Benchmarking QC optimisation methods for homogeneous TM-based catalysts: A descriptor-based approach for use in high throughput screening

Imme Schoot Uiterkamp

Delft University of Technology

TU Delft

Benchmarking QC optimisation methods for homogeneous TM-based catalysts: A descriptor-based approach for use in high throughput screening

by

Imme Schoot Uiterkamp

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on Tuesday July 11, 2023 at 09:00 AM.

Performed at:
Inorganic Systems Engineering
Chemical Engineering
Faculty of Applied Sciences

Under supervision of:
Prof. Dr. E. A. Pidko
Msc. A. Kalikadien

Student number:	5422558/s2659506	
Project duration:	April 24, 2023 – July 4, 2023	
Thesis committee:	Prof. Dr. E. A. Pidko,	TU Delft, supervisor
	Prof. Dr. F. C. Grozema,	TU Delft

TU Delft

ISEgroup ChemE @ TU Delft

Abstract

Computational chemistry is about making models that simulate the behaviour of real chemical entities. These models can then be used as a predictive tool. For example, in high throughput screening of a library of computationally optimized molecules to find catalysts for drug design. In the screening there need to be values on which it is screened, one type of possible value can be descriptors, a numerical representation of the molecule. It is thus important that these molecules are as accurate as possible to get representative descriptors, but also relatively fast. As to make a library a lot of optimizations are performed, thus driving the costs up if the optimization takes a long time.

To optimize a molecule its energy is needed. Calculating the energy of a system is done using quantum mechanics and is an integral part of computational chemistry. Many methods can be used to approximate the energy of a system. One such way of calculating the energy is DFT. DFT depends on a functional and a basis set, which the user must choose.

This study is the descriptor-based benchmarking of three basis sets, def2-SV(P), def2-TZVPP, def2-QZVPP and five functionals, PBE, TPSS, PBE0, B3LYP and MN15. The optimisations' results are compared using different methods using three molecular descriptors (bite angle, buried volume, and HOMO-LUMO gap). The third part of this study uses a combination of optimisation methods to try and improve the previous results.

The different methods were compared by comparing the optimisation times and the descriptors to the standard of PBE0/def2-SV(P). The structures were optimised using Gaussian, and the descriptors were calculated using the in-house workflow OBeLiX.

When comparing the basis sets, it was noted that def2-QZVPP took too much time to be of use and was thus not used in further comparisons. def2-TZVPP took substantially longer to complete than def2-SV(P). Looking at the descriptors, there was no difference between them. This led to the conclusion that def2-SV(P) was the optimum basis set for these 192 complexes as it was faster but had the same accuracy.

When comparing the functionals, there was the surprising result that the choice of functional did not impact the chosen geometric and steric descriptors of bite angle and buried volume. The electronic descriptor, the HOMO-LUMO gap, differed greatly per method. The lower-level theory PBE and TPSS had a very low value compared to the hybrid functionals but were close to each other. MN15 had a HOMO-LUMO gap that was substantially higher than B3LYP and PBE0. The assumption was thus made that B3LYP and PBE0 were the most accurate functionals in this case. Looking at the time needed for the bulk of the optimisations to complete, PBE was by far the faster functional and MN15 the slowest, PBE0 was located in the middle of the pack. As PBE0 has a shorter optimisation time than B3LYP, the conclusion was that PBE0 was the optimal functional to use in this case.

The third part of the study looked at combining optimisation methods to see if a faster optimisation could be achieved with the same accuracy. Here the base was a fast way, such as GFN2-xTB and PBE, to calculate the geometric and steric properties and then use a PBE0 calculation to make the electronic descriptor as accurate as when doing a general PBE0 optimisation. The best option was doing a single-point PBE0 calculation after a PBE optimisation. It was much faster than a PBE0/def2-SV(P) optimisation and as accurate. Making it an excellent way to optimise the TM complexes.

Contents

List of Figures iii

List of Tables v

1 Introduction 1

2 Theory 4
2.1 DFT . . . 4
2.1.1 Quantum mechanics . . . 4
2.1.2 Exchange-correlation functional . . . 5
2.1.3 Dispersion correction . . . 6
2.1.4 Basis sets . . . 6
2.1.5 Single-point calculations . . . 6
2.2 GFN2-xTB . . . 6
2.3 Descriptors . . . 6
2.3.1 Geometric descriptors . . . 6
2.3.2 Steric descriptors . . . 6
2.3.3 Electronic descriptors . . . 7

3 Methods 8
3.1 optimisation . . . 8
3.1.1 Basis sets . . . 8
3.1.2 Functionals . . . 9
3.2 Descriptor calculation . . . 9
3.3 Comparisons . . . 9

4 Results & Discussion 11
4.1 Comparison of basis sets . . . 11
4.1.1 Comparison of duration of optimisation . . . 11
4.1.2 Comparing descriptors . . . 11
4.2 Different functionals . . . 12
4.2.1 Time needed for optimisation with different functionals . . . 12
4.2.2 Comparing descriptors for the different functionals . . . 13
4.3 Combination of optimisation methods . . . 15
4.3.1 Comparison of time for combinations of optimisation methods . . . 16
4.3.2 Comparison of descriptors for combinations of optimisation methods . . . 17

5 Conclusions and Outlook 19
5.1 Conclusions . . . 19
5.2 Outlook . . . 20
5.2.1 Machine-learning . . . 20
5.2.2 Different programs . . . 20

Bibliography 21

A Extra figures 25
A.1 Benchmarking the basis sets . . . 25
A.2 Benchmarking the different functionals . . . 27
A.2.1 Scatter plots . . . 27

List of Figures

1.1 Potential energy surface of three oxygen atoms, starting as isoozone and moving through a transition state on a saddle point to ozone at the global minimum [3]. 1
1.2 Different methods of optimisation. Placed based on their calculation time and accuracy methods. 2

2.1 Jacob's Ladder showing the different levels of DFT functionals trying to come closer to the ultimate functional. 5
2.2 Visual representation of the bite angle and buried volume. The red atom represents the metal centre, the green atoms are the chelating atoms and the blue the rest groups. a) θ showing the bite angle of the molecule. b) Purple sphere representing the sphere of 3.5 Å for which the percentage of used space will be calculated. 7
2.3 Showing the different energy levels of molecular orbitals with the HOMO, LUMO and their gap. 7

3.1 a) Structure that coordinated towards metal center, creating steric hindrance from a substrate [23]. b) same structure as a only added NBD to the metal centre. c) Structure of NBD that can be added to to go against the coordination of the ligand towards the metal centre. Metle centre shown in blue-green. 8

4.1 Violin plots comparing the 88 ligands optimised with basis set def2-SV(P) vs def2-TZVPP for specific descriptors (a) Results of the bite angle for the optimised ligands. (b) Results of the buried volume with a radius of 3.5Å for optimised ligands. (c) Results of the HOMO-LUMO gap for the optimised ligands. 12
4.2 Histogram of time needed per functional to optimise all the ligands. The bin size is 10 CPU hours. 13
4.3 Violin plots comparing the results of the optimised ligands for the different functionals. (a) Violin plot of the bite angle for each functional. (b) Violin plot of the buried volume with a radius of 3.5 Å for each functional. (c) Violin plot of the HOMO-LUMO gap for each functional. 14
4.4 two figures of the outlier structure from buried volume. a) The structure of the complex with PBE0 optimisation, in red is NBD. b)the structure of the complex with TPSS optimisation, in red is NBD. 14
4.5 Scatter plots showing the correlation and p-value for the HOMO-LUMO gap of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15 15
4.6 histogram of time needed for different functionals or combined calculations. For the combined calculations, the histogram shows the cumulative time. The bin size is 10 CPU hours. 16
4.7 Violin plots comparing the results of the different optimisation methods for the ligands, comparing PBE0 optimisation to a PBE0 single-point calculation after a GFN2-xTB optimisation (a) Violin plots of the bite angle for the optimised ligands. (b) Violin plots of the buried volume with a radius of 3.5Å for optimised ligands. 17
4.8 Violin plots of the HOMO-LUMO gap for each optimisation method. 17
4.9 Scatter plots of the HOMO-LUMO gap of PBE0/def2-SV(P) against different methods. (a) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against sp-PBE0/def2-SV(P) // GFN2-xTB. (b) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against sp-PBE0/def2-SV(P) // PBE/def2-SV(P). (c) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against PBE/def2-SV(P) as comparison 18

A.1 Time taken for the optimizations to converge. For the 88 ligands that first converged with the def2-TZVPP basisset. The size of the bins in the histogram is 50 cpu hours 25
A.2 Scatterplots showing the correlation and p-value for the 88 ligands optimized with basis set def2-SV(P) vs def2-TZVPP for specific descriptors (a) Results of the bite angle for the optimized ligands. (b) Results of the buried volume with a radius of 3.5 Å for optimized ligands. (c) Results of the HOMO-LUMO gap for the optimized ligands. 26

A.3 Density plots for each descriptor for the different methods. (a) Densities of the bite angle for the different methods. (b) Densities of the buried volume with a radius of 3.5 Å for the different methods (c) Densitiees of the HOMO-LUMO gap for the different methods. 27
A.4 Scatter plots showing the correlation and p-value for the bite angle of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15 27
A.5 Scatter plots showing the correlation and p-value for the buried volume with a radius of 3.5 Å of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15 28

List of Tables

3.1 Information on the different functionals used for DFT calculations 9

4.1 Table of CPU hours needed for the fastest optimisation using the method in the first column. The hours are all rounded up to whole hours 11
4.2 Table showing insight into time needed to complete optimisations. By looking first at the percentage of optimisations done after 100 hours of CPU time and then how much time has passed when 90% of the optimisations are done. 12
4.3 Showing the process done and the notation used to describe that process 16

Abbreviations and acronyms

DFT	Density functional theory
DFTB	Density functional tight binding
GGA	Generalized gradient approximation
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
HTS	High throughput screening
ISE	Inorganic Systems Engineering
LDA	Local density approximation
LUMO	Lowest unoccupied molecular orbital
mGGA	meta-GGA
PES	Potential energy surface
RPA	Random phase approximation
TM	Tranistion metal

1

Introduction

Computers have evolved immensely over the last couple of decades, and the rise of computational chemistry came with that development. The term 'computational chemistry' was first used in 1970 and refers to the modelling and simulating of chemical systems [1, 2]. These models are very complex, and a lot of power is needed to run these models. Thus having more easily accessible computational power makes computational chemistry more easily used.

The philosophy behind computational chemistry is that the models simulate the real behaviour of physical entities and that improving the model will reflect the real world more closely [3]. Using these models, computational chemistry can be used to verify or explain experimental results, but it is also becoming a predictive tool [4]. A field where the predictive nature of the models gets used regularly is in drug design [5]. Drug discovery and design is a long process, and pharmaceutical companies want to manage the speed and cost of the process [6]. Using computational methods to screen large amounts of data, High throughput screening (HTS), from a library, for specific target values, can lead to finding new drugs or catalysts needed in drug creation. Doing this computationally saves on performing costly trial and error experimentation in the lab [7]. For reliable results, it is important that the molecules in a library are as accurate as possible. This is done by optimizing the electronic structure of a molecule.

Calculating the electronic structure using quantum mechanics is a big factor in computational chemistry [8]. Finding the electronic structure of a molecule is done by optimizing the molecular energy to a desired point. Each step in the optimisation changes the geometry and, with it, the energy. The relation between geometry and energy can be found on the potential energy surface (PES). This surface shows the potential energy of possible arrangements of a molecule [9]. An example of a PES is shown in figure 1.1, the figure shows a very simple surface, as it is just a three-atom molecule, but it works on the same principle as more complicated surfaces.

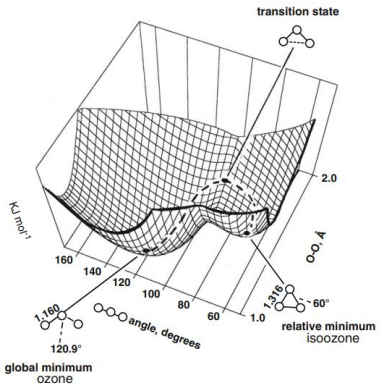


Figure 1.1: Potential energy surface of three oxygen atoms, starting as isoozone and moving through a transition state on a saddle point to ozone at the global minimum [3].

The figure shows the energy path taken from a relatively stable molecule in isoozone to a stabler form in ozone. The PES plots the molecule's energy against factors of its geometry, as shown in the figure. Generally, structure optimisation is done in search of two kinds of points on a PES, the saddle point or a minimum. Saddle points represent a transition state of the molecule, while minima represent a stable molecule which occurs at the lowest energies [10]. Having found the structure's geometry and ground state energy, they are conventionally used to gain insights into the studied chemical system. This is done by calculating properties from the electronic structure [11].

With each step of the optimisation, the energy is calculated to check if the optimisation was done in the right direction. Calculating the energy is however complex as the energy cannot be calculated exactly for a molecule or system larger than one electron. This means that calculating the molecular energy is always an approximation, an important part of developing computational chemistry is thus improving these approximations. There are many methods to approximate the energy of a model. Some are more simplistic than others, meaning they have different calculation accuracies and times. Some are based on molecular mechanics and are very simple, getting a bit more complex are those that are called semi-empirical, like density functional tight binding (DFTB), Density functional theory (DFT) is a level above that and the even more complex are named post-Hartree Fock calculations. These methods are shown in figure 1.2, showing the connection between accuracy and calculation time [12, 13].

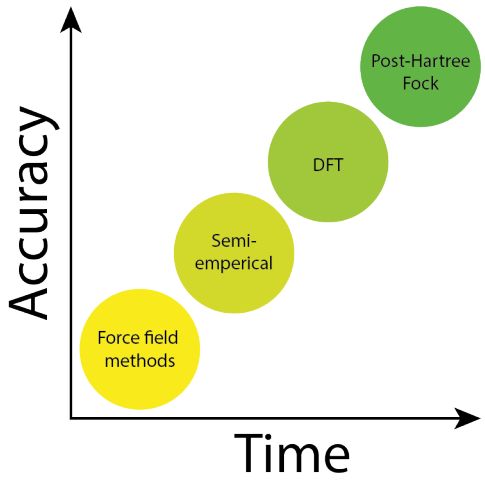


Figure 1.2: Different methods of optimisation. Placed based on their calculation time and accuracy methods.

DFT is widely used in computational chemistry and is based on the Schrödinger equation [8]. As DFT is only an approximation to solve for the energy, certain choices that influence the result must be made. Mainly a choice for a functional and a basis set. Functionals are a function of the electron density solving a specific part of the system's energy, officially named exchange-correlation functionals. Basis sets are a combination of functions to approximate the wave function. Choosing the functional and basis set determines the accuracy of the result. The optimal functional needs to be determined for the specific structures that will be used as some functionals may be more accurate for organic molecules, while some may give more accurate approximations for transition metal (TM) complexes. It is, therefore important to be able to make an informed decision about which ones to use. This decision can be made on accuracy but also on cost.

The time needed for calculations is very important as that determines the cost of the calculations. Most calculations are done on a supercomputer as they have the capacity to handle quantum chemical systems (~ 100 atoms). Running the supercomputer, however, comes at a cost. Supercomputers need a lot of power, so much in fact that the power some supercomputers use could sustain a city of 40000 people [14]. This amount of power means that supercomputing has a large carbon footprint and, thus, a large environmental impact. When picking the methods for the DFT calculation, it is thus important to balance accuracy and time. In the case of preparing a library for HTS this balance is especially important as there is optimisation done for a large number of molecules.

A way to find the best options for a DFT calculation is by benchmarking the different methods. Benchmarking is the process of comparing different calculation methods [15]. There are a lot of benchmarking studies already done for DFT methods [16]. However, as mentioned earlier, some methods might work well for one type of molecule but not for the other [17, 18]. Benchmarking is done by running the DFT calcula-

tion for a specific subset of molecules for each functional or basis set and comparing the results on a specific factor.

This study is a descriptor-based benchmarking study of DFT optimisation on 192 Rhodium catalysts. Descriptors are numerical values that describe the chemical characteristics of a molecule [19]. The 192 complexes were chosen based on a high-throughput experimentation campaign to explore commercially available ligands' performance. Different functionals and basis sets will be used for optimisation calculations of these complexes, and compared on their descriptors.

This research aims to find an optimisation method that takes a minimal amount of time but still yields good results. This is done to use the optimal method in a workflow to create a virtual library of these ligands, that is being created by the Inorganic Systems Engineering (ISE) group of the TU Delft. The descriptors will be used to compare the molecules in HTS. A possible Machine-learning model for the screening of the catalysts would also use the descriptors as training data.

This study is a data-driven approach to benchmarking, which means that the results will only be compared to other data and not to experimental values, as is quite common in benchmarking [20]. This has the benefit that strong correlations between different methods, are also a good result even if the values differ. With a strong correlation, the accurate results could be extrapolated from the less intensive method, reducing the optimisation time. Using this data-driven approach based on descriptors differs from usual benchmark studies. The usual benchmark studies do not use descriptors to compare but use the system's energies [17]. This was not done in this study as the further uses of the optimised molecules need the descriptors and not the energies.

The benchmarking can be split into three parts, the first is benchmarking the different basis sets, the second part is comparing different functionals and the third part of the study is looking at the combination of different optimisations.

The functional/basis set combination used for homogeneous catalyst structures similar to the Rh catalyst in the ISE group is PBE0/def2-SV(P). This is, therefore, also the basis for the chosen functionals and basis sets to compare in this benchmark study.

For the benchmarking of the basis sets, it was chosen to compare def2-SV(P) to 2 other basis sets from the Alrichs group. These basis sets are appealing to use as they have accuracy for all atoms until Radon [21]. This includes all atoms in our ligands, and possible changes for the metal core could be made without losing accuracy. The other two are def2-TZVPP and def2-QZVPP. From these three, def2-SV(P) is the least extensive, and def2-QZVPP is the most extensive. It is expected that the time for the more extensive ones will be longer as they have more functions that describe the system leading to more computer power needed.

Functionals have different levels of theory. How higher the level of theory, how more accurate the functional should be. This study will compare five different functionals from three levels of theory, GGA, mGGA and hybrid. The GGA functional PBE was chosen as that one has the most overlap with PBE0. For mGGA, the TPSS functional was chosen as that is one of the more widely used ones. The same goes for B3LYP in the hybrid category, B3LYP is one of the most used functionals in DFT and, thus, interesting to compare. PBE0 is also a hybrid functional and is included as the standard. Lastly, another hybrid functional called MN15 was chosen as it is very different from the other two as it is specifically made with transition metals in mind [22]. Comparing these functionals, the expectation is that for each category, the steric and geometric descriptors will be close to the ones in the same category but that the electronic descriptors will be highly affected by the functional even in the same category.

The third part of the study is looking at possible combinations of optimisations methods, a good example of this is using the semi-empirical GFN2-xTB optimisation, which according to previous research done in the ISE group should already provide a good basic geometry [23]. Then using a more extensive DFT calculation on the GFN2-xTB optimised structures, to try and get the electronic elements of the molecule more accurate.

The main body of this report will start with a theoretical background on DFT and molecular descriptors. After the theory, the methodology of this benchmarking study will be described. Then the results of this research will be presented and discussed, and lastly, a conclusion will be made and an outlook given.

2

Theory

This chapter describes the theory behind the methods used in this study. Section 2.1 is focused on the theory behind DFT, Section 2.2 focuses shortly on GFN2-xTB and lastly, Section 2.3 goes into the theory behind the molecular descriptors used in this study.

2.1. DFT

Section 2.1 will explain the theory behind DFT in multiple steps, first the derivation will be explained, after that separate elements of DFT will be looked at, the functional, the dispersion correction and the basis set. Lastly, there will be looked at the theory of a single-point calculation.

2.1.1. Quantum mechanics

An important fact of QM (quantum mechanics) is that all chemical systems have a wavefunction and that specific operators, which are functions, that act on the wavefunction return the properties of a system [9]. One of these operators is the Hamiltonian, which acting on a wavefunction returns the system energy which is shown in equation 2.1, the time-independent Schrödinger equation [24].

$$\hat{H}\psi = E\psi \tag{2.1}$$

The $\hat{H}$, is the symbol for the Hamiltonian operator, ψ is the wavefunction and E is the system's energy. The Hamiltonian is a way to describe the kinetic and potential energies of all particles in a system. It takes into account the contribution of all electrons and nuclei in the system. A way to simplify it is using the Born-Oppenheimer approximation, which states that the nucleus is stationary from the point of view of an electron due to the mass difference [25]. The Hamiltonian shown in equation 2.2 is the Hamiltonian simplified with the Born-Oppenheimer approximation.

$$\hat{H} = -\sum_{i}^{elec.} \frac{\nabla_i^2}{2} - \sum_{i}^{nucl.} \sum_{j}^{elec.} \frac{Z_i}{r_{ij}} + \sum_{i<j}^{elec.} \sum \frac{1}{r_{ij}} \tag{2.2}$$

The first term shows the kinetic energy of the electrons, and the second and third terms show the potential energy of the attraction between nuclei and electrons and the repulsion between electrons respectively [26].

DFT solves the Hamiltonian through the use of electron density. The basis of this is described in the Hohenberg-Kohn (HK) theorems [27]. There are two HK theorems, the first states that the ground state energy of a system is a unique functional of the electron density. A functional is a function that takes another function as an input. The second states that the ground-state electronic density minimizes the energy of the system.

A method to obtain the electron density was found by Kohn and Sham [28]. The electron density is obtained by first considering a set of non-interacting electrons and then adding on a factor that adds the interaction between the electrons. Equation 2.3 shows the energy functional from the Kohn-Sham calculations.

$$E[\rho(r)] = T_{ni}[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + \Delta T[\rho(r)] + \Delta V_{ee}[\rho(r)] \tag{2.3}$$

Where $\rho(r)$ represents the electron density, T_{ni} the kinetic energy of the non-interacting electrons, V_{ne} the nuclear-electron interaction, V_{ee} the classical electron-electron repulsion, ΔT the kinetic energy correction

from interacting electrons and ΔV_{ee} all non-classical correction to the electron-electron repulsion energy [9]. The last two terms add back the interaction between the electrons and are combined to $E_{XC}[\rho(r)]$, the exchange-correlation functional.

2.1.2. Exchange-correlation functional

The $E_{XC}[\rho(r)]$ transforms the density to the exchange correlation energy, E_{XC}. The functional is however unknown and needs to be approximated to solve for the energy. The main problem of DFT is making a good functional. There have been many different functionals made, with different levels of theory. The different levels of theory are shown on a DFT Jacob's ladder, see figure 2.1, which shows the hierarchy of the level of theory trying to reach the ultimate accurate functional, the divine functional [29, 30].

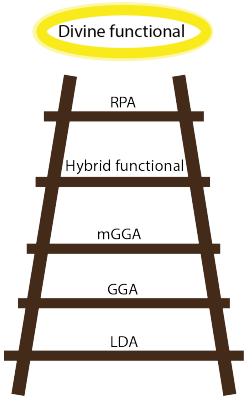


Figure 2.1: Jacob's Ladder showing the different levels of DFT functionals trying to come closer to the ultimate functional.

From low to high the ladder has the local density approximation (LDA), the generalized gradient approximation (GGA), meta-GGA (mGGA), hybrid (meta) GGA also called hybrids, and random phase approximation (RPA). Being lower on the ladder means that it is a simpler approximation. This study only focuses on the middle three levels, GGA, mGGA, and hybrids.

GGA

GGA functionals are a step up from LDA functionals as shown on the Jacob's ladder. LDA assumes a uniform density, GGA takes the gradient of the density also into account. It looks at the first derivative or the electron density to see the gradients. This results in a more complete model of the density. PBE is one of the most used GGA functionals as it is applicable over a wide range of systems [**?**]. It was introduced in 1996 by Perdew, Burke, and Ernzerhof, after who it is named [31].

mGGA

A mGGA functional is an extension of a GGA functional where it not only looks at the gradient but also its direction by incorporating the second derivative of the density. TPSS, published in 2003 is one of those functionals and is one of the standards for mGGA [32].

Hybrid functionals

The exchange-correlation energy can be separated into two parts. The exchange energy and the correlation energy. Hybrid functionals were proposed in 2003 by Becke [**?**]. They incorporate a percentage of exact exchange from the Hartree-Fock (HF) theory into the exchange energy of the GGA functional. HF exact exchange is something that is known exactly, it can however not be directly swapped out for the exchange energy of a different functional, as there are some errors in the correlation energy of the GGA functionals that are cancelled out by the GGA exchange energy. The solution to this problem is implementing a percentage of exact exchange with the exchange energy of the GGA functional. Hybrid functionals can also have different functionals mixed together in certain percentages each having a parameter for added energies. In this study, we will look at three different hybrid functionals. PBE0 was published in 1999 and is the PBE functional with 25% HF exchange [33]. It thus only has one parameter.

B3LYP is probably the most used functional, it has three parameters. A HF exchange of 20%, and adds another exchange energy and has two correlation energies, making up the other two parameters [34].

The last functional is MN15, from the Minnesota functionals. It is a highly parameterized function with a HF exchange of 44% [22]. And made specifically with TM-complexes in mind.

2.1.3. Dispersion correction

In larger molecules London dispersion forces are more apparent, accounting for these is thus very important to reach a better DFT accuracy. Functionals do not, however, always do this. In the case of an Argon dimer, PBE0 captures 99.95% of the total energy, but no more than 15% of the interaction energy [35]. There are some functionals that do have an intrinsic dispersion correction such as MN15. For the other functionals, a dispersion correction is applied. Grimme introduced the dispersion correction D3 which is applicable to a large number of systems which include TM complexes [36]. Applying Becke Johnson damping results in an even more accurate dispersion correction known as D3BJ [37].

2.1.4. Basis sets

Basis sets are mathematical approximations of the wavefunction, ψ, from the Schrödinger equation. They are composed of a set of functions. As they are an approximation there are multiple ways to compose the basis sets. Two main types are STO (Slater-Type orbitals) and GTO (Gaussian-Type orbitals). STO orbitals are more extensive but also more costly. Valence electrons are the main source of bonding, therefor inner-shell and valence atomic orbitals are commonly split. The valence atomic orbitals can be described by two or more functions. This is called split-valence. Examples of this are def2-SV(P), which is a split valence so has two functions per valence atomic orbital, def2-TZVPP which is triple zeta valence and has three functions per valence atomic orbital, and def2-QZVPP which is quadruple zeta valence and thus has four functions per atomic orbital. The P's stand for the amount of polarization in the basis set.

As the number of functions increases the time needed for optimisation will increase drastically as it does not just add one function, but one function for every electron.

2.1.5. Single-point calculations

A single-point calculation is a way to find information for a given geometry. Where an optimisation explores the potential energy surface, a single-point calculation only stays on one point. The DFT calculation will be done for the given geometry.

2.2. GFN2-xTB

There are also less resource-intensive ways to calculate the energy. A level below DFT are semi-empirical methods. GFN2-xTB is one of those methods, GFN stands for Geometries, Frequencies and Non-covalent interactions, and xTB means tight-binding method. It can compute for a lot of elements including transition metals, has a generally good balance between accuracy and efficiency, and can handle very large systems which makes it a very useful way of calculating.

2.3. Descriptors

Descriptors are a way to describe molecular properties. Many different descriptors exist which all describe the system in a slightly different way. They come in different categories, mainly geometric, steric, and electronic. Each category will be shortly explained along with a specific descriptor from that category.

2.3.1. Geometric descriptors

Geometric descriptors are defined by the angles of a structure. They purely depend on the coordinates of the atoms.

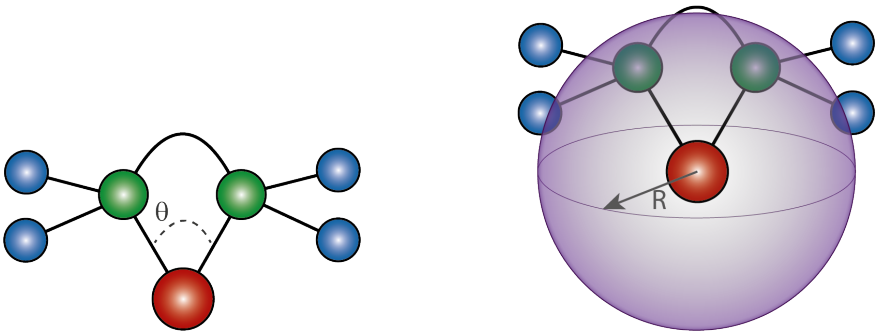
The bite angle is a geometric descriptor which describes the angle between the metal atom and the two chelating ligand atoms. The bite angle is denoted in figure 2.2a with θ. As it needs the angle between two chelating ligands it only works for a bidentate ligand.

2.3.2. Steric descriptors

Steric descriptors describe the steric conditions and thus describe possible steric hindrance.

The buried volume describes the percentage of space filled around the metal centre. A visual is shown in figure 2.2b, as seen in the figure an imaginary sphere comes around the metal centre and then it looks at the percentage of the sphere that is filled by the ligand. The radius of the sphere around the metal centre is

adjustable but generally, it is around 3.5 Å. Because it shows the amount of space around the metal centre it is a good metric showing possible steric hindrance for a possible attachment.



(a) (b)

Figure 2.2: Visual representation of the bite angle and buried volume. The red atom represents the metal centre, the green atoms are the chelating atoms and the blue the rest groups. a) θ showing the bite angle of the molecule. b) Purple sphere representing the sphere of 3.5 Å for which the percentage of used space will be calculated.

2.3.3. Electronic descriptors

Electronic descriptors describe the electronic state of a molecule. The HOMO-LUMO gap is an important electronic descriptor as it influences other descriptors as well such as the electronegativity.

The HOMO-LUMO gap describes the filling of the molecular orbitals with electrons. HOMO stands for highest occupied molecular orbital and LUMO for lowest unoccupied molecular orbital. The HOMO-LUMO gap is visually explained in figure 2.3. Showing different molecular orbitals, filled and unfilled, and the gap in between.

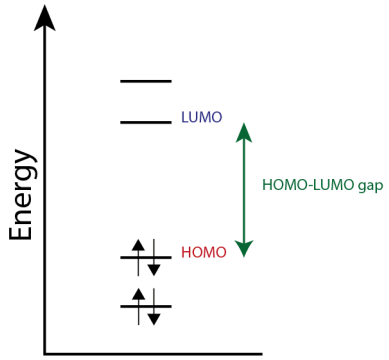


Figure 2.3: Showing the different energy levels of molecular orbitals with the HOMO, LUMO and their gap.

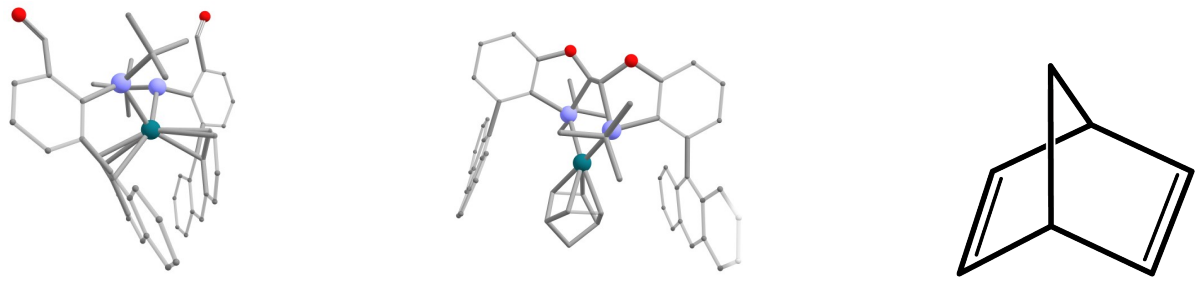
3

Methods

This chapter describes the method and workflow used for this study. Section 3.1 describes the methods used for optimising the metal-ligand complexes. Section 3.2 explains the methods behind the calculations of the molecular descriptors. In section 3.3, the methods of comparison of the different ways of calculation are presented.

3.1. optimisation

The 192 catalysts used in this study all had a rhodium metal centre with charge 1. Figure 3.1 showcases two optimised structures which were the same before optimisation except the left did not have NBD (Norbadiene), shown on the right, attached to the metal atom. In this optimisation the ligand atoms coordinated towards the metal atom, which creates steric hindrance for a possible substrate. The middle molecule does have NBD attached and as can be seen during the optimisation the atoms did not coordinate to the metal centre, leaving space for a substrate instead of the NBD.



(a) (b) (c)

Figure 3.1: a) Structure that coordinated towards metal center, creating steric hindrance from a substrate [23]. b) same structure as a only added NBD to the metal centre. c) Structure of NBD that can be added to to go against the coordination of the ligand towards the metal centre. Metle centre shown in blue-green.

Because of this all structures in this study had NBD added to the complex. The starting structures for the different ligands were created manually or partially automated.

Structure optimisation was done using DFT in the Gaussian 16 C.01 suite [38]. Gaussian was run on the Snellius supercomputer which is the Dutch national supercomputer [39]. All calculations included natural bond order (NBO), were performed in the gas phase using varying basis sets and functionals. These calculations were conducted on an ultrafine grid.

3.1.1. Basis sets

To be able to compare the effects of the different basis sets on the optimisation, the functional is kept constant in this part of the research. The functional used was PBE0 with the dispersion correction GD3BJ from the grimes group [37]. The three basis sets that are compared in this study are listed below.

- def2-SV(P) (split valence, polarization on heavy atoms)

- def2-TZVPP (Triple zeta valence, two sets of polarization functions)
- def2-QZVPP (Quadruple zeta valence, two sets of polarization function)

Def2-SV(P) is the standard basis set used in this part of the research of the ISE group and is thus included as the standard. The other two are expansions on it, every time a function is added per atomic orbital. All 192 ligands were started for each basis set. The limit for jobs on Snellius is five days of wall time. If more than five days were needed then the optimisation was restarted reading the geometry and wavefunction from the checkpoint file.

3.1.2. Functionals

Shown in table 3.1 are the five functionals compared in this study with added information.

Functional	Dispersion	Level	Parameterization
PBE	D3BJ	GGA	Low
TPSS	D3BJ	mGGA	Low
B3LYP	D3BJ	hybrid	medium
MN15	intrinsic	hybrid	very high
PBE0	D3BJ	hybrid	low

Table 3.1: Information on the different functionals used for DFT calculations

This table shows all functionals, the used dispersion correction, the level of theory and the level of parametrization. All functionals use the GD3BJ dispersion correction except for MN15, which has an intrinsic dispersion correction. As can be seen in table 3.1 above all functionals use the D3BJ dispersion correction except for MN15, which has an intrinsic correction. The different functionals fall into three levels of theory. The hybrid category contains three different functionals, each with a different level of parameterization.

The basis set with which these will run depends on the results of the benchmark for the basis sets done previously. Then as with the basis sets, all 192 ligands will be optimised. In case of errors, for example, with certain angles, the optimisation will fail. If this happens the calculation was restarted, but the guess for the wavefunction is read from the checkpoint file, a machine readable file containing the progress of the optimisation, to hopefully nudge the optimisation in the right direction.

3.2. Descriptor calculation

The molecular descriptors were calculated using the descriptor calculator contained in *descriptor_calculator.py* from OBeLiX (Open Bidentate Ligand eXplorer) [40]. OBeLiX is a Python-based workflow made by the ISE group. The descriptor calculator functionality uses the log files obtained from the DFT optimisations and extracts the optimised *.xyz* files. From these structures, 75 descriptors are calculated and exported in a *.csv* file. The descriptors are calculated using the MORFEUS (MOleculaR FEatureS for machine learning) and cclib python packages [41, 42].

3.3. Comparisons

The goal of this research is to compare the results of the optimisation by comparing specific descriptors and optimisation time. The descriptors that will be compared are one from each of the three categories steric, geometric and electronic. The three descriptors used for comparison are listed below.

- Bite angle (Geometric)
- Buried volume with a radius of 3.5 Å (Steric)
- HOMO-LUMO gap (electronic)

The choice to use bite angle and buried volume as the geometric and steric descriptors were made based on their easy interpretability and their extensive previous use in other research [43–45]. For a lot of electronic descriptors, the HOMO-LUMO gap is a parent metric thus it is the most useful electronic descriptor to use as a way to compare the optimisation methods.

Alongside descriptors, optimisation time is also an important comparing factor. Each structure optimisation yields two time measurements: wall time and CPU time. Wall time represents the elapsed time from the start to the completion of the optimisation process. However, this metric alone is not ideal for comparison as it does not accurately reflect the computational workload. On the other hand, CPU time provides a more meaningful metric for comparing optimisation efficiency. It measures the cumulative time spent by all CPUs involved in performing the optimisation. By focusing on CPU time, the actual computational effort invested in the optimisation can be assessed. Thus when comparing optimisation times, this research looks at CPU time in hours, further referenced as CPU hours.

The comparisons were mostly done by generating different plots to be able to compare the different methods visually. Histograms were generated to be able to compare the time needed for optimisations of the structures. Violin plots were used as a way to compare all methods at one time. It makes it easy to spot differences in values and differences in densities. Scatter plots are used to see if there is a correlation between the values even if the actual values may not be the same. The scatterplots are all of the different methods plotted against the standard method of PBE0/def2-SV(P).

Lastly, an ANOVA (Analysis Of Variance) statistical test was performed. ANOVA is a way to compare multiple data groups working from the null hypothesis that the two groups are the same. From the test a p-value will be found, if the p-value is lower than 0.05 then the null hypothesis is rejected. The ANOVA test was performed using the Python package scipy.stats [46]. The ANOVA test is used to compare the values of the descriptors for a specific method against the descriptor values from the standard method of PBE0/def2-SV(P).

If there are some surprising results, the structures of the optimised metal-ligand complexes are compared by overlaying the structures using chemcraft 1.8 [47]. To try and explain possible differences.

4

Results & Discussion

4.1. Comparison of basis sets

4.1.1. Comparison of duration of optimisation

The first step in comparing the different basis sets is comparing the duration of the optimisation. This is done by comparing the CPU hours needed to complete the calculations. Table 4.1 shows the time required for the first structure optimisation to be completed in CPU hours.

Functional/Basis set	CPU hours fastest optimisation
PBE0-D3BJ/def2-SV(P)	5
PBE0-D3BJ/def2-TZVPP	175
PBE0-D3BJ/def2-QZVPP	9640

Table 4.1: Table of CPU hours needed for the fastest optimisation using the method in the first column. The hours are all rounded up to whole hours

The different times correspond with the ranking of the different levels of theories, with def2-SV(P) having the lowest time and def2-QZVPP the highest. The increase in time is expected to have significant gaps as with each extra level; a new function is added for each atomic orbital, leading to a drastic time increase. The time needed for optimisations using the basis set def2-QZVPP takes so long that after around 10240 CPU hours, only 2 structures converged. This led to the decision to stop running the optimisations with this basis set, as the time it takes already makes it so costly that it would not be a feasible option to use further.

For the def2-TZVPP basis set, the structures were optimised for a maximum of 4600 CPU hours; in that time, 88 structures converged. Due to limited time, it was decided that these 88 were enough to compare the two remaining basis sets. For def2-SV(P), all ligands converged in a reasonable time, but in comparing the two basis sets, only the same 88 will be used that there is data on from the def2-TZVPP optimisations. These 88 for def2-SV(P) were all optimised in 100 CPU hours. Figure A.1 in the appendix shows a histogram of the time needed for these optimisations to converge.

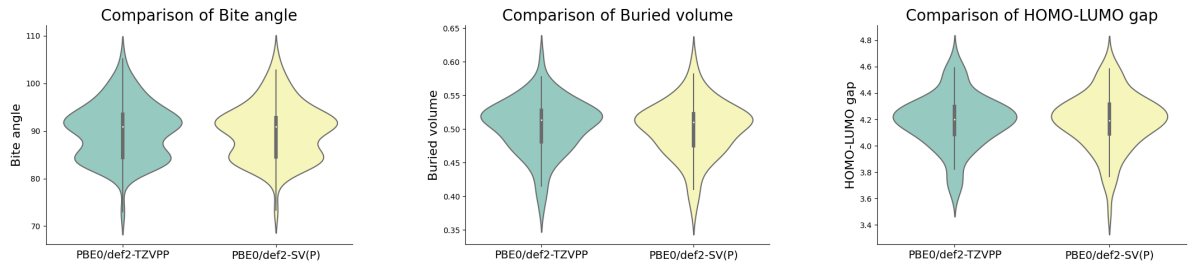
From these numbers, it can be concluded that when only looking at the time needed for the optimisation, def2-SV(P) seems more optimal.

4.1.2. Comparing descriptors

After a time comparison, the actual results of the optimisation need to be compared using descriptors to make a conclusion. Violin plots shown in figure 4.1 are used to get a general idea of how similar or different the results are.

As shown in figure 4.1, the optimisation results with the basis set def2-SV(P) vs def2-TZVPP do not give a big difference; their distribution looks almost identical. To further see if they are the same, the correlations are checked, and the ANOVA statistical test is done, the results are shown in figure A.2 shown in Appendix A. They all have a very strong correlation and similarity.

For the bite angle, the r^2 is 0.99 which means that almost all points fall in a straight line. There are a couple of outliers, but there was no specific reason found. , and fall almost on one line, except for a couple of outlier



(a) (b) (c)

Figure 4.1: Violin plots comparing the 88 ligands optimised with basis set def2-SV(P) vs def2-TZVPP for specific descriptors (a) Results of the bite angle for the optimised ligands. (b) Results of the buried volume with a radius of 3.5Å for optimised ligands. (c) Results of the HOMO-LUMO gap for the optimised ligands.

points, but those do not have an apparent reason or similarity. the p-value from the ANOVA statistical test had a value of 0.97, this means that besides the strong correlations, the actual values are almost identical. For the buried volume, there was again a correlation of 0.98, but in this case, the p-value was lower at 0.52 meaning that there is less similarity between the datasets, but they are not statistically different. Lastly, the HOMO-LUMO gap had high values for r^2 and the p-value, 0.96 and 0.97 respectively. There is thus, almost no difference between using def2-SV(P) and def2-TZVPP when comparing the descriptors.

With time already on the side of the basis set def2-SV(P) and almost no difference between the results in the descriptors of the different basis sets, it can be concluded that def2-SV(P) is superior compared to def2-TZVPP.

4.2. Different functionals

Having chosen basis set def2-SV(P) as the best option, the optimisations for the 192 ligands were run with that basis set and then the five different functionals which are outlined in table 4.2. Again, the time will be analyzed afterwards, the descriptor results will be compared.

4.2.1. Time needed for optimisation with different functionals

The time it takes to optimise the 192 ligands with the different functionals is shown in figure 4.2

The histogram clearly shows a big difference between the times for the different functionals. The MN15 functional takes the longest time with its bulk and outliers. This is not surprising as it is a hybrid functional, meaning that it is already expected to be slower than PBE and TPSS, and it is also the most parameterized functional, which leads to a longer optimisation time. PBE is unsurprisingly the faster functional as it has the lowest level of theory.

Functional	Percentage done after 100 hours CPU time	CPU time where 90% is done
PBE0	89	105.5
B3LYP	87	113.1
MN15	72	152.7
TPSS	93	88.6
PBE	95	58.9

Table 4.2: Table showing insight into time needed to complete optimisations. By looking first at the percentage of optimisations done after 100 hours of CPU time and then how much time has passed when 90% of the optimisations are done.

Table 4.2 summarizes the results of figure 4.2. On the left side, the percentage of structures done after 100 hours are summed per functional, and the right side shows after which time 90% of the 192 complexes are done with optimisation. This shows more insight into the time distribution of each functional. The left side should be as high as that means a larger percentage got done in the same amount of time, the right side should be low meaning that the 90% got finished faster. From the table, now the information gotten from the histogram can be expanded. The table confirms that MN15 is the slowest functional. Between B3LYP and PBE0, PBE0 is a bit faster in both columns of the table. Between PBE0 and TPSS is a gap, with PBE0 being slower. TPSS and PBE have similar percentages for the amount done after 100 CPU hours but are very different for when 90% was completed. This can be explained by looking at the histogram, at 100 hours they

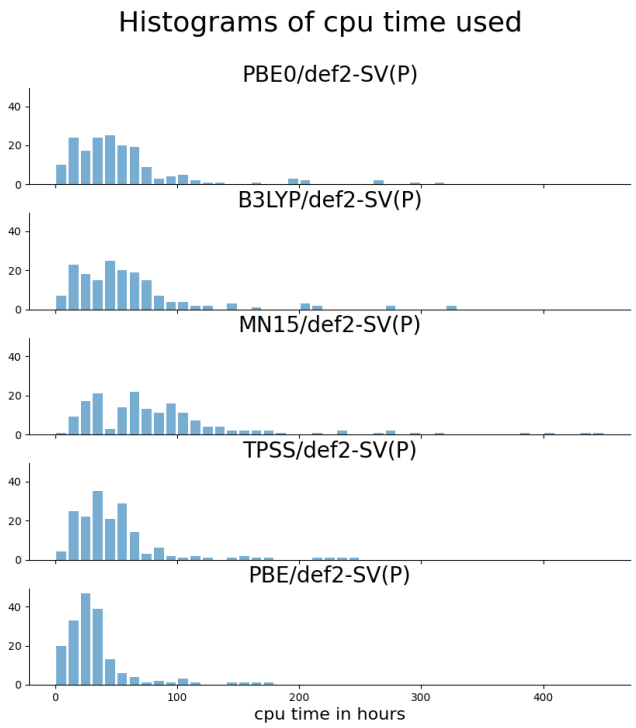


Figure 4.2: Histogram of time needed per functional to optimise all the ligands. The bin size is 10 CPU hours.

both have only outliers left, the percentage done is thus similar but when the bulk of the optimisations got done is very different. In the end, ranking the functionals based on the optimisation times gives, from fastest to slowest:

1. PBE
2. TPSS
3. PBE0
4. B3LYP
5. MN15

This follows the level of theory of each functional, with PBE being the lowest and PBE0, B3LYP and MN15 being the highest level of theory. The hybrid functionals just mentioned then follow the order of parameterization, with PBE0 having the lowest parameterization and MN15 the highest.

4.2.2. Comparing descriptors for the different functionals

To choose an optimal functional, the descriptors need to be compared. This, again, is firstly done with violin plots, as shown in figure 4.3

The violin plots of the bite angle, figure 4.3a, show almost no difference between the different functionals. Their distribution is identical, shown in figure A.3a in Appendix A. Looking further into their similarities by checking the r^2 and p-value found in figure A.4 in the appendix we find that all the r^2 are between 0.965 and 0.995, which shows a very strong correlation, the p-values were between 0.718 and 0.884 showing that indeed the values were very similar.

For the buried volume, the violin plots, figure 4.3b, looked also almost identical. There was a noticeable difference, every violin plot except from TPSS shows a spike to 0.8. The outlier was the same for each functional so its structure was looked at, figure 4.4.

The figure shows the disassociation of NBD for PBE0 this happened for all functionals except for TPSS as shown on the right. The disassociation of the metal atom resulted in a lack of steric hindrance allowing the

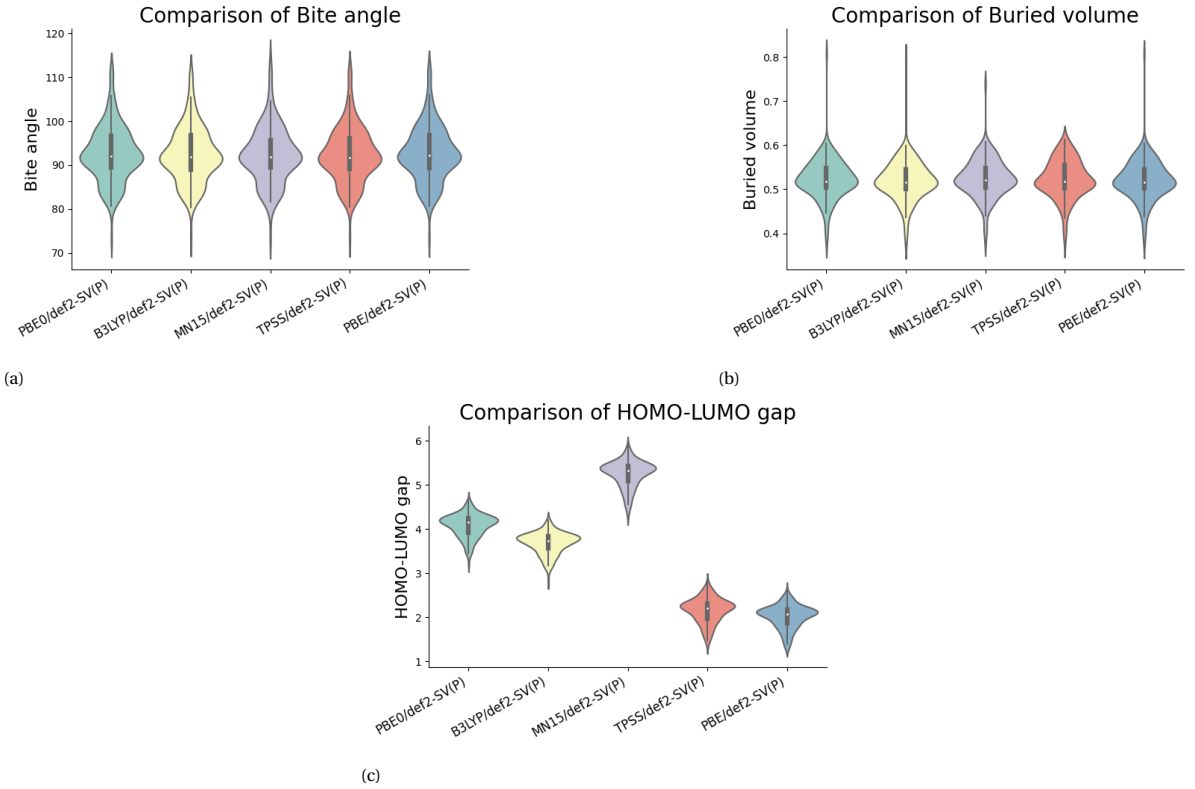


Figure 4.3: Violin plots comparing the results of the optimised ligands for the different functionals. (a) Violin plot of the bite angle for each functional. (b) Violin plot of the buried volume with a radius of 3.5 Å for each functional. (c) Violin plot of the HOMO-LUMO gap for each functional.

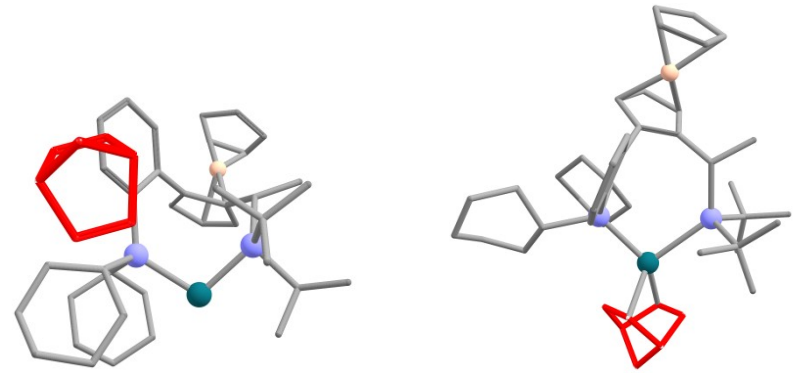


Figure 4.4: two figures of the outlier structure from buried volume. a) The structure of the complex with PBE0 optimisation, in red is NBD. b)the structure of the complex with TPSS optimisation, in red is NBD.

ligand to coordinate with the rhodium centre. This results in a larger buried volume. Apart from that spike, the violin plots look identical which the density plot, figure A.3b appendix A, also shows. The correlations, figure A.5, had an r^2 between 0.825 and 0.982. The r^2 of 0.825 is the correlation between TPSS and PBE0 which is understandably lower as there is a very big outlier as mentioned earlier. The next lowest r^2 had a value of 0.946. The p-values from the ANOVA statistical test were between 0.485 and 0.840, the null hypothesis of statistically same datasets still stands for all functionals.

Figure 4.3c shows that the chosen functional substantially influences the HOMO-LUMO gap. This is expected as the functional affects the electronic interactions specifically. It is also known in the literature that PBE0 has a consistently higher HOMO-LUMO gap than B3LYP [48]. These results are thus not unexpected.

It is interesting to look at possible correlations even though the values are not the same. As having a

strong correlation which can be used to find a more correct HOMO-LUMO gap quickly by extrapolation. The scatterplots for the HOMO-LUMO gap for each functional are shown in figure 4.5

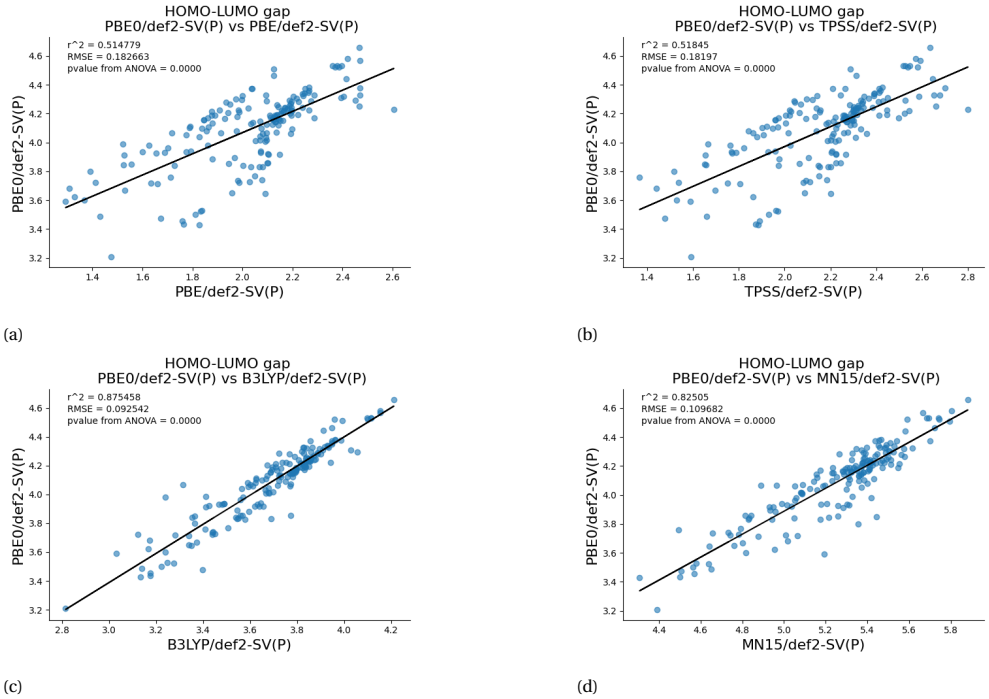


Figure 4.5: Scatter plots showing the correlation and p-value for the HOMO-LUMO gap of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15

The figure shows that there is a strong correlation between PBE0 and the other hybrid functionals, r^2 of 0.875 for B3LYP and 0.825 for MN15. Their p-values are zero for all the functionals against PBE0, this means that all the datasets are statistically different from PBE0 which was already seen in the violin plot. Looking at the correlation between PBE0 and the functionals with a lower level of theory shows a weak correlation of 0.51 for both. The points are even more widespread making extrapolation again not possible.

It is surprising to note that the bite angle and buried volume are the same for all functional, as it implies that the functional does not affect the chosen geometric and steric descriptors but does affect the electronic ones. This is important as it means that the structural differences are thus not as big as the electronic differences between functionals. The similarity cannot be attributed to the dispersion correction as MN15 has an intrinsic correction and still has the same values.

The functionals can thus only be compared by the HOMO-LUMO gap and time. There were big differences between the hybrid functionals and the mGGA and GGA functionals for the HOMO-LUMO gap. Between the hybrid functionals there was a s big difference between MN15 and the other two hybrids, PBE0 and B3LYP were pretty close. As PBE and TPSS are a lower level of theory we can say that their HOMO-LUMO gap results are less accurate than those of the hybrids. The functional MN15 might be inaccurate for our structures as it has a big gap with B3LYP and PBE0. Comparing B3LYP and PBE0 on the HOMO-LUMO gap is not really possible, taking the time into account PBE0 comes out favourably over B3LYP.

From these results, it seems that PBE0/def2-SV(P) is the best option to get accurate and fast results.

4.3. Combination of optimisation methods

Looking further into the results, because of the result that the functional does not influence the chosen geometric and steric descriptor much. A PBE optimisation could be done to reach the geometric/steric accuracy fast. Then getting the electronic accuracy from PBE can be done by a PBE0 single-point calculation or a full PBE0 optimisation.

Another option is based on previous research by the ISE group, it was found that for a part of the complexes, the GFN2-xTB would also give a close correlation for bite angle and buried volume [23]. It is thus interesting to see if combining it with the PBE0 single-point calculation will cause a more accurate result.

Table 4.3 shows in the left column a short description of the optimisation method and on the right the official notation to describe that optimisation.

optimisations	Notation
PBE0 single point calculation after PBE optimisation	sp-PBE0/def2-SV(P) // PBE/def2-SV(P)
PBE0 optimisation after PBE optimisation	PBE0/def2-SV(P) // PBE/def2-SV(P)
single point PBE0 calculation after GFN2-xTB optimisation	sp-PBE0/def2-SV(P) // GFN2-xTB

Table 4.3: Showing the process done and the notation used to describe that process

The different methods will be compared on the cumulative time needed for the complete optimisation and the result of the descriptors against the result from a PBE0/def2-SV(P) optimisation. The time is as previously compared using a histogram and the descriptors using violin plots and scatter plots.

4.3.1. Comparison of time for combinations of optimisation methods

The histogram in figure 4.6 shows the time needed for the different recommendations and the time required for PBE0/def2-SV(P) and PBE/def2-SV(P). The time is the total CPU hours needed to complete the whole calculation.

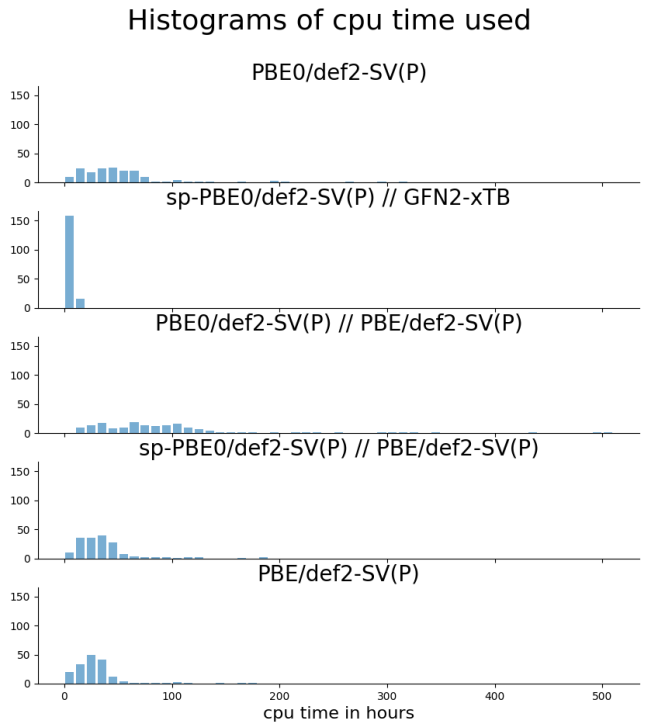
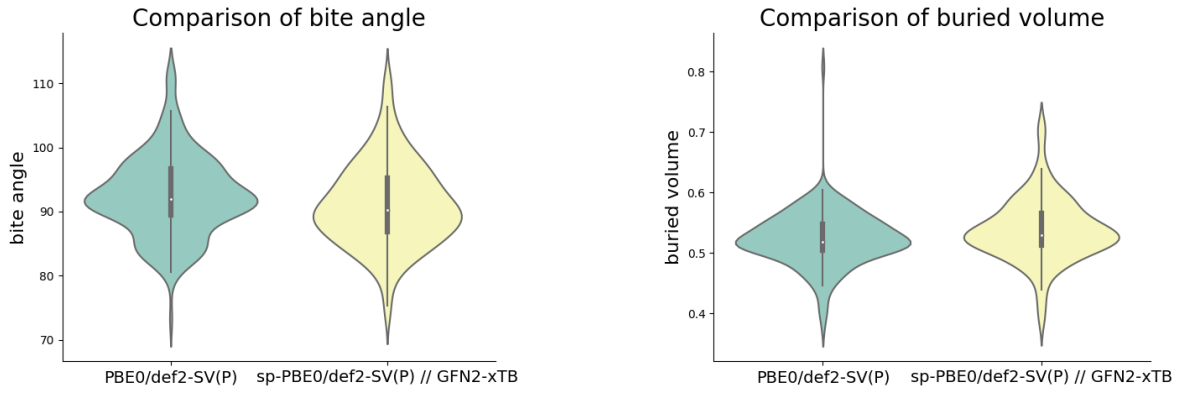


Figure 4.6: histogram of time needed for different functionals or combined calculations. For the combined calculations, the histogram shows the cumulative time. The bin size is 10 CPU hours.

In the figure, the middle three histograms are of combined methods, the PBE0/def2-SV(P) and PBE/def2-SV(P) are put in as references. sp-PBE0/def2-SV(P) // GFN2-xTB is by far the fastest method only needing 20 hours to complete all optimisations. On the other side adding the PBE0 optimisation after the PBE optimisation is not worth it as it takes longer than just running the PBE0 optimisation on its own. The PBE0 single-point calculation after the PBE optimisation is a bit slower than just a PBE optimisation but is still faster than doing only a PBE0 optimisation. As the PBE0/def2-SV(P) // PBE/def2-SV(P) is slower than the PBE0 optimisation it will not be taken into consideration when comparing the descriptors. The other two combinations are, however, still viable options.



(a) (b)

Figure 4.7: Violin plots comparing the results of the different optimisation methods for the ligands, comparing PBE0 optimisation to a PBE0 single-point calculation after a GFN2-xTB optimisation (a) Violin plots of the bite angle for the optimised ligands. (b) Violin plots of the buried volume with a radius of 3.5Å for optimised ligands.

4.3.2. Comparison of descriptors for combinations of optimisation methods.

As the single-point calculation only changes the electronic descriptor, the bite angle, and buried volume do not have to be compared against PBE0 for the sp-PBE0/def2-SV(P) // PBE/def2-SV(P) calculation. GFN2-xTB has not been used previously in this study so it will be compared for all descriptors. Figure 4.7 shows the violin plots for PBE0 optimisation and the single-point PBE0 calculation after GFN2-xTB optimisation for the bite angle and the buried volume. It can be seen that they have values that differ a bit and very different distributions. This results in weak correlations with a r^2 of 0.73 for the bite angle and 0.54 for the buried volume. This is much lower than expected as the previous research indicated values of 0.85 and 0.9, respectively [23]. A reason for this can be the fact that the previous research only used structures that were represented as SMILES and thus simpler. The current set of ligands is more extensive and more complex.

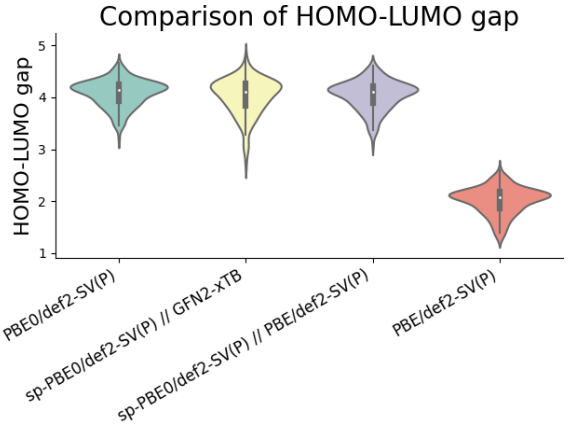


Figure 4.8: Violin plots of the HOMO-LUMO gap for each optimisation method.

Figure 4.8 below shows each method's violin plots of the HOMO-LUMO gap. Both recommendations are closer to PBE0 than PBE, which is good. As the sp-PBE0/def2-SV(P) // PBE/def2-SV(P) looks to have the same distribution as the PBE0/def2-SV(P) optimisation, it seems like that may be the best option. To really compare the recommendations, it is necessary to look at the scatter plots as shown in figure 4.9.

Figure 4.9a combined with the results from the other descriptors shows that sp-PBE0/def2-SV(P) // GFN2-xTB is not a usable option to replace PBE0 optimisation as there is a lot of difference in the values with for the HOMO-LUMO gap an ANOVA p-value of 0.055 which is just at the edge of being statistically different. Furthermore, the points do not fall on a line, meaning they do not correlate and are, therefore, not helpful. Looking at 4.9b shows this is a perfect option. There is a very strong correlation between these results. There are some differences in values, but because of the strong correlation, this does not matter. Comparing it with

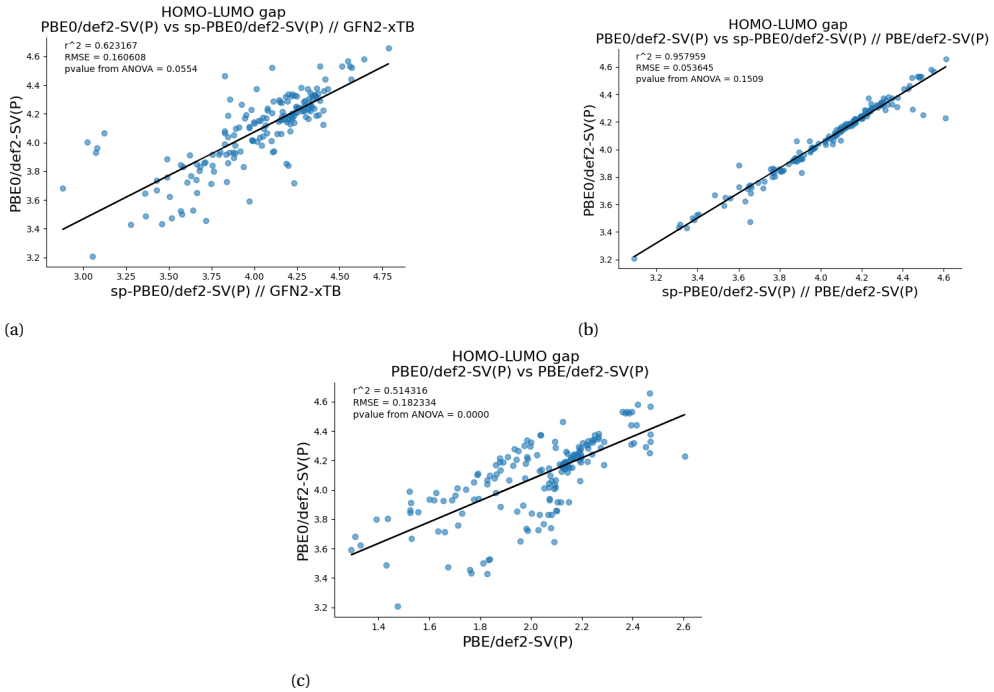


Figure 4.9: Scatter plots of the HOMO-LUMO gap of PBE0/def2-SV(P) against different methods. (a) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against sp-PBE0/def2-SV(P) // GFN2-xTB. (b) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against sp-PBE0/def2-SV(P) // PBE/def2-SV(P). (c) Scatter plot showing the correlation and ANOVA result for PBE0/def2-SV(P) against PBE/def2-SV(P) as comparison

the scatterplot where just PBE optimisation was done without the single-point calculation in figure 4.9c, an obvious improvement can be seen.

From this in combining the optimisation methods a faster optimisation time was achieved with the same level of accuracy as PBE0/def2-SV(P)

5

Conclusions and Outlook

5.1. Conclusions

This study was done to find the basis set and functional that had the best balance between time and accuracy, for the optimization of 192 TM complexes. The comparisons between the different methods were made with the CPU time needed for optimization and a selected set of descriptors (bite angle, buried volume, HOMO-LUMO gap) were used for comparison between the different methods.

The first question to answer was, do the more extensive def2 basis sets achieve more accurate results, and are they still relatively low in CPU time? The second problem was which functional between PBE, TPSS, PBE0, B3LYP and MN15 would provide the best balance between time and accurate results. The third part of the study was if combining different optimization methods could lead to a faster and as an accurate result as just the best performing one-time optimization.

Comparing the three different basis sets from the def2 family from the Ahlrichs group; def2-SV(P), def2-TZVPP, and def2-QZVPP. Optimizing the structures using the def2-QZVPP basis set took very long, with only one structure optimized before 10000 CPU hours. Therefore in comparing the basis sets for possible later uses, it does not stand up to the other two as the optimizations take so long that using it is just not effective.

From the other two basis sets, there was a significant gap in the time needed to optimize, with def2-TZVPP taking way longer than def2-SV(P). Due to time constraints on the project, only 88 structures were compared between the two basis sets, as waiting to have all the structures converge with def2-TZVPP would take too long. Comparing the three descriptors for the 88 structures, all the results were surprisingly similar. The datasets from def2-SV(P) and def2-TZVPP had a strong correlation for all three descriptors, with a r^2 of 0.99, 0.99 and 0.96 for the bite angle, buried volume and HOMO-LUMO gap respectively. Each also had a p-value above 0.05 meaning that the datasets were not statistically different. For all three descriptors, there was no statistical difference between the two datasets. Thus from the tested basis sets, def2-SV(P) is the most practical to use, as selecting def2-TZVPP will result in longer optimization times but no clear improvement in the descriptors.

In the second part of this research using def2-SV(P) as the basis set, five different functionals were compared. One GGA functional, PBE, one mGGA functional, TPSS, and three hybrid functionals, PBE0, B3LYP and MN15. The time in which the bulk of the structures was done, showed the following ranking, PBE took the shortest amount of time, then TPSS, then PBE0, then B3LYP and MN15 took the longest amount of time. The ranking follows the order of the level of theory and parameterization of the functionals, a higher level of theory, and a higher level of parameterization resulting in a longer time.

Comparing the descriptors for the different functionals gave surprising results. The bite angle and buried volume were the same across all functionals. The HOMO-LUMO gap was different for each functional with three levels, PBE and TPSS on one level, PBE0 and B3LYP on another, and lastly MN15 on a separate level. As PBE and TPSS have a lower level of theory, the results from those were deemed less accurate. They were also not usable as they did not have a strong correlation to the HOMO-LUMO gaps of PBE0. As B3LYP and PBE0 were on a similar level and MN15 was separate, MN15 was also deemed inaccurate. Choosing between B3LYP and PBE0 for the best functional for these TM-complexes came down to time.

The best option from the benchmarking is thus a PBE0/def2-SV(P) optimization. Since the geometric/steric descriptors were the same across all functionals, and only the electronic varied some combinations

of optimization may be interesting.

Combining optimization methods was done based on two factors, the first was the fact that by first doing a PBE optimization the geometric and steric descriptors are already accurate using a fast method. Adding a PBE0 optimization or a single-point PBE0 calculation can lead to the electronic descriptor being as accurate as a full PBE0 optimization in less time. The other factor was previous research done in the ISE group suggesting that using GFN2-xTB would lead to an already good correlating bite angle and buried volume. The same goes than as was mentioned above, using a fast method to calculate the geometric and steric descriptors and then adding another method to calculate the more accurate HOMO-LUMO gap.

Doing a PBE0 optimization after a PBE optimization was slower than only doing a PBE0 optimization and is thus not useful. The other two options were faster than a PBE0 optimization. The descriptors showed however that sp-PBE0/def2-SV(P) had a very strong correlation with the HOMO-LUMO gap of the PBE0 optimization, with a r^2 of 0.96. The p-value was quite low with 0.15 but that does not really matter as the correlation is so strong. From there, the recommendation for a single-point PBE0 calculation on the PBE-optimized structures was made. This resulted in similar results to PBE0 but in a way shorter time. The best way to optimize these TM complexes to balance accuracy and time is by doing a PBE/def2-SV(P) optimization and running a single-point PBE0 calculation on those optimized structures.

5.2. Outlook

From the results in this thesis, a clear conclusion was drawn, the fastest accurate way to optimize a molecule is by performing a sp-PBE0/def2-SV(P) // PBE/def2-SV(P) optimization.

This study only looked at a small subset of functionals and basis sets. A follow-up study could look at even less extensive basis sets such as def2-SV to see if there is a difference between those results, or basis sets from a completely different family. Another follow-up could be checking if using GFN2-xTB with a full PBE0 optimization would give better results than the one now found for GFN2-xTB and if the time would still be faster than a PBE0 optimization.

This study was done looking at only three descriptors, while there are many more calculated. Some tests could be done to see if the results from these three carry over towards the other descriptors.

5.2.1. Machine-learning

The reason for benchmarking based on descriptors was that the descriptors are needed in a machine-learning workflow of the ISE group. Depending on the goal of the machine learning workflow maybe only a PBE optimization needs to be done if it only looks at geometric and steric descriptors, this would also save time. An important part in saving time is also that when doing the full optimization it would be better if it was so coded that once an optimization completes the single-point calculation starts immediately instead of waiting for everything to finish as that would make sure the calculations keep being fast.

5.2.2. Different programs

This study was fully done on Gaussian except for the GFN2-xTB calculation. Trying to see if using a different programme might be an option could be good for research. There are programmes which are more easily accessible as they are free or open-source, or would be more easily integrated in a Machine learning workflow.

Bibliography

[1] KI Ramachandran, Gopakumar Deepa, and Krishnan Namboori. *Computational chemistry and molecular modeling: principles and applications*. Springer Science Business Media, 2008. ISBN 3540773045.

[2] Xiaoyue Ma. Development of computational chemistry and application of computational methods. *Journal of Physics: Conference Series*, 2386(1):012005, 2022. ISSN 1742-6596 1742-6588. doi: 10.1088/1742-6596/2386/1/012005. URL https://dx.doi.org/10.1088/1742-6596/2386/1/012005.

[3] Errol Lewars. Introduction to the theory and applications of molecular and quantum mechanics. *J. Computational chemistry, Ontario Canada*, 2003.

[4] Seihwan Ahn, Mannkyu Hong, Mahesh Sundararajan, Daniel H Ess, and Mu-Hyun Baik. Design and optimization of catalysts based on mechanistic insights derived from quantum chemical reaction modeling. *Chemical reviews*, 119(11):6509–6560, 2019. ISSN 0009-2665.

[5] S. Ekins, A. M. Clark, K. Dole, K. Gregory, A. M. McNutt, A. C. Spektor, C. Weatherall, N. K. Litterman, and B. A. Bunin. Data mining and computational modeling of high-throughput screening datasets. *Methods Mol Biol*, 1755:197–221, 2018. ISSN 1064-3745 (Print) 1064-3745. doi: 10.1007/978-1-4939-7724-6_14.

[6] Lu Zhang, Jianjun Tan, Dan Han, and Hao Zhu. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discovery Today*, 22(11):1680–1685, 2017. ISSN 1359-6446. doi: https://doi.org/10.1016/j.drudis.2017.08.010. URL https://www.sciencedirect.com/science/article/pii/S1359644616304366.

[7] Xiaoqian Lin, Xiu Li, and Xubo Lin. A review on applications of computational methods in drug screening and design. *Molecules*, 25(6):1375, 2020. ISSN 1420-3049.

[8] Frank Jensen. Computational chemistry: The exciting opportunities and the boring details. *Israel Journal of Chemistry*, 62(1-2):e202100027, 2022. ISSN 0021-2148. doi: https://doi.org/10.1002/ijch.202100027. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/ijch.202100027.

[9] Christopher J Cramer. *Essentials of computational chemistry: theories and models*. John Wiley Sons, 2013. ISBN 1118712277.

[10] Ho Ryu, Jiyong Park, Hong Ki Kim, Ji Young Park, Seoung-Tae Kim, and Mu-Hyun Baik. Pitfalls in computational modeling of chemical reactions and how to avoid them. *Organometallics*, 37(19):3228–3239, 2018. ISSN 0276-7333. doi: 10.1021/acs.organomet.8b00456. URL https://doi.org/10.1021/acs.organomet.8b00456.

[11] John A. Keith, Valentin Vassilev-Galindo, Bingqing Cheng, Stefan Chmiela, Michael Gastegger, Klaus-Robert Müller, and Alexandre Tkatchenko. Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chemical Reviews*, 121(16):9816–9872, 2021. ISSN 0009-2665. doi: 10.1021/acs.chemrev.1c00107. URL https://doi.org/10.1021/acs.chemrev.1c00107.

[12] Fernand Spiegelman, Nathalie Tarrat, Jérôme Cuny, Leo Dontot, Evgeny Posenitskiy, Carles Martí, Aude Simon, and Mathias Rapacioli. Density-functional tight-binding: basic concepts and applications to molecules and clusters. *Advances in Physics: X*, 5(1):1710252, 2020. ISSN null. doi: 10.1080/23746149.2019.1710252. URL https://doi.org/10.1080/23746149.2019.1710252.

[13] Pekka Koskinen and Ville Mäkinen. Density-functional tight-binding for beginners. *Computational Materials Science*, 47(1):237–253, 2009. ISSN 0927-0256. doi: https://doi.org/10.1016/j.commatsci.2009.07.013. URL https://www.sciencedirect.com/science/article/pii/S0927025609003036.

[14] W. c. Feng and K. Cameron. The green500 list: Encouraging sustainable supercomputing. *Computer*, 40 (12):50–55, 2007. ISSN 1558-0814. doi: 10.1109/MC.2007.445.

[15] Thomas Bligaard, R. Morris Bullock, Charles T. Campbell, Jingguang G. Chen, Bruce C. Gates, Raymond J. Gorte, Christopher W. Jones, William D. Jones, John R. Kitchin, and Susannah L. Scott. Toward benchmarking in catalysis science: Best practices, challenges, and opportunities. *ACS Catalysis*, 6(4):2590–2602, 2016. doi: 10.1021/acscatal.6b00183. URL https://doi.org/10.1021/acscatal.6b00183.

[16] Antti Siiskonen and Arri Priimagi. Benchmarking dft methods with small basis sets for the calculation of halogen-bond strengths. *Journal of Molecular Modeling*, 23(2):50, 2017. ISSN 0948-5023. doi: 10.1007/s00894-017-3212-4. URL https://doi.org/10.1007/s00894-017-3212-4.

[17] Lars Goerigk, Andreas Hansen, Christoph Bauer, Stephan Ehrlich, Asim Najibi, and Stefan Grimme. A look at the density functional theory zoo with the advanced gmtkn55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Physical Chemistry Chemical Physics*, 19(48):32184–32215, 2017. ISSN 1463-9076. doi: 10.1039/C7CP04913G. URL http://dx.doi.org/10.1039/C7CP04913G.

[18] Lars Goerigk and Stefan Grimme. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Physical Chemistry Chemical Physics*, 13(14):6670–6688, 2011. ISSN 1463-9076. doi: 10.1039/C0CP02984J. URL http://dx.doi.org/10.1039/C0CP02984J.

[19] Danishuddin and Asad U. Khan. Descriptors and their selection methods in qsar analysis: paradigm for drug design. *Drug Discovery Today*, 21(8):1291–1302, 2016. ISSN 1359-6446. doi: https://doi.org/10.1016/j.drudis.2016.06.013. URL https://www.sciencedirect.com/science/article/pii/S1359644616302318.

[20] Yuri A. Aoto, Ana Paula de Lima Batista, Andreas Köhn, and Antonio G. S. de Oliveira-Filho. How to arrive at accurate benchmark values for transition metal compounds: Computation or experiment? *Journal of Chemical Theory and Computation*, 13(11):5291–5316, 2017. ISSN 1549-9618. doi: 10.1021/acs.jctc.7b00688. URL https://doi.org/10.1021/acs.jctc.7b00688.

[21] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.

[22] Haoyu S. Yu, Xiao He, Shaohong L. Li, and Donald G. Truhlar. Mn15: A kohn–sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chemical Science*, 7(8):5032–5051, 2016. ISSN 2041-6520. doi: 10.1039/C6SC00705H. URL http://dx.doi.org/10.1039/C6SC00705H.

[23] Aydin Hossaini. Transferability of descriptors for in silico catalyst screening. 2023.

[24] Erwin Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical review*, 28 (6):1049, 1926.

[25] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927. ISSN 0003-3804. doi: https://doi.org/10.1002/andp.19273892002.

[26] David C Young. A practical guide for applying techniques to real-world problems. *Computational Chemistry, New York*, 9:390, 2001.

[27] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3B):B864–B871, 1964. doi: 10.1103/PhysRev.136.B864. URL https://link.aps.org/doi/10.1103/PhysRev.136.B864.

[28] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133–A1138, 1965. doi: 10.1103/PhysRev.140.A1133. URL https://link.aps.org/doi/10.1103/PhysRev.140.A1133.

[29] John P Perdew and Karla Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings*, volume 577, pages 1–20. American Institute of Physics. ISBN 0735400164.

[30] Ann E. Mattsson. In pursuit of the "divine" functional. *Science*, 298(5594):759–760, 2002. doi: doi: 10.1126/science.1077710.

[31] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters.*, 77(18):3865, 1996. ISSN 0031-9007.

[32] Jianmin Tao, John P. Perdew, Viktor N. Staroverov, and Gustavo E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, 91(14):146401, 2003. doi: 10.1103/PhysRevLett.91.146401. URL https://link.aps.org/doi/10.1103/PhysRevLett.91.146401.

[33] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, 1999. ISSN 0021-9606. doi: 10.1063/1.478522. URL https://doi.org/10.1063/1.478522.

[34] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994. ISSN 0022-3654. doi: 10.1021/j100096a001. URL https://doi.org/10.1021/j100096a001.

[35] Martin Stöhr, Troy Van Voorhis, and Alexandre Tkatchenko. Theory and practice of modeling van der waals interactions in electronic-structure calculations. *Chemical Society Reviews*, 48(15):4118–4154, 2019. ISSN 0306-0012. doi: 10.1039/C9CS00060G. URL http://dx.doi.org/10.1039/C9CS00060G.

[36] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15), 2010. ISSN 0021-9606. doi: 10.1063/1.3382344. URL https://doi.org/10.1063/1.3382344.

[37] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem*, 32(7):1456–65, 2011. ISSN 0192-8651. doi: 10.1002/jcc.21759.

[38] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian~16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[39] Dutch national supercomputer snellius. URL https://www.surf.nl/en/dutch-national-supercomputer-snellius. Accessed: 20-06-2023.

[40] Obelix. URL https://github.com/EPiCs-group/obelix. Accessed: June 2023.

[41] Lauriane Jacot-Descombes, Lucas Turcani, and Kjell Jorner. morfeus. URL https://github.com/digital-chemistry-laboratory/morfeus.

[42] N O'Boyle, A. L. Tenderholt, and K. M. Langner. cclib. URL https://github.com/cclib/.

[43] Rylan J. Lundgren and Mark Stradiotto. *Key Concepts in Ligand Design*, pages 1–14. 2016. doi: https://doi.org/10.1002/9781118839621.ch1. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118839621.ch1.

[44] P. W. van Leeuwen, P. C. Kamer, J. N. Reek, and P. Dierkes. Ligand bite angle effects in metal-catalyzed c-c bond formation. *Chem Rev*, 100(8):2741–70, 2000. ISSN 0009-2665. doi: 10.1021/cr9902704.

[45] Hervé Clavier and Steven P. Nolan. Percent buried volume for phosphine and n-heterocyclic carbene ligands: steric properties in organometallic chemistry. *Chemical Communications*, 46(6):841–861, 2010. ISSN 1359-7345. doi: 10.1039/B922984A. URL http://dx.doi.org/10.1039/B922984A.

[46] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

[47] Chemcraft - graphical software for visualization of quantum chemistry computations. URL https://www.chemcraftprog.com. Version 1.8, build 648.

[48] N Novikov, M Maslov, K Katin, and V Prudkovskiy. Effect of dft-functional on the energy and electronic characteristics of carbon compounds with the unconventional geometry of the framework. *Letters on materials*, 7(4):433–436, 2017. ISSN 2218-5046.

A

Extra figures

A.1. Benchmarking the basis sets

Graphs comparing the different basis sets

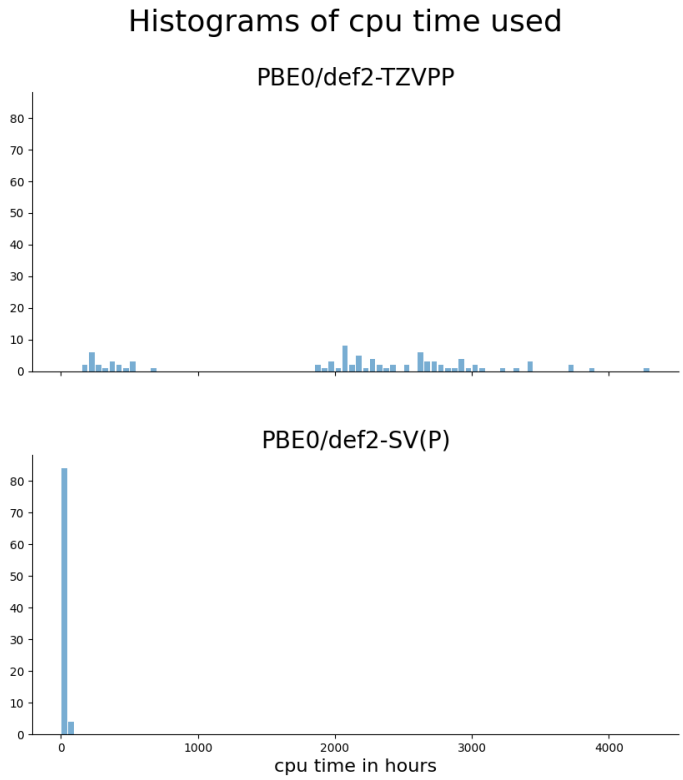
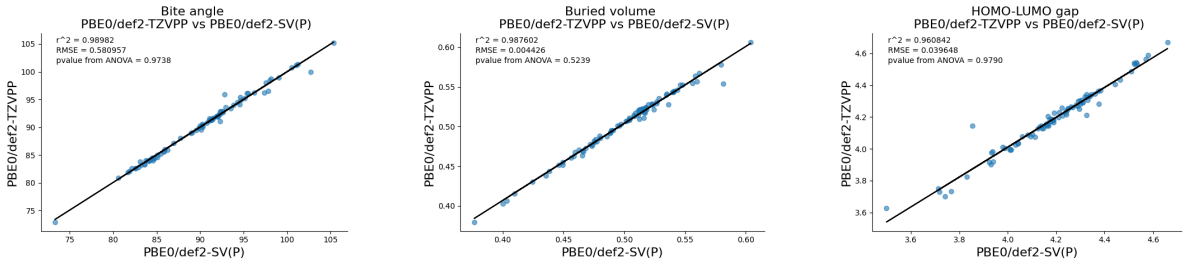


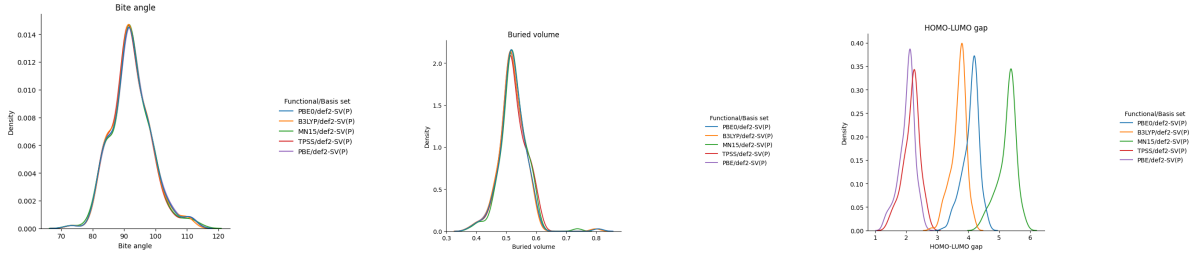
Figure A.1: Time taken for the optimizations to converge. For the 88 ligands that first converged with the def2-TZVPP basisset. The size of the bins in the histogram is 50 cpu hours



(a) (b) (c)

Figure A.2: Scatterplots showing the correlation and p-value for the 88 ligands optimized with basis set def2-SV(P) vs def2-TZVPP for specific descriptors (a) Results of the bite angle for the optimized ligands. (b) Results of the buried volume with a radius of 3.5 Å for optimized ligands. (c) Results of the HOMO-LUMO gap for the optimized ligands.

A.2. Benchmarking the different functionals

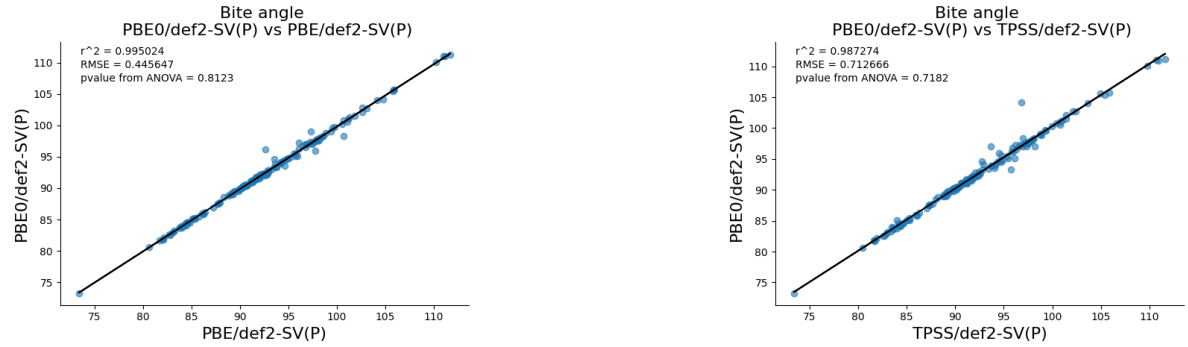


(a) (b) (c)

Figure A.3: Density plots for each descriptor for the different methods. (a) Densities of the bite angle for the different methods. (b) Densities of the buried volume with a radius of 3.5 Å for the different methods (c) Densitiees of the HOMO-LUMO gap for the different methods.

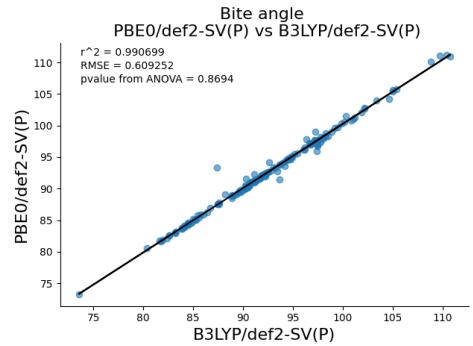
A.2.1. Scatter plots

Bite angle

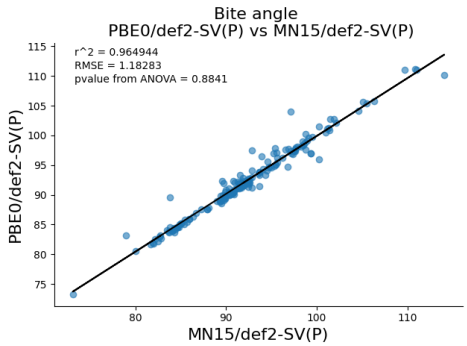


(a)

(b)



(c)



(d)

Figure A.4: Scatter plots showing the correlation and p-value for the bite angle of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15

Buried volume

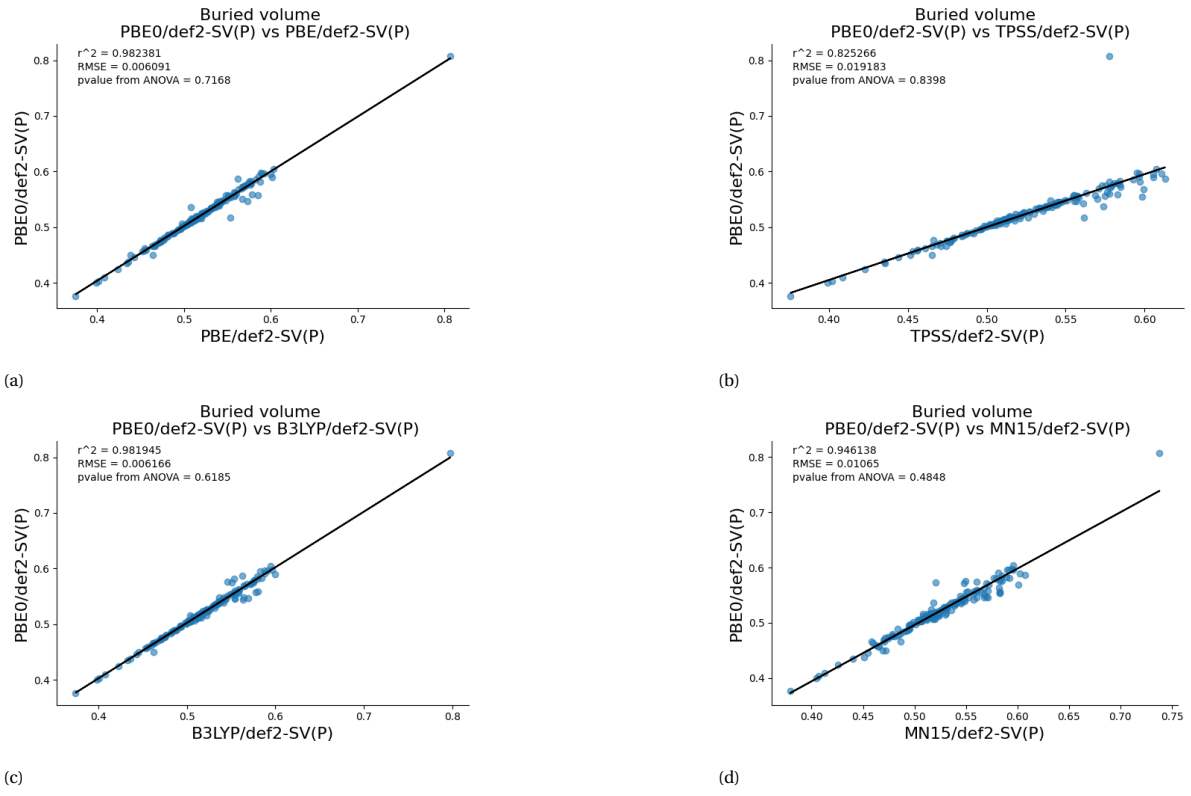


Figure A.5: Scatter plots showing the correlation and p-value for the buried volume with a radius of 3.5 Å of the different functionals against PBE0. (a) PBE0 vs. PBE (b) PBE0 vs. TPSS (c) PBE0 vs. B3LYP (d) PBE0 vs MN15