.7950, H: 3.200)] Angstrom. All of these files follow the first part of the python program.

The second part of python program is creating the function with 4 parameters (sigma (σ), epsilon (ε), distance (r) and constant (E)) to calculate force field energies.
The third part of python program is evaluating the potential using the element-indexed parameters plus the globals (charge, asymptote) and calculate the difference between the energies of ab initio calculation that are existing in XYZ files and potential energies of the force field.

The last part of the python program is fitting the energies that obtained from the difference between the energies of ab initio calculation and potential energies of force field using a least square fit procedure, followed by calculation of the root-mean square (RMS) deviation as mentioned above, where it is considered one of the most commonly used measures of the performance indicators or success for numerical prediction.

4.4 Results and discussion

In Table 4-2, the potential energies surface was calculated for different electronic structure methods of theory and basis sets, the results were used to obtain values for the Lennard-Jones parameters (see Table 4-3). The results are summarized in Table 4-2.

Furthermore, In Table 4-3, the Lennard-Jones parameters obtained in the different electronic structure calculations are given for the three directions X, Y and Z (see Figs. 4-1 and 4-2). Where the σ is the zero interaction point and ε denotes the well depth derived from fits to the ab initio potential energy surface at a variety of levels of theory and basis sets so the behaviour of the parameters could be examined. The results are summarized in Table 4-3. The different factors of these calculations are discussed below in more detail.

Table 4-2. Optimized binding energies for H_2 ...imidazole several electronic structure methods of theory and basis sets and associated error (RMS deviation). Binding energy is given in hartree.

Methods/Basis	Binding energy of the lowest point in the grid (ab initio)/hartree	Binding energy of the lowest point in the grid (lennard-jones)/hartree	RMS deviation/hartree
MP2/A-2	-0.00122	-0.00103	0.000198
MP2/A-3	-0.00152	-0.00136	0.000189
MP2/A-4	-0.00161	-0.00145	0.000185
MP2/A-5	-0.00164	-0.00148	0.000184
*LMP2/A-3	-0.00143	-0.00128	0.000189
CCSD(T)/A-2	-0.00101	-0.00089	0.000191
CCSD(T)/A-3	-0.00118	-0.00094	0.000184
*LCCSD(T)/A3	-0.00123	-0.00103	0.000220
MP2/ [23]	-0.00164	-0.00149	0.000185
MP2/ [34]	-0.00167	-0.00151	0.000183
MP2/ [45]	-0.00167	-0.00152	0.000184
CCSD(T) / [23]	-0.00142	-0.00134	0.000180

A-X: aug-cc-pVXZ basis sets, where X=2,3,4,5. *: Without counter poise. [23]: extrapolation basis set [aug-cc-pVDZ: aug-cc-pVTZ], [34]: Extrapolation basis set [aug-cc-pVTZ: aug-cc-pVQZ] and [45]: Extrapolation basis set [aug-cc-pVQZ: aug-cc-pV5Z]. Binding energy = energy of complex (H_2…imidazole) – energy of monomer1 (H_2) – monomer 2 (imidazole).

1. Effects of the computational method and Basis set choice

As expected, the MP2/CCSD(T) calculations shown in Table 4-2 yielded a deep well since most of the van der Waals well originated from the electronic correlation[32-34]. In Table 4-2, for MP2 level of theory with different size of augmented basis sets, it can be observed that ab initio binding energy and Lennard-Jones binding energy decline with increase the size of basis set where larger basis set gives deeper wells. Also, there is an

effect of the calculation method. CCSD(T) binding energy is slightly larger than MP2 at the same basis set. As known MP2 overestimates binding energy[35-37]. In addition, from last column in Table 4-2, it can be seen that root mean square (RMS) deviation for all calculations is stable, only about 0.0002 hartree. This means the RMS deviation is insensitive to the improving of calculation methods and increases the size of basis set.

Table 4-3. Comparison of optimized Lennard- Jones parameters, derived from fits to the ab initio potential energy surface for several electronic structure methods of theory and basis sets. σ and ε given in Å and kJ mol^{-1} respectively.

Methods/Basis set	σ(N1)	σ(C)	σ(N2)	σ(H)	ε(N1)	ε(C)	ε(N2)	ε(H)
*MP2/A-2	1.530	3.079	1.657	3.200	153.26	0.096	26.28	0.1841/ 0.0183
MP2/A-3	1.478	2.558	1.620	3.200	215.8	0.586	30.42	0.0172
MP2/A-4	1.464	2.459	1.595	3.200	235.36	0.911	34.72	0.01690
MP2/A-5	1.462	2.412	1.594	3.200	237.85	1.124	34.48	0.0168
**LMP2/A-3	1.368	2.548	1.419	3.200	194.65	0.731	73.73	0.016145
***CCSD(T)/A-2	1.192	2.705	1.289	2.886/ 3.597	530.27	0.314	106.37	0.0418
***CCSD(T)/A-3	1.140	2.641	1.300	2.886/ 3.568	631.83	0.355	81.00	0.0418
**LCCSD(T)/A-3	1.243	2.689	1.213	2.886/ 3.576	557.09	0.319	216.35	0.0418
MP2/ [23]	1.462	2.409	1.591	3.200	238.75	1.133	35.41	0.01686
MP2/ [34]	1.456	2.390	1.587	3.200	246.49	1.240	36.03	0.01675
MP2/ [45]	1.456	2.387	1.585	3.200	247.2	0.261	36.42	0.01675
CCSD(T) / [23]	1.494	2.641	1.601	3.200	229.71	0.466	38.10	0.01616

A-X: aug-cc-pVXZ basis sets, where X=2,3,4,5. [23]: extrapolation basis set [aug-cc-pVDZ: aug-cc-pVTZ]. [34]: Extrapolation basis set [aug-cc-pVTZ: aug-cc-pVQZ] and [45]: Extrapolation basis set [aug-cc-pVQZ: aug-cc-pV5Z]. *: the value of epsilon of hydrogen is taken from UFF parameters. **: denotes for calculation without counterpoise correction. ***: the value of sigma of hydrogen is taken from UFF parameters.

In Table 4-3, the variation in the values of ε is relatively large for all atoms, unlike the variation in the values of σ, where it can be seen that σ values for all atoms, are insensitive to the size of basis sets and sensitive to the calculation method except for hydrogen that is kept fixed (see rows (2-5), (7-8), (10-12)). There is a reduction in the values of zero interaction point σ. The average of the reduction in the zero-interaction point σ achieved by the MP2 method (with counterpoise correction), with a gradual use of the basis set from aug-cc-pVDZ to aug-cc-pV5Z, is approximately 0.4% for N1, with an average decrease of 5% and 1% in C and N2 respectively and fixed zero interaction point for H_2 for a good fit of the ab initio potential energy where during the optimization procedure when we optimize the hydrogen zero interaction point with other, we obtain very high value of RMS deviation, negative values of σ and ε and

some point of potential energy surface dose not on the curve of fitting. It was also observed that all values of σ and ε were similar for all atoms and via the MP2 method with the aug-cc-pVQZ to aug-cc-pV5Z basis set.

Also, it is clear that, the zero interaction point σ for N1 in CCSD(T) is significantly smaller, by 0.338 and 0.337 Å, respectively, than the MP2 zero interaction point and that is explained that the values of σ is sensitive to the calculation method, where the average of zero interaction point σ is 1.6% smaller in CCSD(T) than in MP2. In the case of C and N2, the average of zero interaction points is smaller by 9% and 22.9% for C and N2 respectively in CCSD(T) than in MP2.

Moreover, in rows (2,3) and (7,8), the zero interaction point σ and well depth ε parameters are shown for the MP2 and CCSD(T) calculations, with the same basis set aug-cc-pVDZ and set aug-cc-pVTZ, and counterpoise correction for BSSE[14] and it is clear, that these values are sensitive to the calculation method, as mentioned above for zero interaction point σ and regarding to the well depth ε parameter, there is an increase in the values of the well depth. For N1, the CCSD(T) well depth is significantly larger, by 377.01 and 541.4 kJ mol^{-1} respectively, than the MP2 well depth. The average well depth is 28.75% larger in CCSD(T) than in MP2. By contrast, the average well depths are larger by 43.95%, 23.25% and 42.47% for C, N2 and H respectively in CCSD(T) than in MP2, considering that the value of the hydrogen well depth is fixed in CCSD(T) to obtain a good fitting as explained above.

Furthermore, in the rows (2-5) of Table 4-3, the improvement in the size of augmented basis set results in an increase in the values of the well depth. The average increase in the well depth values was approximately 87.2% for N1, 53.9% for C and 59% for N2. The well depth values for H were similar, reflecting the weakness of the interaction between hydrogen atoms (guest and host hydrogen).

Additionally, in rows (10-13) of Table 4-3, the effect of applying a basis set extrapolation to complete the basis set limit on the values of zero interaction point σ and well depth ε parameters are shown for MP2[23] (aug-cc-pVDZ:aug-cc-pVTZ), MP2[34] (aug-cc-pVTZ:aug-cc-pVQZ), MP2[45] (aug-cc-pVQZ:aug-cc-pV5Z) and CCSD(T) [23] (aug-cc-pVDZ:aug-cc-pVTZ). Zero interaction point σ values are

insensitive to the size of basis set extrapolation and sensitive to calculation method. While well depth ε values are sensitive to the size of basis set extrapolation and calculation method.

As unexpected, the MP2[34] and MP2[45] calculations with counterpoise correction gave the same result, except for the value of ε for the C atom, which declined from 1.240 kJ mol^{-1} to 0.261 kJ mol^{-1}. A similar result was observed for MP2[23] as for MP2[34] and MP2[45]. Furthermore, an increase in the values of the zero interaction point and a decrease in the well depth values were observed for MP2 and CCSD(T) with [23], except for the well depth of N2 which increased by 92.9%. The values of N1, C and H reduced by 4%, 59% and 4% respectively, while the increases in the zero interaction point σ for N1, C, N2 were 97.9%, 91.2% and 99.4% respectively.

2. Effect of local correlation methods

In rows (6,9) of Table 4-3, the effect of using local correlation methods on the values of the zero interaction point σ and well depth ε parameters are shown for the LMP2 and LCCSD(T) calculations, with the same basis set aug-cc-pVTZ. These values are sensitive to the using local correlation methods. In the case of N1 and N2, there was a decrease in the values of the zero-interaction point, with an increase in the well depth of 9% and 14.5% for the σ values and approximately 34.9% and 34.1% for the ε values, respectively. In contrast, for C, there was an increase in the values of the zero interaction point and a decrease in the well depth of approximately 94.8% for σ values and 56.4% for the ε values.

4.4.1 The estimation of Lennard-Jones parameters and binding energy

Indeed, it is not possible to compute the high order electron correlation as a CCSD(T) correction in a larger basis to calculate the binding energy and Lennard-Jones parameters. Thus, the binding energy and parameters must be estimated by fitting a scheme, such as Hobza's scheme, that depends on adding the difference between the CCSD(T) and MP2 correlation energies evaluated with a small basis set as aug-cc-pVDZ or aug-cc-pVTZ basis sets to the energy of MP2, with the big basis set as shown

in the Table 4-4. The best fit for binding energy (lowest value of binding energy) were obtained when we apply the following schemes:

CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23])

CCSD (T)/ [45] = MP2/ [45] + (CCSD (T)/ [23] – MP2 [23])

where the binding energies for these two equations are the best estimated energies as Table 4-4 shown, with reasonable RMS deviation and variance. More specifically, the binding energies for these schemes were -0.002383 and -0.002353 hartree respectively, approximately 0.000255 hartree lower than the average of the others (-0.002098 hartree) in Table 4-2. Furthermore, in terms of RMS for these schemes, it was also found to be lower (0.1E-05) than the average of others (0.00017865) in Table 4-2.

Table 4-4. Optimized binding energies for H_2 ...imidazole several electronic structure methods of theory and basis sets and associated error (RMS deviation). Binding energy is given in hartree.

Methods/Basis	Binding energy of the lowest point in the grid (ab initio)/hartree	Binding energy of the lowest point in the grid (lennard-jones)/hartree	RMS deviation
a	-0.00130	-0.00120	0.000184
b	-0.00131	-0.00109	0.000176
c	-0.00136	-0.00123	0.000177
d	-0.00142	-0.00133	0.000180
e	-0.00139	-0.00126	0.000176
f	-0.00145	-0.00136	0.000179
g	-0.00142	-0.00129	0.000175
h	-0.00145	-0.00136	0.000179
i	-0.00143	-0.00129	0.000175
j	-0.00145	-0.00136	0.000178
k	-0.00145	-0.00136	0.000179

a: CCSD (T)/A-3 = MP2/ A-3 + [CCSD (T)/A-2– MP2/A-2], **b:** CCSD (T)/A-4 = MP2/ A-4 + [CCSD (T)/A-2– MP2/A-2], **c:** CCSD (T)/A-4 = MP2/ A-4 + [CCSD (T)/A-3– MP2 A-3], **d:** CCSD (T)/A-5 = MP2/ A-5 + CCSD (T)/A-2– MP2 A-2], **e:** CCSD (T)/A-5 = MP2/ A-5 + [CCSD (T)/A-3– MP2 A-3], **f:** CCSD (T)/ [34] = MP2/ [34] + [CCSD (T)/A-2– MP2/A-2], **g:** CCSD (T)/ [34] = MP2/ [34] + CCSD (T)/A-3– MP2 A-3], **h:** CCSD (T)/ [45] = MP2/ [45] + CCSD (T)/A-2– MP2 A-2], **i:** CCSD (T)/ [45] = MP2/ [45] + CCSD (T)/A-3– MP2 A-3], **j:** CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23]), **k:** CCSD (T)/ [45] = MP2/ [45] + (CCSD (T)/ [23] – MP2 [23]). Binding energy = energy of complex (H_2…imidazole) – energy of monomer1 (H_2) – monomer 2 (imidazole).

Regarding to the estimation of Lennard-Jones parameters, it can be summarized in Table 4-5. Zero interaction point σ values are insensitive to estimated scheme. While

the well depth ε values are sensitive to estimated scheme except b scheme where it has especial trend and that because σ of hydrogen does not fix as all schemes.

Table 4-5. Comparison of estimated Lennard- Jones parameters, derived from fits to the ab initio potential energy surface for several electronic structure methods of theory and basis sets. σ and ε given in Å and kJ mol^{-1} respectively.

Methods/Basis set	σ(N1)	σ(C)	σ(N2)	σ(H)	ε(N1)	ε(C)	ε(N2)	ε(H)
a	1.518	2.760	1.645	3.200	190.38	0.283	29.48	0.0167
b	1.179	2.379	1.210	3.527	549.81	1.103	95.49	0.0048
c	1.493	2.763	1.613	3.200	215.03	0.272	34.08	0.01612
d	1.499	2.595	1.619	3.200	214.81	0.573	34.01	0.01629
e	1.491	2.689	1.613	3.200	214.8	0.372	33.7	0.01607
f	1.494	2.570	1.612	3.200	222.59	0.637	35.38	0.01621
g	1.484	2.667	1.605	3.200	226.3	0.407	35.28	0.01596
h	1.493	2.565	1.610	3.200	223.24	0.649	35.77	0.01622
i	1.484	2.662	1.603	3.200	227.0	0.416	35.67	0.01597
j	1.488	2.616	1.597	3.200	237.16	0.521	38.76	0.0161
k	1.487	2.611	1.595	3.200	237.84	0.532	39.17	0.01606

a: CCSD (T)/A-3 = MP2/ A-3 + [CCSD (T)/A-2– MP2/A-2], **b:** CCSD (T)/A-4 = MP2/ A-4 + [CCSD (T)/A-2– MP2/A-2], **c:** CCSD (T)/A-4 = MP2/ A-4 + [CCSD (T)/A-3– MP2 A-3], **d:** CCSD (T)/A-5 = MP2/ A-5 + CCSD (T)/A-2– MP2 A-2], **e:** CCSD (T)/A-5 = MP2/ A-5 + [CCSD (T)/A-3– MP2 A-3], **f:** CCSD (T)/ [34] = MP2/ [34] + [CCSD (T)/A-2– MP2/A-2], **g:** CCSD (T)/ [34] = MP2/ [34] + CCSD (T)/A-3– MP2 A-3], **h:** CCSD (T)/ [45] = MP2/ [45] + CCSD (T)/A-2– MP2 A-2], **i:** CCSD (T)/ [45] = MP2/ [45] + CCSD (T)/A-3– MP2 A-3], **j:** CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23]), **k:** CCSD (T)/ [45] = MP2/ [45] + (CCSD (T)/ [23] – MP2 [23]).

4.4.2 Potential fitting

The 12-6 LJ formula produce unreasonable fit for ab initio calculated potential energy surface PES, for both the equilibrium and attractive regions, as shown in the Fig. 4-4. It is clear that there is slight difference from the fitted 12-6 LJ curve and ab initio PES data in the equilibrium region and attractive region, where Table 4-2 and Table 4-4, show that the average of the difference of binding energy in the equilibrium region is 0.00016 hartree, it is equal about 0.42 kJ mol^{-1}. In contrast, there is good agreement between the two in the repulsive region. To improve this fitting, we tried to do many attempts, the different attempts of the improvement are discussed below in more detail.

Firstly, we tried to apply the 12-6 LJ formula between H_2 molecule (as a quadrupole instead of neutral particles) and imidazole.

Secondly, we tried to apply the 12-6 LJ formula between H_2 molecule (as three sites: H1, H2 and centre of mass of hydrogen molecule) and every atom in imidazole.

Thirdly, we tried to apply the 12-6 LJ formula between H_2 molecule (as three sites: H1, H2 and centre of mass of hydrogen molecule) and every atom in imidazole except 4Hs of imidazole.

Fourthly, we tried to apply the 12-6 LJ formula between H_2 molecule (as two sits: H1and H2) and every atom in imidazole except 4Hs of imidazole.

Unfortunately, all these attempts gave unreasonable fit where the values of σ and ε parameters obtained were negative values. So, we tried to fit the following potential energy equations to the ab initio potential energy surface (PES):

$$V(r) = Ar^{-12} + B\,r^{-6} \quad (9)$$

$$V(r) = Ar^{-12} + B\,r^{-6} + Cr^{-5} \quad (10)$$

$$V(r) = Ar^{-12} + B\,r^{-6} + Cr^{-5} + Dr^{-4} \quad (11)$$

$$V(r) = Ee^{-\alpha r} + B\,r^{-6} + Cr^{-5} + Dr^{-4} \quad (12)$$

$$V(r) = Ee^{-\alpha r} - B\,r^{-6} \quad (13)$$

where A, B, C, D, E and α are constants, r is the distance between H_2 molecule and imidazole in Z direction. This fitting focuses on 22 positions of H_2 molecule above of imidazole in Z direction (not for all positions of the grid). The results of this fitting shown in Fig. (4-5), while the binding energies are shown in Table 4-6.

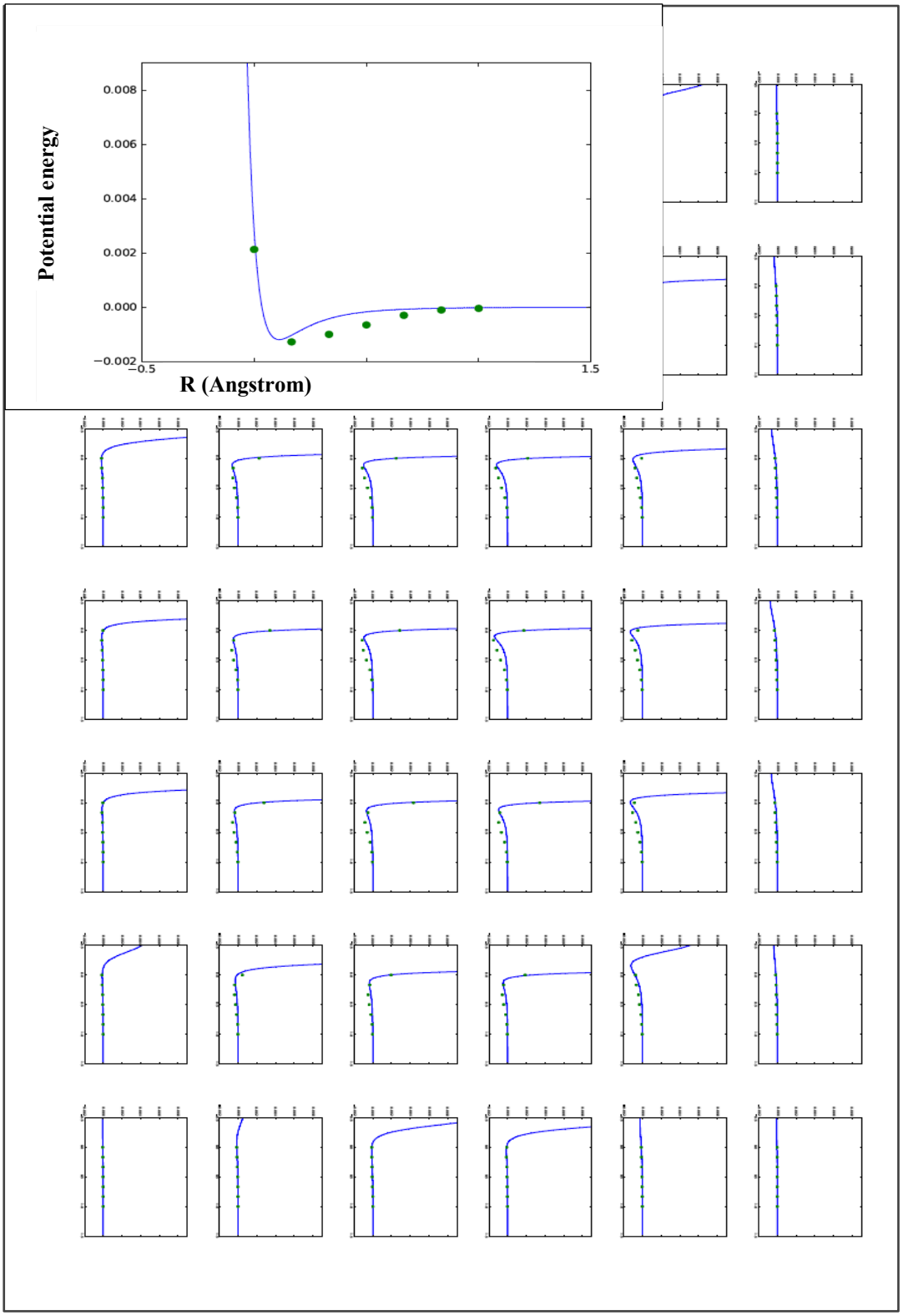


Fig. 4-4: the fitted MP2/aug-cc-pVTZ curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree.

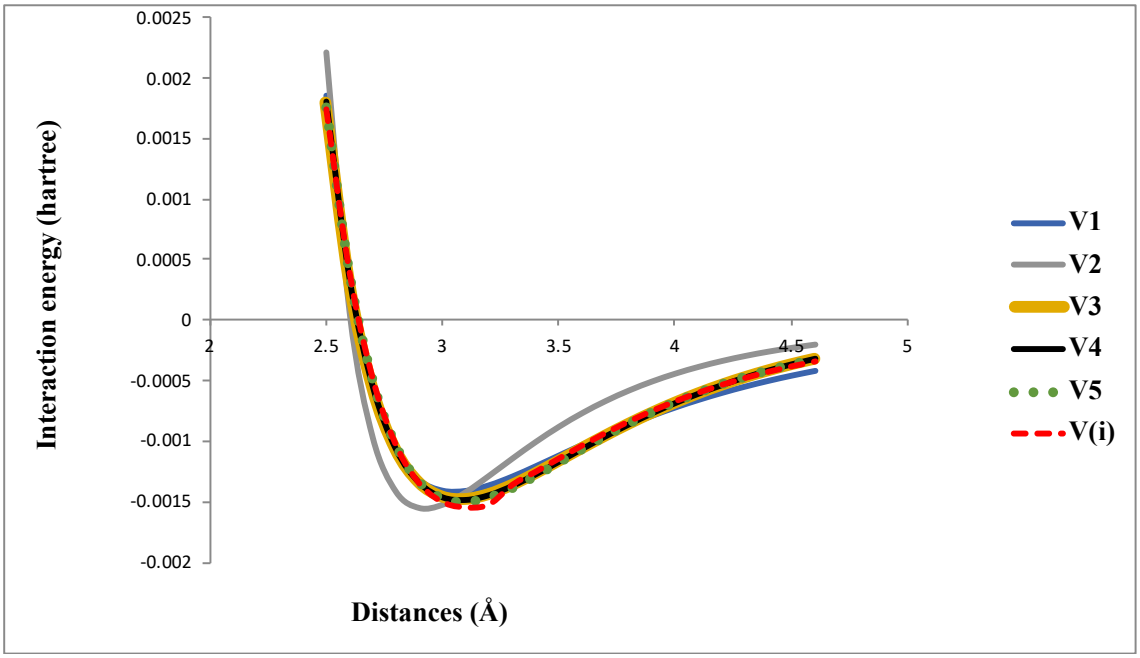


Fig. 4-5: the fitted MP2/aug-cc-pVTZ curve in comparison with the potential energy equations (9-13). V1= Eq. (9), V2 = Eq. (10), V3 = Eq. (11), V4 = Eq. (12), V5 = Eq. (13), V(i): the MP2/aug-cc-pVTZ curve.

Figure 4-5, illustrates that the fitting of MP2/aug-cc-pVTZ curve (as the example of ab initio potential energy surface) to the Eq. (9) that labelled as V1 is nearly the same as the fitting of 12-6 LJ equation where it is clear that the fit in the equilibrium and attractive regions is far from satisfactory, whilst, there is agreement between the two in the repulsive region. The same is true for the H_2-H_2 molecules and H_2-graphite interactions, where the 12-6 LJ potential is inadequate to describe the full range ab initio results covering both the attractive and the repulsive regions[38, 39]. On the other hand, there is good agreement between the MP2/aug-cc-pVTZ curve and Eq. (10) that labelled as V2 in the repulsive region with slight differences in the equilibrium region and attractive region (long range). In terms of fitting of both the Eq. (11) and Eq. (12) equations which labelled as V3 and V4, respectively, it seems that the fitting of these two equations is unreasonable in equilibrium region with good agreement for both the repulsive and attractive regions (short and long ranges). The last fitting was by Eq. (13) that labelled as V5 (it is known as exp-6 LJ) and it has been recently used in the study of hydrogen storage[38, 40]. The green short-dashed line of exp-6 LJ produces an excellent fit for MP2/aug-cc-p VTZ curve, and the exp-6 LJ potential is adequate to describe the full range ab initio results covering both the attractive and the repulsive regions.

Table 4-6.The binding energies of H_2...imidazole and associated error (RMS deviation). Binding energy is given in hartree.

Types of Eqs	Binding energy	Binding energy (ab initio)	RMS deviation/hartree
Eq. (9)	-0.00142	-0.00152	0.0003
Eq. (10)	-0.00141		0.00008
Eq. (11)	-0.00148		0.00004
Eq. (12)	-0.00148		0.00004
Eq. (13)	-0.00150		0.00002

In Table 4-6, the binding energies for the H_2… imidazole system are compared among the Eqs. (9-12) and exp-6 LJ PES (Eq. (13), as well as with the ab initio binding energy. Again, the best fit is only achieved by the exp-6 LJ PES. It is interesting to notice that the simple formula exp-6 LJ is already sufficient for a good fit for the ab initio binding energy, in which the R^{-6} part obviously describes the long-range interaction, and the exponential part describes the short range. It is useful in the future to apply exp-6 LJ PES on the grid for all positions of H_2 molecule above the imidazole and find the best estimated binding energy to use it for molecular dynamic MD and other thermodynamic simulations.

Overall, the recent high-level ab initio results on the interaction between H_2 and imidazole cannot be easily fit to a 12-6 LJ potential, while exp-6 LJ potential can describe both the long range attractive and short range repulsive regions. Such potential should be useful in the future studies on the interaction between H_2 and imidazole materials to store H_2 molecule.

4.5 Conclusion and future recommendations

Ab initio calculations at the MP2/CCSD(T) levels with different basis set, basis set extrapolation and Lennard-Jones potential for the three directions X, Y and Z for 294 positions of H_2 have been performed. Also, we have fitted ab initio binding energy at the MP2/CCSD(T) levels with different basis set and basis set extrapolation to Lennard –Jones binding energy by applying the nonlinear least squares method. Then we estimated the fitted binding energy using Hobza's schemes to reduce the errors. We found that the MP2/CCSD(T) calculations shown yielded a deep well since most of the van der Waals well originated from the electronic correlation. In terms of the improvement in the size of the basis set results in a reduction in the values of the zero interaction point and an increase in the values of the well depth. Also regarding to the basis set extrapolation, we observed that an increase in the values of the zero interaction point and a decrease in the well depth values were observed for MP2 and CCSD(T) with aug-cc-pVDZ: aug-cc-pVTZ, except for the well depth of N2 and the MP2 calculation with counterpoise correction and aug-cc-pVTZ: aug-cc-pVQZ and aug-cc-pVQZ: aug-cc-pV5Z gave the same result, except for the value of ε for the C atom. Furthermore, CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23]) and CCSD (T)/ [45] = MP2/ [45] + (CCSD (T)/ [23] – MP2 [23]) gave the best estimates for binding energy with reasonable RMS deviation where the binding energy of these two schemes is the lowest value obtained it (-0.0024 hartree). On other sides, the 12-6 LJ formula produce unreasonable fit for ab initio calculated potential energy surface PES, for both the equilibrium and attractive regions, as shown in the Fig. 4-4. It is clear that the high-level ab initio results on the interaction between H_2 and imidazole cannot be easily fit to a 12-6 LJ potential. To improve this fitting, we tried to do many attempts and we found when applied many questions of potential that good fit is only achieved by the exp-6 LJ PES (see Table 4-6 and Fig. 4-5).

In the future, it is useful to apply exp-6 LJ PES on the grid for all positions of H_2 molecule above the imidazole and find the best estimated binding energy to use it for molecular dynamic MD and other thermodynamic simulations. Also, it be useful to apply a more complicated formula as exp-6-8-10 potential to fit the potential energy surface and compare the results that obtained from 12-6 LJ, exp-6 LJ and exp-6-8-12 LJ[41] as the study that applied these potential on H_2-graphite and H_2-H_2 molecule where found that the high-level ab initio results on the interaction between H_2 and graphite as modelled by a coronene molecule and H_2-H_2 molecule cannot be easily fit to a 12-6 LJ potential, while both exp-6-8-10 and exp-6 LJ potentials can describe both the short range repulsive regions and the long range attractive, also found that the exp-6-8-10 is a slightly better fit than exp-6 LJ. The formula of exp-6-8-10 is given by the following equation:

$$V(r) = \exp(\alpha - \beta r - \gamma r^2) - \left(\frac{C_6}{r_6} + \frac{C_8}{r_8} + \frac{C_{10}}{r_{10}}\right) \times exp\left[-\left(\frac{r_m}{r} - 1\right)^2\right] \qquad (14)$$

Where α, β and γ are the constants, C_6, C_8 and C_{10} are the coefficients, r is the distance between two fragments and r_m is the positions of potential minimum.

4.6 References

1. https://notendur.hi.is/agust/kennsla/ee10/ees10/Seminars/ees-LongTHESIS-10.pdf.
2. Jalkanen J P, Mahlanen R, Pakkanen T A and Rowley R L, *Journal of Chemical Physics*, 2002, **116**, 1303.
3. https://chem.libretexts.org/LibreTexts/University_of_California_Davis/UCD_Chem_107B%3A_Physical_Chemistry_for_Life_Scientists/Chapters/2%3A_Chemical_Kinetics/2.06%3A_Potential_Energy_Surfaces.
4. http://www.chemistry.gatech.edu/faculty/sherrill/.
5. Rezac J, Simova L and Hobza P, *Journal of Chemical Theory and Computation*, 2013, **9**, 364.
6. Haldar S, Gnanasekaran R and Hobza P, *Physical Chemistry Chemical Physics*, 2015, **17**, 26645.
7. Rezac J, Riley KE and Hobza P, *Journal of Chemical Theory and Computation*, 2011, **7**, 2427.
8. Berka K, Laskowski R, Riley KE, Hobza P and Vondrasek J, *Journal of Chemical Theory and Computation*, 2009, **5**, 982.
9. Jurecka P, Sponer J, Cerny J and Hobza P, *Physical Chemistry Chemical Physics*, 2006, **8**, 1985.
10. Sponer J, Jurecka P and Hobza P, *Journal of the American Chemical Society*, 2004, **126**, 10142.
11. Halkier A, Helgaker T, Jorgensen P, Klopper W, Koch H, Olsen J and Wilson AK, *Chemical Physics Letters*, 1998, **286**, 243.
12. Helgaker T, Klopper W, Koch H and Noga J, *Journal of Chemical Physics*, 1997, **106**, 9639.
13. Halkier A, Helgaker T, Jorgensen P, Klopper W and Olsen J, *Chemical Physics Letters*, 1999, **302**, 437.
14. Boys SF and Bernardi F, *Molecular Physics*, 1970, **19**, 553.
15. Werner HJ, Knowles P J, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O.

Deegan, A. J. Dobbyn, F. Eckert, E. Goll, Hampel, C., A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. Mura, E., A. Nicklaß, O', D. P. Neill, P. Palmieri, D. Peng, K. Pflüger, Pitzer, R., M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson and M. W, ***Molpro*** *quantum chemistry package version 2012.1* http://www.molpro.net.

16. Rappe AK, Casewit CJ, Colwell KS, Goddard WA and Skiff WM, *Journal of the American Chemical Society*, 1992, **114**, 10024.
17. Mayo SL, Olafson BD and Goddard WA, *Journal of Physical Chemistry*, 1990, **94**, 8897.
18. Hart JR and Rappe AK, *Journal of Chemical Physics*, 1992, **97**, 1109.
19. Halgren TA, *Journal of the American Chemical Society*, 1992, **114**, 7827.
20. Li H and Le Roy RJ, *Physical Chemistry Chemical Physics*, 2008, **10**, 4128.
21. Song L, van der Avoird A and Groenenboom GC, *Journal of Physical Chemistry A*, 2013, **117**, 7571.
22. Guilbert C and James TL, *Journal of Chemical Information and Modeling*, 2008, **48**, 1257.
23. *TIOBE Software Index (2011)*. "TIOBE Programming Community Index Python".
24. Nosrati M, *WAP journal*, 2011, **1**, 110.
25. http://redmonk.com/sogrady/2017/06/08/language-rankings-6-17/.
26. http://radar.oreilly.com/2006/08/programming-language-trends.html
27. https://docs.chemaxon.com/display/docs/XYZ+import+and+export+options.
28. http://calistry.org/calculate/xyzviewer.
29. http://www.ccl.net/chemistry/resources/messages/1996/10/21.005-dir/index.html.
30. http://wiki.jmol.org/index.php/File_formats/Formats/XYZ.
31. Knizia G, *Journal of Chemical Theory and Computation*, 2013, **9**, 4834.
32. Palma A, Green S, Defrees D J and McLean A D, *Journal of Chemical Physics*, 1988, **89**, 1401.
33. Chalasinski G, Funk DJ, Simons J and Breckenridge WH, *Journal of Chemical Physics*, 1987, **87**, 3569.
34. Collins JR and Gallup GA, *Chemical Physics Letters*, 1986, **123**, 56.

35. Dunning TH, *Journal of Physical Chemistry A*, 2000, **104**, 9062.
36. Gruneis A, Marsman M and Kresse G, *Journal of Chemical Physics*, 2010, **133**, 074107.
37. He X, Fusti-Molnar L, Cui G L and Merz K M, *Journal of Physical Chemistry B*, 2009, **113**, 5290.
38. Sun DY, Liu JW, Gong XG and Liu ZF, *Physical Review B*, 2007, **75**, 075424.
39. Silvera IF, *Reviews of Modern Physics*, 1980, **52**, 393.
40. Patchkovskii S, Tse JS, Yurchenko SN, Zhechkov L, Heine T and Seifert G, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 10439.
41. Ahlrichs R, Penco R and Scoles G, *Chemical Physics*, 1977, **19**, 119.

Chapter 5
General Conclusions

5 General Conclusions

The aim of the work presented in this thesis was investigating- through high-accuracy electronic structure calculations- the adsorption of H_2 molecule on imidazole as organic fragments, with a view to understanding how to carry out calculations of the properties of larger systems, such as metal-organic frameworks to provide a "good" application to store H_2 molecule conveniently and safely.

The first chapter introduces the overview of H_2 storage, imidazole, Metal Organic Frameworks (MOFs), Zeolitic Imidazole Frameworks (ZIFs) and the interaction between H_2 and Imidazole Frameworks (ZIF) in an attempt to illustrate the environment of studied system.

The second chapter presents the subject of electronic structure theory and some of the most common methods employed in approximate solution of the Schrödinger equation. We then focus on the need for a single reference methods and describe some common single reference methods. Particularly MP2, CCSD and CCSD(T) methods, with gave overview about the basis sets and the errors that produce from using these basis set (BSSE) and shed the light on the appropriate procedure to solve this error by using basis set extrapolation and counterpoise procedure. Also, in this chapter there are introduction about the density function theory method (DFT) and Force field method (Molecular mechanics method) represented in Lennard-Jones parameters and formula. Additionally, there is presenting of increasing the accuracy of calculated intermolecular interaction energies (Composite CCSD(T)/CBS Schemes). The last part in this chapter was simple explanation of the data fitting and error estimation represented in nonlinear least squares method.

The third chapter describes the calculations through high-accuracy electronic structure methods (MP2, CCSD and CCSD(T)), with controlled errors to investigate the adsorption of small molecules on organic fragments. Also, there are established and calibrated computational protocols for accurately predicting the binding energy and structure of weakly bound complexes. For example, some systems of noncovalently bound complexes are built here namely [H_2…benzene, H_2…imidazole, CO…imidazole, N_2…imidazole, NH_3…Imidazole and H_2O…imidazole] and the

geometries of these systems optimized through calculating numerical gradients at MP2/CP level and LMP2 level of theory and extrapolated from aug-cc-pVTZ and aug-cc-pVQZ basis set are calculated to evaluate binding energy by using Hobza's scheme to obtain correct interaction energies. The overall of results were in this chapter as the following: firstly, the parallel hydrogen position has a highest potential energy surface. On the other hand, using the perpendicular position of hydrogen has a lowest potential energy surface. Additionally, that have confirmed by using a high level of basis sets at MP2 as cc-pVXZ (x= Q, 5, 6) and aug- cc-pVXZ (x=D, T, Q, 5, 6) and by using the same basis sets at CCSD and CCSD(T) as the high level of theory. Secondly, the MP2-F12 calculation gets good and accurate values with a small basis set. We used the aug-cc-pVTZ as a basis set with MP2-F12 and we have found that the interaction energy is much more accurate than the MP2/AVQZ results. Thirdly, LMP2 method is useful, where we found that the MP2/CP and LMP2 methods yield very similar results at the basis set limit. Also, the convergence of MP2 and LMP2 with increasing size of basis sets is different since the BSSE in the LMP2 is reduced. Fourthly, this scheme CCSD(T)/ [34] = MP2/ [34] + (CCSD(T)/ [23] – MP2 [23]) achieved a most high accurate of interaction energy for H_2...imi and N_2 …imi. On another hand, this scheme CCSD(T)/ [34] = MP2/ [34] + [CCSD(T)/ aug-cc-pVDZ – MP2/ aug-cc-pVDZ] produced a most high accurate of interaction energy for CO...imi and for H_2 …Benzene.

The fourth chapter introduces the calculations of ab initio potential surface for the rigid imidazole molecule and hydrogen molecule system (H_2 ... imidazole) using the Møller-Plesset perturbation theory and CCSD(T). The potential was calculated in the three physically distinct Cartesian directions X, Y, and Z, the effects of the basis set and the counterpoise correction were examined. Ab initio calculations at the MP2/CCSD(T) levels with different basis set, basis set extrapolation and Lennard-Jones potential for the three directions X, Y and Z for 294 positions of H_2 have been performed. Also, fitting of ab initio binding energy at the MP2/CCSD(T) levels with different basis set and basis set extrapolation to Lennard –Jones binding energy have performed by applying the nonlinear least squares method. Then there was estimating for the fitted binding energy using Hobza's schemes to reduce the errors. The results of these calculations have been discussed in this chapter. The counterpoise corrected interaction energies converge significantly faster and in a smoother manner than the uncorrected

ones. Additionally, the MP2/CCSD(T) calculations shown yielded a deep well since most of the van der Waals well originated from the electronic correlation. In terms of the improvement in the size of the basis set results in a reduction in the values of the zero interaction point and an increase in the values of the well depth. Also regarding to the basis set extrapolation. Moreover, there is an observation that an increase in the values of the zero interaction point and a decrease in the well depth values for MP2 and CCSD(T) with aug-cc-pVDZ:aug-cc-pVTZ, except for the well depth of N2 and the MP2 calculation with counterpoise correction and aug-cc-pVTZ:aug-cc-pVQZ and aug-cc-pVQZ:aug-cc-pV5Z gave the same result, except for the value of ε for the C atom. Furthermore, CCSD (T)/ [34] = MP2/ [34] + (CCSD (T)/ [23] – MP2 [23]) and CCSD (T)/ [45] = MP2/ [45] + (CCSD (T)/ [23] – MP2 [23]) gave the best estimates for binding energy with reasonable RMS deviation where the binding energy of these two schemes is the lowest value obtained it (-0.0024 hartree). On other sides, the 12-6 LJ formula produce unreasonable fit for ab initio calculated potential energy surface PES, for both the equilibrium and attractive regions, as shown in the Fig. 4-4. It is clear that the high-level ab initio results on the interaction between H_2 and imidazole cannot be easily fit to a 12-6 LJ potential. To improve this fitting, we tried to do some attempts and we found when applied these questions of potential that the best fitting is only achieved by the exp-6 LJ PES (see Table 4-6 and Fig. 4-5).

.

Appendix

1. Potential fitting

The 12-6 LJ formula produce unreasonable fit for ab initio calculated potential energy surface PES, for both the equilibrium and attractive regions, as shown in the following Figures:

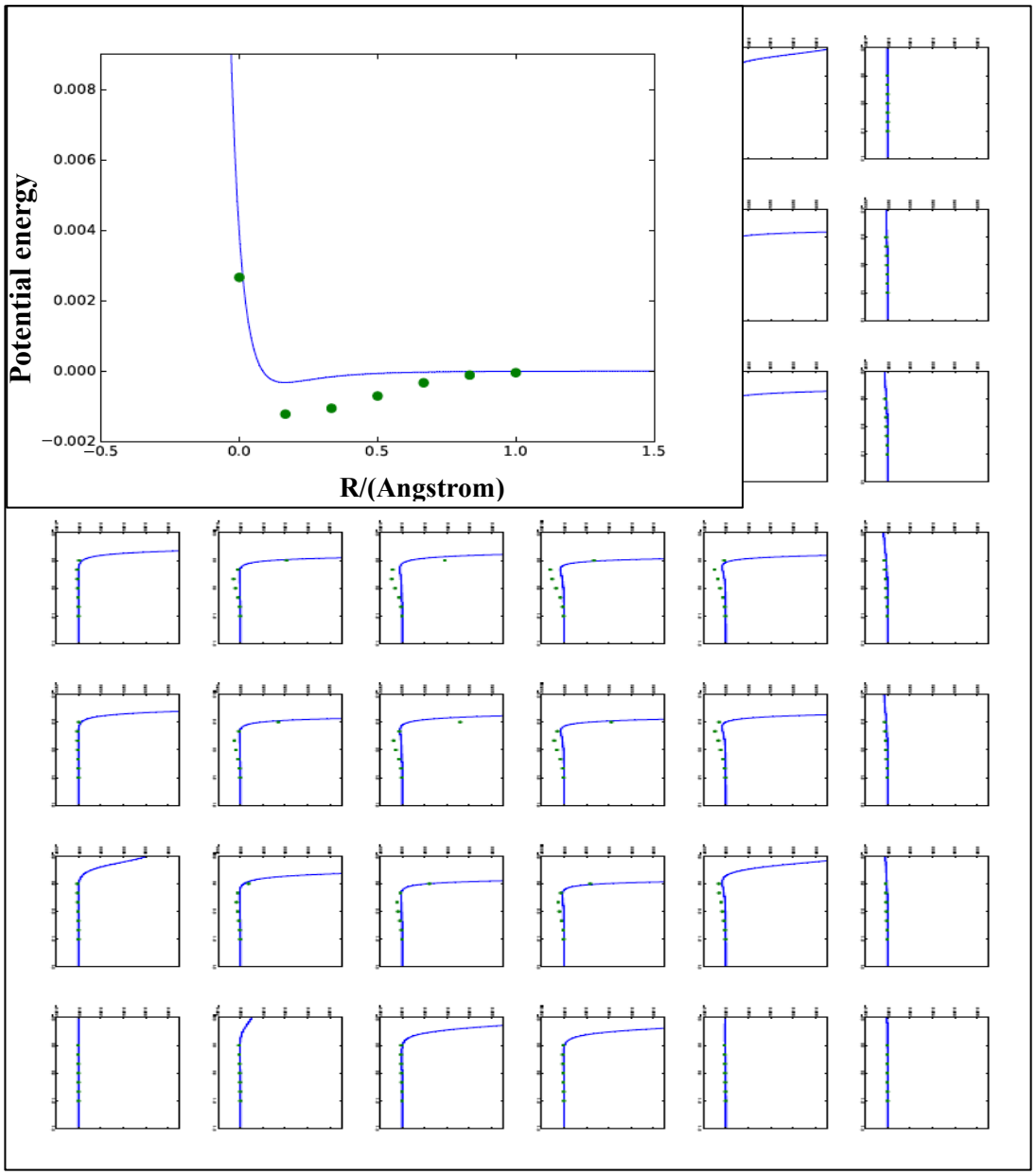


Fig. 1: the fitted MP2/aug-cc-pVDZ curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree.

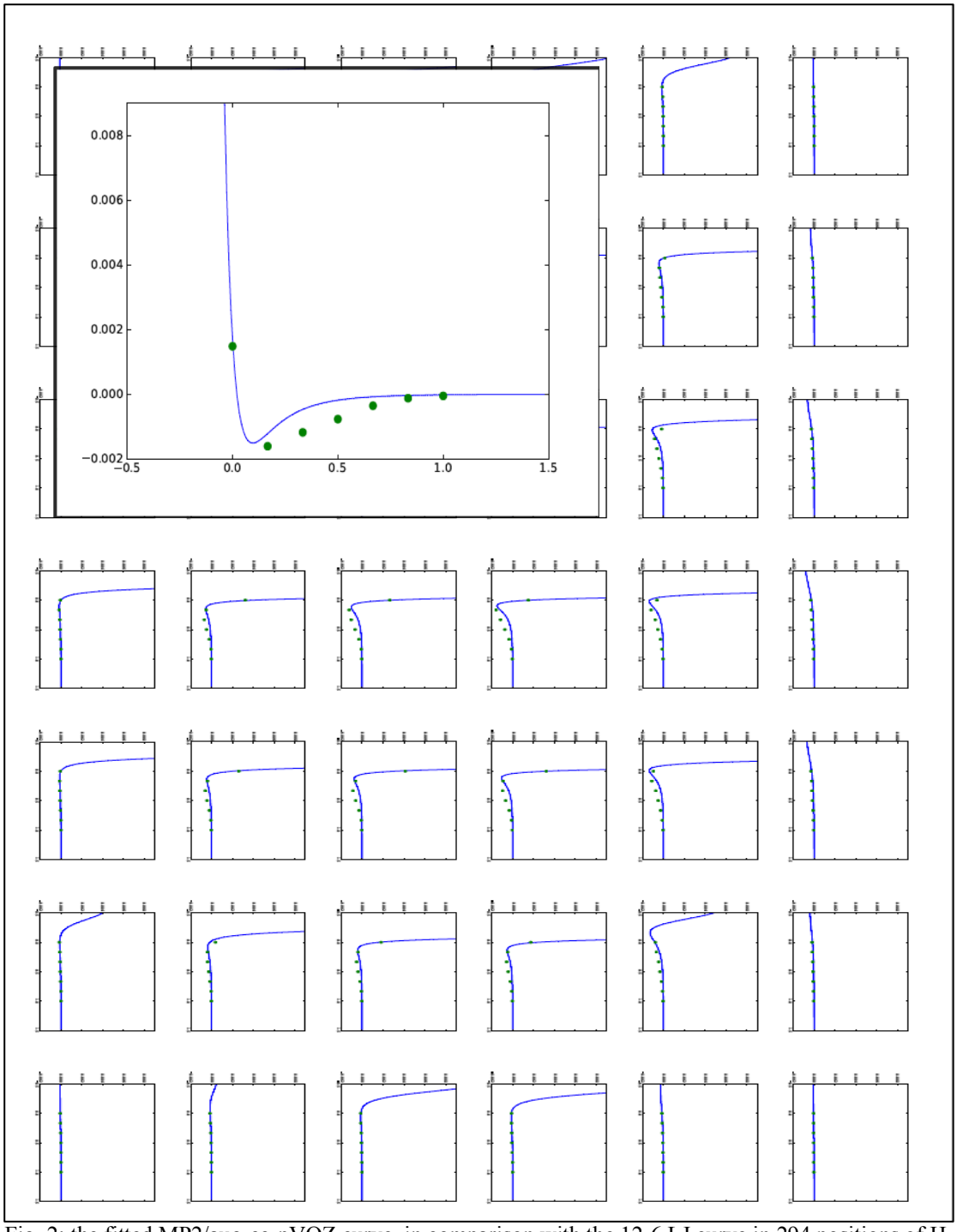


Fig. 2: the fitted MP2/aug-cc-pVQZ curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree. X-axis is R (Angstrom) and Y-axis is potential energy.

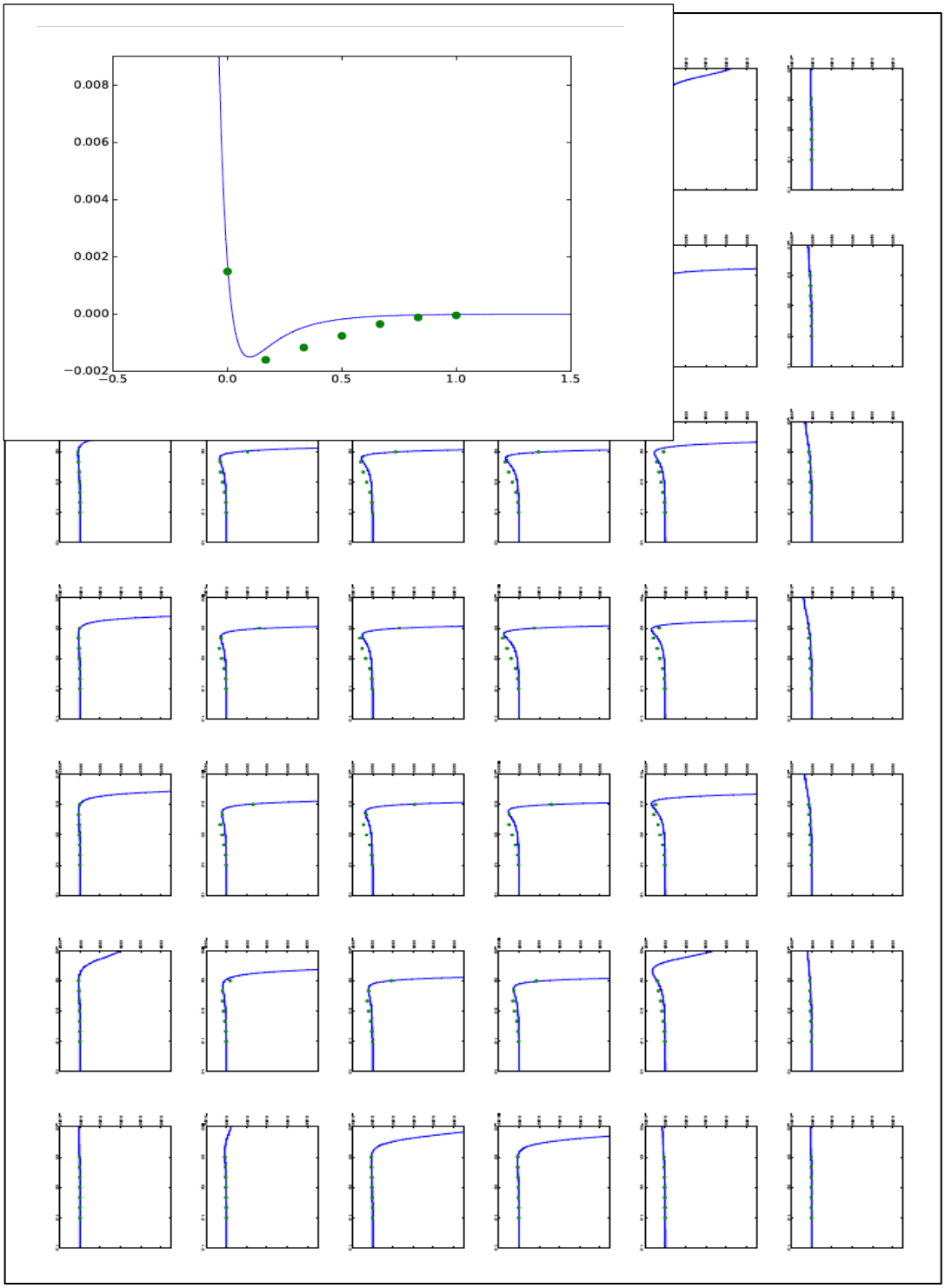


Fig. 3: the fitted MP2/aug-cc-pV5Z curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree. X-axis is R (Angstrom) and Y-axis is potential energy.

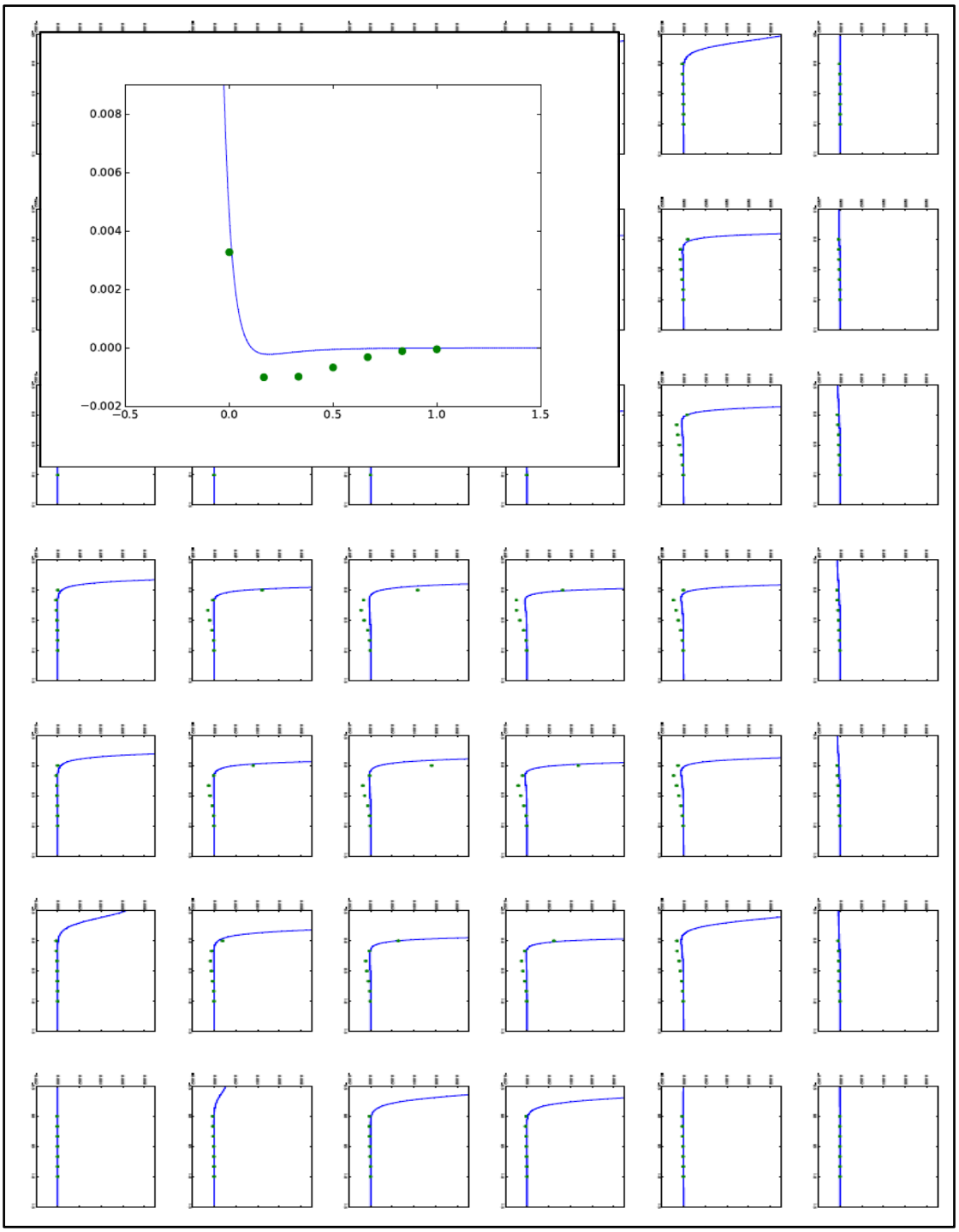


Fig. 4: the fitted CCSD(T)/aug-cc-pVDZ curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree. X-axis is R (Angstrom) and Y-axis is potential energy.

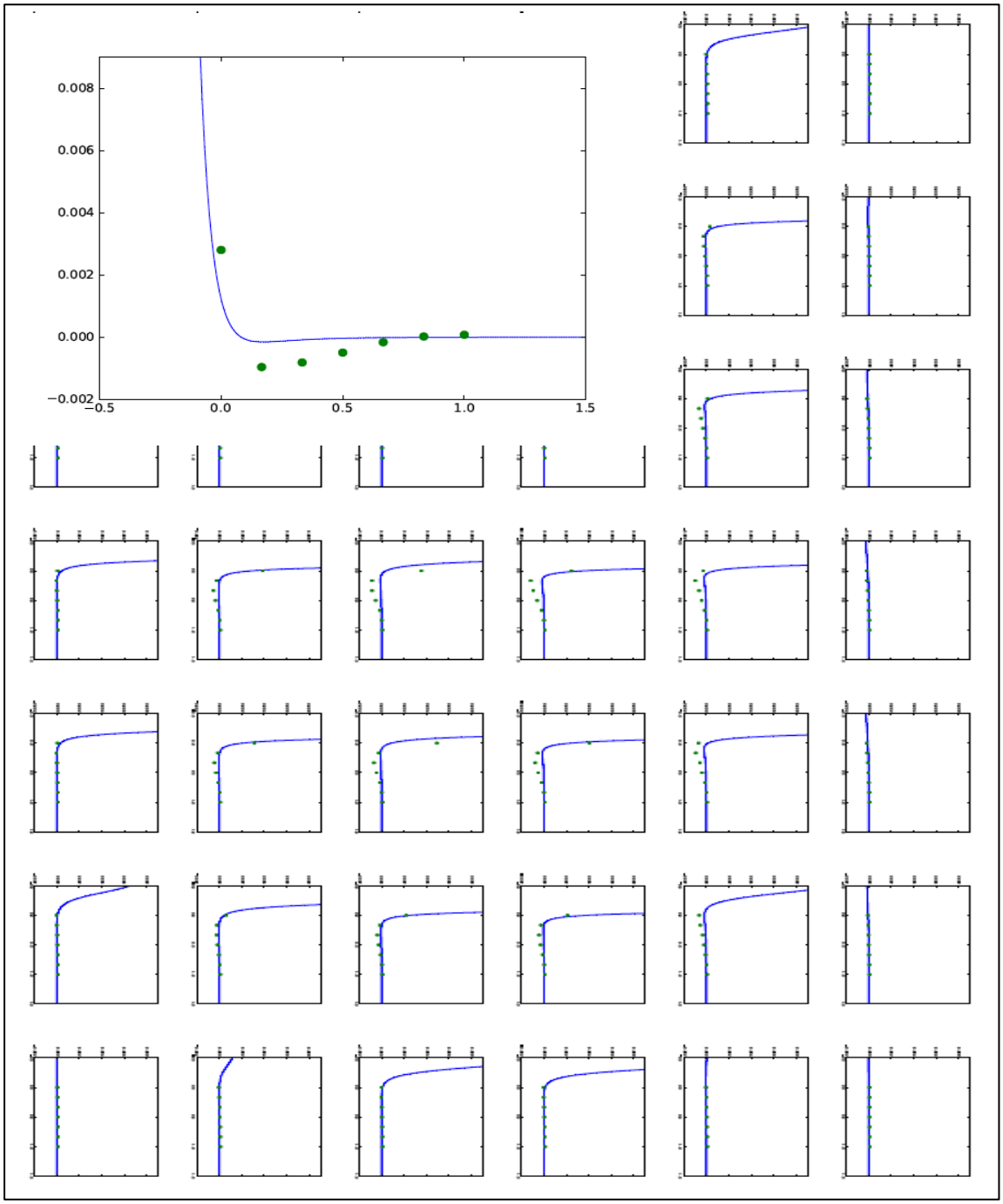


Fig. 5: the fitted CCSD(T)/aug-cc-pVTZ curve, in comparison with the 12-6 LJ curve in 294 positions of H_2 above of the imidazole. Potential energy is given in hartree. X-axis is R (Angstrom) and Y-axis is potential energy.

2. The structures of the systems

2.1 The structure of H_2 …imidazole (parallel)

Geometry= {

n

c 1 rn1c1

n 2 rc1n2 1 ang1

c 3 rn2c2 2 ang2 1 0

c 4 rc2c3 3 ang3 2 0

h 2 rc1h1 1 ang4 5 180

h 3 rn2h2 2 ang5 6 0

h 4 rc2h3 3 ang6 7 0

h 5 rc3h4 4 ang7 8 0

x1 1 rx1 2 ax1 3 0

x2 x1 zh1 1 90 2 90

h1 x2 rhh/2 x1 90 2 0

h2 x2 rhh/2 x1 90 2 180}.

Where x1 and x2 are dummy atoms.

2.2 The structure of H_2 …imidazole (perpendicular)

Geometry= {

n

c 1 rn1c1

n 2 rc1n2 1 ang1

c 3 rn2c2 2 ang2 1 0

c 4 rc2c3 3 ang3 2 0

h 2 rc1h1 1 ang4 5 180

h 3 rn2h2 2 ang5 6 0

h 4 rc2h3 3 ang6 7 0

h 5 rc3h4 4 ang7 8 0

x1 1 rx1 2 ax1 3 0

h1 x1 zh1 1 90 2 90

h2 h1 rhh x1 180 2 90}.

Where x1 is a dummy atom.

2.3 The structure of Co …imidazole

Geometry= {

n

c 1 rn1c1

n 2 rc1n2 1 ang1

c 3 rn2c2 2 ang2 1 0

c 4 rc2c3 3 ang3 2 0

h 2 rc1h1 1 ang4 5 180

h 3 rn2h2 2 ang5 6 0

h 4 rc2h3 3 ang6 7 0

h 5 rc3h4 4 ang7 8 0

x1 1 rx1 2 ax1 3 0

c6 x1 zc6 1 90 2 90

o c6 rco x1 180 2 90}.

Where x1 is a dummy atom.

2.4 The structure of N_2 …imidazole

Geometry= {

n

c 1 rn1c1

n 2 rc1n2 1 ang1

c 3 rn2c2 2 ang2 1 0

c 4 rc2c3 3 ang3 2 0

h 2 rc1h1 1 ang4 5 180

h 3 rn2h2 2 ang5 6 0

h 4 rc2h3 3 ang6 7 0

h 5 rc3h4 4 ang7 8 0

x1 1 rx1 2 ax1 3 0

n3 x1 zn3 1 90 2 90

n4 n3 rn3n4 x1 180 2 90}.

Where x1 is a dummy atom.

2.5 The structure of NH_3 …imidazole

Geometry= {
n1
c2 1 rn1c2
n3 2 rc2n3 1 ang1
c4 3 rn3c4 2 ang2 1 0
c5 4 rc5c4 3 ang3 2 0
h6 2 rc2h6 1 ang4 5 180
h7 3 rn3h7 2 ang5 6 0
h8 4 rc4h8 3 ang6 7 0
h9 5 rc5h9 4 ang7 8 0
x1 1 rx1n1 2 ax1 3 0
n11 x1 zn 1 90 2 90
h12 11 rn11h12 x1 90 2 180
h13 11 rn11h13 x1 90 12 120
h14 11 rn11h14 x1 90 12 -120}.

Where x1 is a dummy atom.

2.6 The structure of H_2O …imidazole

Geometry= {
n
c 1 rn1c1
n 2 rc1n2 1 ang1
c 3 rn2c2 2 ang2 1 0
c 4 rc2c3 3 ang3 2 0
h 2 rc1h1 1 ang4 5 180
h 3 rn2h2 2 ang5 6 0
h 4 rc2h3 3 ang6 7 0
h 5 rc3h4 4 ang7 8 0
x1 1 rx1 2 ax1 3 0
o x1 zo 1 90 2 90
h o roh x1 90 2 180
h o roh x1 90 3 180}.

Where x1 is a dummy atom.

2.7 The structure of H_2 …benzene

Geometry= {
c1
c2 c1 rc1c2
c3 c2 rc2c3 c1 ang1
c4 c3 rc3c4 c2 ang2 c1 0
c5 c4 rc4c5 c3 ang3 c2 0
c6 c5 rc5c6 c4 ang4 c3 0
h7 c1 rc1h7 c2 ang5 c3 180
h8 c2 rc2h8 c3 ang6 c4 180
h9 c3 rc3h9 c4 ang7 c5 180
h10 c4 rc4h10 c5 ang8 c6 180
h11 c5 rc5h11 c6 ang9 c1 180
h12 c6 rc6h12 c1 ang10 c2 180
x1 c1 rx1 c2 ax1 c3 0
h1 x1 zh1 c1 90 c2 90
h2 h1 rhh x1 180 c2 90}.

Where x1 is a dummy atom.