The influence of ligand configurations on TM-complex chemical space exploration

Niels van der Lem

TU Delft

The influence of ligand configurations on TM-complex chemical space exploration

by

Niels van der Lem

to obtain the degree of Master of Science
at the Delft University of Technology,
to be publicly defended on Friday, February 9, 2024 at 10:00 AM.

Performed at:
Inorganic Systems Engineering
Faculty of Applied Sciences

Under supervision of:
Prof. Dr. E. A. Pidko
A.V. Kalikadien, MSc.

Student number:	5188903	
Project duration:	September 7, 2023 – February 9, 2024	
Thesis committee:	Prof. Dr. E. A. Pidko,	TU Delft, AS, supervisor
	Prof. Dr. F. C. Grozema,	TU Delft, AS
	Dr. G. A. Filonenko,	TU Delft, 3ME

TU Delft

Abstract

Various industries rely upon transition metal complexes to efficiently catalyze chemical reactions. These transition metal complexes often consist of precious metals, which are scarce and expensive. Therefore, a shift towards catalysts containing earth abundant metals is necessary. Computational catalyst screening has the potential to accelerate this shift by reducing catalyst discovery time. The first step of computational catalyst screening consists of obtaining a digital representation of the catalyst. Usually, a fixed a priori ligand configuration of TM-complexes is assumed to represent the catalyst. However, this approach of catalyst representation might not capture the influence of alternative ligand configurations on the observed catalytic behavior.

In the context of high-throughput *in-silico* catalyst screening, this study aims to evaluate the influence of ligand configurations on the stability and physical-chemical properties of transition metal complexes. An automated workflow for the generation of complexes, complex sorting based on ligand configuration, DFT geometry optimization and descriptor extraction is employed. Ensembles of ligand configurations are generated for iridium(III), ruthenium(II) and manganese(I) complexes featuring 88 bisphosphine bidentate ligands. Based on DFT optimized geometries, analysis reveals a preference for a specific ligand configuration for iridium(III) complexes. However, this preference is not observed for ruthenium(II) and manganese(I) complexes. Furthermore, for the majority of ruthenium(II) and manganese(I) complexes multiple ligand configurations are found within a 10 kJ/mol range from the most favorable one. The analysis of physical-chemical of thermodynamic, electronic, geometric and steric descriptors reveals that none of the descriptors consistently correlates to the preferred ligand configuration. These findings indicate that a priori selection of ligand configuration may result in insufficient coverage and representation of key catalyst features for predictive *in-silico* chemical space exploration.

Contents

List of Figures

Acronyms

BO	Born-Oppenheimer
BV	Buried volume
DFT	Density functional theory
FF	Force field
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
HTE	High-throughput experimentation
LUMO	Lowest unoccupied molecular orbital
NBO	Natural bonding orbital
OH	Octahedral
PES	Potential energy surface
QSAR	Quantitative structure-activity relationship
QM	Quantum mechanics
SMILES	Simplified molecular-input line-entry system
SP	Square planar
TM	Transition metal

1

Introduction

The chemical industry is responsible for providing the world with many necessities of modern society, including medicine, food and materials. In these industries, catalysis plays a fundamental role, with estimates indicating that over 90% of chemical products are produced utilizing catalysts [1]. A catalyst is a chemical substance increasing the reaction rate, without being consumed. Furthermore, catalysts may allow reactions to proceed through new reaction pathways with lower energy barriers. These advantage are especially important for sustainable chemistry, which focuses on minimizing energy expenditure and waste generation [2]. In general, catalysts can be divided into three categories, each presenting their own advantages and disadvantages. These categories are heterogeneous, homogeneous and biocatalysis.

Heterogeneous catalysis stands out as the most applied type of catalysis. In heterogeneous catalysis, the catalyst exists in a different phase from the reactants. This type of catalyst has the principal advantage of easy separation and recovery of the catalyst [3]. However, heterogeneous catalysis generally suffers from mass transfer limitations. Furthermore, reaction conditions may reach extreme levels of temperature and pressure. In contrast to heterogeneous catalysis, homogeneous catalysts exist in the same phase as the reactants, which has the general benefit of achieving high selectivity and activity under mild reaction conditions [4]. These advantages are important for reactions, typically difficult to perform, such as cross-coupling and hydrogenation [5, 6]. Similar to homogeneous catalysts, biocatalysts, which primarily consists of enzymes, also coexist in the same phase as the reactants [7]. Nonetheless, due to large differences in the nature of the catalysts, biocatalysis is recognized as a separate class of catalysis.

A versatile class of homogeneous catalysts consists of organometallic complexes, which are characterized by a transition metal center (TM) alongside coordinated ligands. These ligands provide stability and reactivity to the metal [8]. Phosphine ligands are often used due to their ability to provide flexibility in modification of steric and electronic properties of the complex [9]. Bisphosphine bidentate ligands (PP) in particular, excel at influencing the substrate coordination sites by effectively blocking off parts of the metal center. In addition to ligands, metals are also crucial in influencing TM-complex properties, as metals in different rows and groups of the periodic table display differences in properties. Metals in different rows exhibit differences in atomic orbital energies [10]. Additionally, metals in different groups differ in electron configuration of the d electrons, leading to variation in energy reduction due to complex formation[10, 11].

TM-complexes play an important role as catalysts in hydrogenation reactions. These hydrogenation reactions are employed to synthesise various chemicals, ranging from pharmaceuticals to

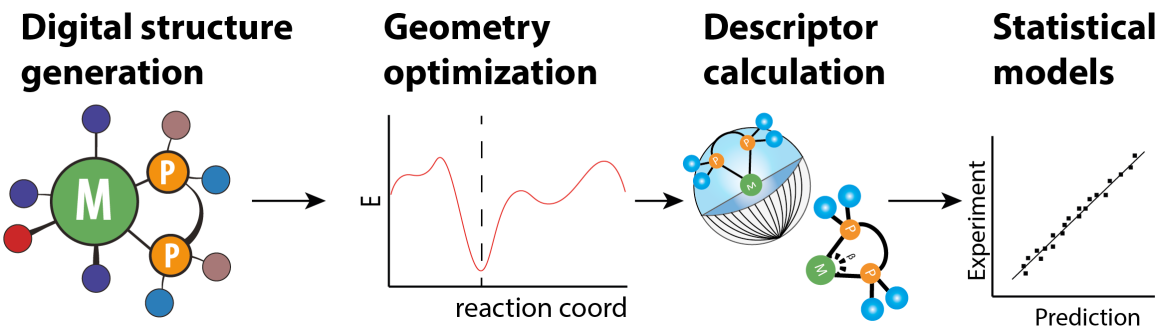


Figure 1.1: General scheme of main components of computational workflows for HT catalyst discovery and screening

fuels. [12, 13]. As an example, the asymmetric hydrogenation of nitriles provides an efficient pathway to produce amines, integral compounds for the pharmaceutical and agrochemical industries [14]. Catalysts for nitrile hydrogenation need the right combination of metal and ligands to achieve high selectivity and minimum waste product generation [15]. Therefore, precious metals (Pd, Pt, Rh, Ru, Ir) are frequently employed as the metal center in these hydrogenation catalysts [16, 17]. However, the scarceness and high cost of precious metals makes a shift towards earth-abundant metals, such as manganese and cobalt, necessary [18, 19].

The substitution of precious metals with earth-abundant alternatives poses a significant challenge, due to differences in metal properties. Therefore, novel catalysts have to be discovered, using ligands as a bridge to address variations in properties of different metals [20]. However, the available metal-ligand combinations are nearly endless. Various methods are employed to explore the chemical space, such as the popular high throughput experimentation (HTE) approach. During these HTE approaches, a large number of reactions are performed in parallel, each performed by different catalysts. This approach allows many catalysts to be effectively analyzed on a specific target, such as selectivity. [21]. Despite the effectiveness of HTE, this method is still time consuming due to the dependency on trial and error. A potential approach to reduce catalyst discovery time is the implementation of HT computational chemical space exploration.

Different methods are employed for HT computational exploration of the chemical space [22–24]. In general, the computational workflows contain four main components, namely: structure generation, structure optimization, descriptor extraction, and finally statistical modeling relating descriptors with catalytic behavior [25]. Figure 1.1 visualizes the main components of computational workflows for digital catalyst discovery and screening. First, an initial digital representation of the catalyst is generated. Subsequently, geometry of the generated catalyst is optimized by searching for a minimum on the potential energy surface (PES). This optimization can be performed using various methods, such as density functional theory (DFT). The descriptors of the optimized geometries are obtained and used to build statistical models relating catalyst properties to catalytic performance through quantitative structure-activity/selectivity relations (QSAR/QSSR) [22–24, 26].

The digital catalysts representation may be the most crucial part of any HT computational workflow for chemical space exploration, as the selected structure greatly influences subsequent workflow outcomes. In octahedral TM-complexes, ligands can be coordinated to the metal center in various arrangements, as shown in Figure 1.2. This ligand arrangement influences the interaction between metal center and substrate, for example, through the trans-effect. The trans-effect describes the influence of a ligand on the rate of substitution of the ligand trans to itself [27]. The

strength of the trans-effect depends primarily on electronic factors of the ligands, such as the type of metal-ligand bonding [28]. Generally, computational workflows a priori assume a specific ligand arrangement based on an energy ranking or intuition, thus neglecting parts of the chemical space.

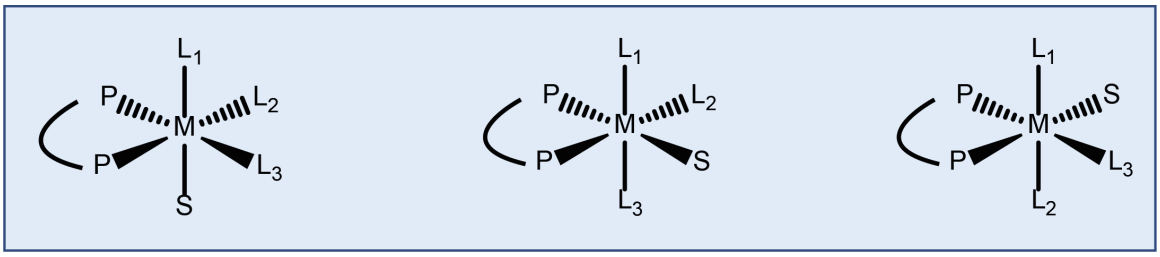


Figure 1.2: Set of possible ligand configurations for an octahedral TM-complex featuring a bidentate ligand (PP), other ligands (L) and a substrate (S)

Different ligand configurations of TM-complexes may be similar in stability. Following a Boltzmann distribution, these metastable ligand configurations may coexist under reaction conditions. The ligand configurations present in low concentrations may provide favorable reaction pathways, leading to a major influence on the observed catalytic behavior [29–35]. Therefore, the question arises whether virtual catalyst screening methods can correctly consider catalytic systems when part of the chemical space is neglected due to the initial human choices and intuition.

In the context of high-throughput *in silico* hydrogenation catalyst screening, this study aims to evaluate the influence of ligand configurations on the stability and physical-chemical properties of transition metal complexes. An automated workflow for for the generation of complexes, complex sorting based on ligand configuration, DFT geometry optimization and descriptor extraction is employed. Using this workflow, ensembles of ligand configurations for iridium, ruthenium and manganese complexes featuring 88 bisphosphine bidentate ligands are created. DFT-optimized geometries and relevant descriptors of the different generated complex ensembles are compared and analyzed to explain differences between different ligand configurations. The thesis starts by providing a theoretical background for applied computational methods. Afterwards, the main findings are discussed in the form of a paper. This paper also discusses applied methods and models. At last, a summary of the project is provided and the future outlook is discussed.

2

Theoretical background

In this chapter, theory behind applied geometry optimization methods is introduced. The applied methods were based on DFT and force fields. Afterwards, a discussion of investigated physical-chemical properties, also known as 'descriptors', is included.

2.1 Density Functional Theory

Density Functional Theory (DFT) is a cornerstlone in modern quantum-mechanical calculations. DFT separates itself from other quantum-mechanical methods by the reliance on electron density, instead of wavefunctions [36]. This reliance on electron density results in a method for quantum-mechanical calculations that can excellently balance efficiency and accuracy. Therefore, DFT is widely applied in the field of homogeneous catalysis [37–39]. Tasks for which DFT is often employed include geometry optimization, energy calculations and transition state searching [40]. In this study, DFT is the primary method for geometry optimization. The fundamental equation of DFT is the non-relativistic, time-independent Schrödinger equation, shown in Equation 2.1.

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

In this equation, $\hat{H}$ and ψ represent the Hamiltonian operator and wavefunction, respectively. To obtain solutions for the Schrödinger equation, the Hamiltonian operator has to be described. However, depending on the system, the Hamiltonian operator takes different forms. For a molecular system, consisting of N electrons and M nuclei, the Hamiltonian operator is expressed by the following equation [41, 42]:

$$\boldsymbol{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{i=1}^{N}\frac{\nabla_i^2}{M_A} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{A=1}^{M} r_{ij}^{-1} + \sum_{A=1}^{M}\sum_{A'>A}^{M}\frac{Z_{A'}Z_A}{R_{AA'}} \tag{2.2}$$

In this equation, Z_A represents the nucleus charge, r_{ij} is the distance between electrons, r_{iA} is the distance between electron and nucleus and $R_{AA'}$ represents the distance between nuclei. The Hamiltonian operator, as shown in Equation 2.2 consists of kinetic and potential energy terms. The first two terms are operators for the kinetic energy of electrons and nuclei, respectively. The last three terms are potential energy operators, representing coulomb attraction and repulsion terms between electrons and between nuclei.

The Hamiltonian, as depicted in Equation 2.2, enables the Schrödinger equation to be solved. However, analytical solutions can only be obtained for a one electron system [36]. Due to pairwise electron interactions, approximations are required to obtain solutions for the Schrödinger equation for larger systems. The Hamiltonian can be simplified through the Born-Oppenheimer (BO) approximation. The BO approximation states that nuclear and electronic motions can be decoupled, due to a significant weight difference between electrons and nuclei. Consequently, nuclei move much slower than electrons, leading to nuclei being much less reactive to changes in the system. As a result, the kinetic energy associated with the nuclei can be neglected from the Hamiltonian [43]. Applying these assumptions to Equation 2.2, results in the following simplified Hamiltonian.

$$\boldsymbol{H_{el}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{A=1}^{M} r_{ij}^{-1} \tag{2.3}$$

Different approaches, each with different assumptions, are employed to solver the simplified Hamiltonian depicted in Equation 2.3. The Hartree-Fock method for example, solves the optimized single electron wavefunctions by applying a single Slater determinant [44]. However, this method neglects electron correlations. In contrast, DFT does consider electron correlations by describing the electronic wavefunction in terms of electron density. This reliance on electron density reduces complexity, due to the dependency on only 3 variables, instead of 3N, for an N-electron system [45]. Hohenberg and Kohn first described two theorems to solve the many-electron wavefunction, known as the Hohenberg-Kohn theorems [46]. The first theorem states that all electronic properties are determined by the electron density of the ground state, while the second theorem establishes that the electronic energy of the system is inherently a functional of the density. This results in a minimum electronic energy for the ground state electron density. Together, these two theorems state that solely obtaining the density is sufficient to describe the molecular system. The energy functional can now be expressed in terms of density, by the Kohn-Sham equation [47]:

$$E[\rho(\boldsymbol{r})] = V_{AA} + V_{eA}[\rho(\boldsymbol{r})] + J[\rho(\boldsymbol{r})] + T[\rho(\boldsymbol{r})] + E_{XC}[\rho(\boldsymbol{r})] \tag{2.4}$$

In this equation, The repulsion between nuclei is denoted by V_{AA}, V_{eA} is the attraction term between electrons and nuclei, J represents the mean field electron-electron interaction, T is the kinetic energy functional and E_{XC} is the exchange-correlation functional.

2.1.1 Exchange-Correlation functionals

The exchange-correlation functional consists of all electron-electron interactions, outside of the Hartree-Fock mean field interactions. However, the exact form of the exchange-correlation functional is unknown, and only approximations can be used. Describing accurate exchange-correlation functionals is a challenge due to difficult-to-predict dependencies required for improved estimations. Generally, exchange-correlation functionals can be placed on the so-called Jacob's ladder, based on the dependencies of the functionals [48].

Figure 2.1 shows the levels of theory for the exchange-correlation functionals. Each increase in level of theory adds new dependencies, improving the accuracy of the exchange-correlation functional. However, computational cost also drastically increases with added dependencies. In literature, hybrid functionals are most often used, as the balance between chemical accuracy and computational cost is favorable [37]. In this research specifically, the PBE0 functional has been used [49]. PBE0 applies the Hartree-Fock and Perdew-Burke-Emzerhof (PBE) exchange energy in a 1:3

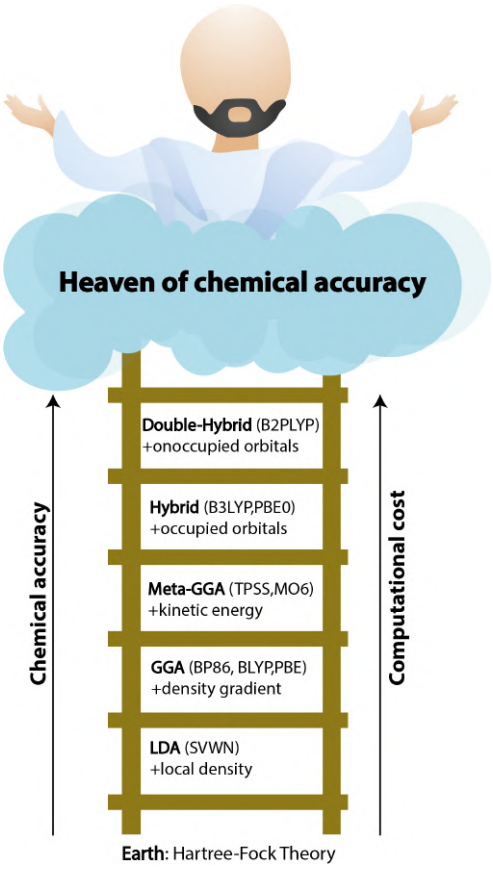


Figure 2.1: Jacob's ladder of DFT exchange-correlation functionals. Starting from the bottom, Hartree-Fock, dependencies are added along the ladder, increasing chemical accuracy and computational cost.

ratio, alongside full PBE correlation energy [50].

2.1.2 Basis sets

The atomic wave functions consist of a set of functions called the basis set. Similarly to the exchange-correlation functionals, basis sets apply approximations balancing accuracy and computational cost. The basis set transforms the mathematical form of molecular orbital (MO), resulting in equations which are efficiently solvable by computers. To determine MOs in the first place, the Linear Combination of Atomic Orbitals (LCAO) method can be applied. As the name implies, this method states that MOs can be expressed as a linear combination of atomic orbitals (AO). Most accurate basis sets are formed by applying Slater-type orbitals. These orbitals resemble hydrogen atomic orbitals, which are excellent for expanding MOs due to their correct shape both near and far from the nucleus. However, Slater functions are computationally expensive. For that reason, Gaussian type orbitals are preferred in practice. Multiple Gaussian type orbitals can be used to represent a single AO to improve accuracy. A valence double zeta basis set obtains two coefficients for an AO, describing both charge distribution near and far from the nucleus. By using a valence double zeta basis set, AOs of the same type are not identical and therefore distinguishable. The number of functions can be increased to triple or even quadruple zeta to improve accuracy and computational cost [51]. However, to obtain an accurate basis set, other types of functions also have to be

considered. Polarization functions account for polarization of electron density of the atom in the molecule. Diffuse functions are important for improving accuracy far away from the nucleus.
In this study, the double zeta def2-SVPP basis set and triple zeta def2-TZVPP basis set were used [52]. These basis sets contain extra polarization functions compared to the def2-SVP basis set. The triple zeta def2-TZVPP basis set contains three algebraically solvable orbital functions for each atomic orbital, resulting in higher accuracy. For that reason, the def2-SVPP basis set is used for geometry optimization. Afterwards, energies of optimized geometry are refined with the higher order def2-TZVPP basis set.

2.1.3 Dispersion correction

For large molecules, like transition metal complexes, London dispersion forces are essential to consider to improve the chemical accuracy of Kohn-Sham DFT [53]. Without correcting for the dispersion forces, most common hybrid functionals are not able to provide correct $-C_6R^{-6}$ dependence of the dispersion interaction energy on the molecular distance. Different methods have been made to take dispersion forces into account, with most methods being modifications of the so-called DFT-D method (DFT + dispersion). For this research, the DFT-D3 [54] is used, due to its high accuracy and broad range of application.

2.1.4 Geometry optimization

In this study, DFT is applied to perform geometry optimization of TM-complexes. The goal of geometry optimizations is to characterize the molecule configurations, for which a minimum total energy on the potential energy surface (PES) is found. Geometry optimizations are performed by providing a molecule structure, believed to be close to the minimum energy structure. This initial molecular geometry will in be changed until the potential energy gradient with respect to position equals zero, corresponding to a minimum on the PES. Moreover, the second derivatives of the total energy with respect to the configuration is determined in a hessian matrix. The Hessian matrix is examined to check whether negative (imaginary) frequencies are present in this matrix. If so, the found geometry is not actually a minimum on the PES.

2.2 Force Field methods

In order to create initial geometries for DFT geometry optimization, force field (FF) methods are applied. These FF methods employ simple empirical energy functions and parameterization, enabling molecular structure optimization at relatively low computational costs [55]. The parameters of FF methods are specifically optimized using experimental or quantum mechanical data of specific systems to accurately model those specific systems. For this research, the Universal Force Field (UFF) method is applied, due to the broad range of applicability and proven use for TM-complexes [56]. The UFF provides FF parameters for a large part of the periodic table, using the element, hybridization and connectivity [57].

2.3 Descriptors

In order to evaluate properties of DFT geometry optimized TM-complexes, mathematical representations of physical-chemical properties, called descriptors, can be obtained [58]. Each different complex and ligand configuration consists of a unique set of descriptors. These descriptors are regularly used to relate catalyst properties to catalyst characteristics through QSAR models [25]. For this study, descriptors are applied to study differences in material properties between different complex ligand configurations. Therefore, a selection of electronic, steric and geometric descriptors is obtained. A list of all studied descriptors, including their category and a description is available in the appendix.

2.3.1 Electronic descriptors

Electronic descriptors aim to describe the electronic structure properties of TM-complexes, providing insight into the reactivity and stability of TM-complexes. Examples of electronic descriptors are: dipole, dispersion, nucleophilicity, natural bonding orbital (NBO) charge and more. Explicitly studied electronic descriptors are the energy of the highest occupied molecular orbital (HOMO), energy of the lowest unoccupied molecular orbital (LUMO) and the energy gap (HOMO-LUMO gap) between them. These orbitals are prime reactivity and stability indicators, as they describe the electron accepting and donating capabilities of a molecule [59]. An illustration of the HOMO-LUMO orbitals and gap is shown in Figure 2.2.

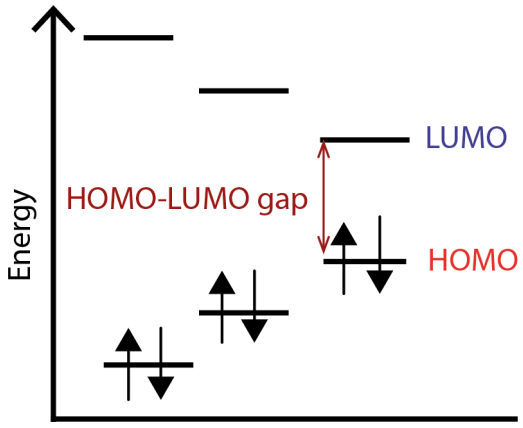


Figure 2.2: Illustration of the highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) and HOMO-LUMO gap

2.3.2 Steric descriptors

Steric effects describe non-bonding interactions between atoms in a molecule. Each atom in a molecule occupies a specific amount of space, leading to overlap in electron clouds of different atoms. In turn, this overlap results in repulsive inter-atomic interactions, influencing reactivity of the TM-complex [60]. Bisphosphine bidentate ligands may consist of many atoms, defining the direction of substrate-metal reactions through steric effects [9]. A key descriptor describing steric hindrance is the buried volume (BV), which describes the fraction of a sphere occupied by a specific ligand, using the metal center as a core [61]. An illustration of the buried volume is shown in

Figure 2.3.

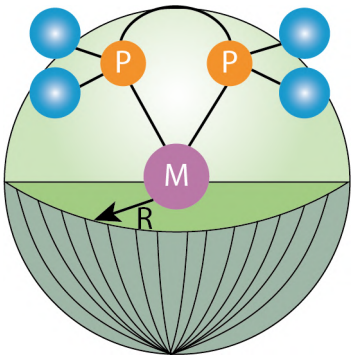


Figure 2.3: Illustration of the buried volume of a bidentate ligand for a sphere with radius R

2.3.3 Geometric descriptors

Geometric descriptors focus on describing the shape and size of molecular structures. This descriptor class entails (bond) angles and lengths. Two important angles for bisphosphine bidentate ligands are the bite and cone angles, of which illustrations are shown in Figure 2.4 The bite angle (Figure 2.4a) describes the angle between the central metal and the two donor atoms of a bidentate ligand. Correlations between the bite angle and catalyst selectivity have been found for various reactions [62]. The Tolman cone angle (Figure 2.4b) is one of the most common used measures of phosphine ligand size [63]. This cone angle covers the angle of a cylindrical cone, using the metal center as its core and spanning the edges of the van der Waals spheres of the outer ligand atoms [64].

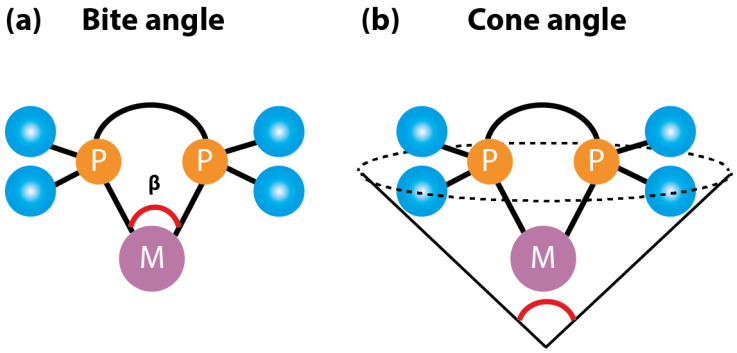


Figure 2.4: Illustrations of the bite angle (a) and cone angle (b) geometric descriptors

3

Results

In this chapter the obtained results will be discussed. The main findings are presented in the form of a paper. This paper includes applied methods and models, analysis and discussion of the main findings and research conclusions. The corresponding supporting information of the paper is found in the appendix.

The influence of ligand configurations on automated chemical space exploration of transition-metal complexes

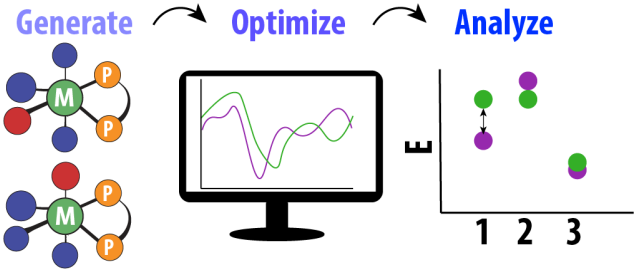
Niels J. van der Lem,[†,‡] Adarsh V. Kalikadien,[†,‡] and Evgeny A. Pidko[*,†]

†*Inorganic Systems Engineering, Department of Chemical Engineering, Faculty of Applied Sciences, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, The Netherlands.*

‡*Contributed equally to this work*

E-mail: e.a.pidko@tudelft.nl

Abstract



Computational chemical space exploration is becoming increasingly important for the design of organometallic homogeneous catalysts. Catalytic properties of metal-ligand complexes are not only influenced by the nature of ligands coordinated to the metal center, but also by their spatial arrangement around the metal center. However, computational catalyst screening workflows usually operate under the assumption of a particularly fixed ligand arrangement for broad classes of compounds, thus neglecting part of the chemical space. In the context of high-throughput *in silico* catalyst

screening, this study aims to evaluate the influence of ligand configurations on the stability and physical-chemical properties of transition metal complexes. We employed an automated workflow for the generation of complexes, sorting based on bite angles of ligands, DFT optimization and descriptor extraction. Ensembles of possible ligand configurations are generated for 88 bisphosphine bidentate ligands, alongside iridium(III), ruthenium(II) and manganese(I) metal centers. Based on DFT optimized geometries, analysis reveals a preference for a specific ligand configuration for iridium(III) complexes. However, this preference is not observed for ruthenium(II) and manganese(I) complexes. Furthermore, for the majority of ruthenium(II) and manganese(I) complexes multiple ligand configurations are found within a 10 kJ/mol range from the most favorable configuration. The analysis of thermodynamic, electronic, geometric and steric descriptors reveal that none of the descriptors consistently correlates to the preferred ligand configuration. These findings indicate that a priori selection of ligand configuration may result in insufficient coverage of the chemical space and insufficient representations of the key catalyst features for predictive *in-silico* catalyst screening campaigns.

Introduction

Homogeneous catalysis serves as the enabling technology for numerous organic chemical transformations.[1–3] An important and versatile class of homogeneous catalysts consists of organometallic complexes with tunable ligands.[4,5] Considering the nearly endless possible metal-ligand combinations, exploration of the chemical space is a challenge. Therefore, many methods for that help navigating this multidimensional space have been developed. Relevant examples are the approaches involving virtual high-throughput (HT) screening[6–10] and combinatorial chemistry.[8,11,12] Virtual screening of the metal-ligand space has been enabled by several (semi-)automated workflows utilizing data-driven quantitative structure activity/selectivity relationships (QSAR/QSSR).[13–17] These methods often consist of four fundamental components: structure generation, electronic structure calculation, descriptor extraction, and finally statistical modeling to relate descriptors to catalyst behavior.[18,19]

Common structure generation methods typically select only a single ligand arrangement based on either an energy ranking or heuristics for a selected chemical system. Considering the flexibility of these ligands, one thus assumes that the preferred ligand arrangement is retained for all members of a given ligand family and/or metal centers. Conformational search aiming at identifying low-energy rotamers and isomeric structures is commonly carried out for this selected coordination polyhedron with the pre-defined ligand configuration, which is preserved at this stage.[20,21]

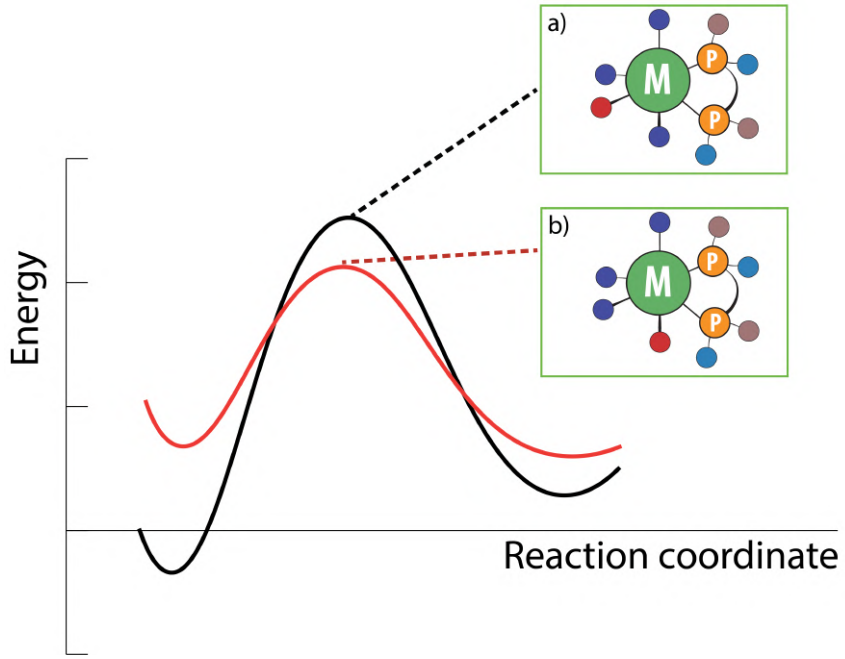


Figure 1: Potential reaction energy diagram of two different ligand configurations (a) and (b) of a TM-complex

However, different ligand configurations of TM-complexes vary in both stability and reactivity.[22] The relationship between the stability and observed catalytic properties of a complex is not known a priori. Consider a scenario where a meta-stable configuration of a TM complex, existing at a low concentration in the reactive system, establishes a favorable reaction channel. This minor catalytic component would provide a major impact on the reaction rate and even determines the nature and characteristics of the primary reaction product (Figure 1).[23–29] As an example, two possible ligand arrangements are depicted in Figure 1a and b. Although the complexes with a different ligand arrangement are in equilibrium, one configuration may provide a reaction path with a significantly higher energy barrier than the other configuration. The overall ensemble of ligand configurations ultimately contribute to the observed catalytic properties. In the context of the virtual HT screening, the conformational isomery of the organic ligand backbone has been recognized by the community.[19,30–32] However,the influence and contribution of the metastable configurations featuring varied

coordination environments and ligand arrangements is often overlooked due to the initial configuration choice. Furthermore, this choice assumes that the preferred configuration does not change with relatively minor variations in the ligand structure and, often, even the nature of the metal center. Consequently, the question arises whether catalytic systems can be fully accounted for when part of the chemical space is neglected due to initial human choices and intuition?

In this work, we investigated TM-complexes that are relevant to homogeneous hydrogenation catalysis, where bidentate ligands provide high enantioselectivity.[33–35] These complexes employed various bidentate ligands with iridium, ruthenium and manganese metal centers. Model auxiliary ligands were hydrides and CO (to yield neutral TM-complexes), while acetonitrile served as a model substrate, to ensure their minor impact on the overall conformational freedom of the complexes. We employed an automated workflow for construction, sorting and descriptor calculation of ensembles of ligand configurations for TM-complexes. The primary questions we would like to answer are whether a specific ligand configuration is abundantly more favorable, whether ligand configurations display notable differences in descriptors, and whether different ligand configurations are significantly different in terms of stability and chemical properties. To answer these questions, we constructed ensembles containing different ligand configurations for 88 bidentate ligands connected to 3 different metal centers, namely iridium(III), ruthenium(II) and manganese(I), giving us a total of 919 octahedral TM-complexes. Afterwards, DFT-optimized geometries and relevant descriptors of the different complex ensembles are compared and analyzed to explain differences between different ligand configurations.

Methods

Ligands and transition metal complexes

Herein, we have explored the configurational freedom and physical-chemical properties of an extended catalyst dataset featuring neutral $PPIrH_3(CH_3CN)$, $PPRuH_2(CO)CH_3CN$ and $PPMnH(CO)_2CH_3CN$ complexes. The dataset was constructed using 88 chiral bisphosphine (PP) ligands and without any a priory assumption of the preferred ligand arrangement or TM stereochemistry using a fully automated TM complex generation workflow.[36] Figure 2 a illustrates the studied ligand configurations for the iridium, ruthenium and manganese complexes. The 88 selected bidentate PP ligands belong to different ligand families, a subset of which is shown in Figure 2b. The complete set of bidentate ligands is available in the Supporting Information section A.

Complex generation and sorting workflow

The general workflow for complex structure generation employed in this study is visualized in Figure 3. Structures for TM-complexes were generated using the in-house built OBeLiX (Open Bidentate Ligand Explorer) workflow, available at `https://github.com/EPiCs-group/obelix`. This workflow aims to aid computational exploration of organometallic chemistry space through the automated generation, featurization and descriptor calculation. OBeLiX utilized the MACE python package for the automated generation of 3D structures and stereochemistry assessment of TM-complexes.[36,37] MACE is an open source python package, which allows bias-free generation of 3D TM-complexes starting from molecular SMILES strings[38] for ligands and metal centers. Furhthermore, MACE generates all possible stereoisomers, explores conformations and filters out identical and "impossible" configurations for the given metal-ligand combination.

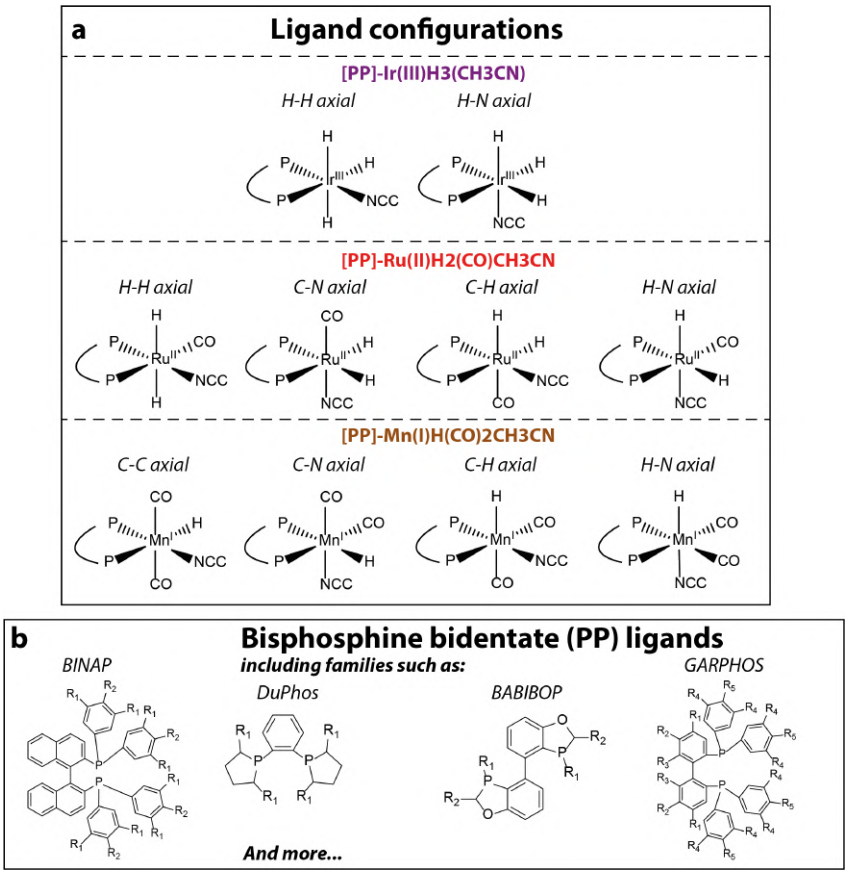


Figure 2: List of possible ligand configurations for each metal center (a) and a selection of representative selected biphosphine bidentate ligand families (b)

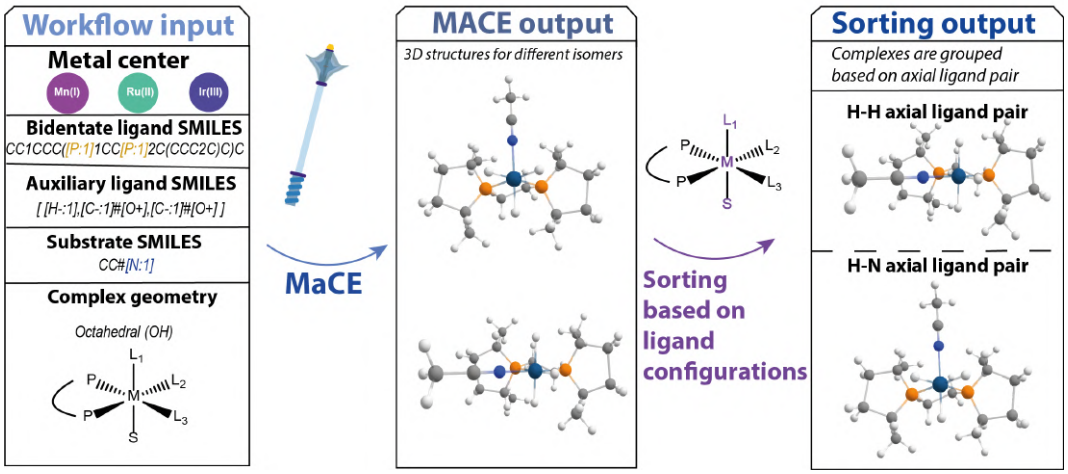


Figure 3: Workflow for generating & sorting TM-complex geometries from specified user input

After complex generation via MACE, the complexes were sorted according to their axial

ligand configuration, which were identified using bite angles. The coordinate system of TM-complexes are defined with respect to the bidentate ligand. For that reason, the bidentate ligand is always present in the equatorial position. Therefore, the ligands in the axial position are the only non-bidentate ligand containing pair forming a bite angle of 180°. After generating and sorting the TM-complexes, geometries were optimized at the level of Density Functional Theory (DFT).

Density Functional Theory Calculations

The geometries and energies of the generated geometries were further refined by DFT calculations using the Gaussian 16 C.02[39] software. For geometry optimizations, the PBE0[40] exchange-correlation functional was used with Grimme's DFT-D3(BJ) dispersion corrections[41] and the def2-SVPP basis set.[42] The selected combination of basis set and exchange correlation function are provide a good balance between computational cost and chemical accuracy.[43,44] Normal mode analysis was carried out to confirm that the optimized geometries correspond to local minima on the potential energy surface. For structures with present imaginary frequencies, the PyQRC python package[45,46] was used to remove these imaginary frequencies and restart geometry optimizations. After geometry optimization, energies were refined with single point (SP) calculations at the PBE0-D3 level using the def2-TZVPP basis set.[42,47]

Investigation of activity

To screen catalyst activity, one needs to define a measure of this emerging property. In this study, the substrate binding energy of acetonitrile (Equation 1) was used as the simplified metric for catalyst reactivity,[48] In 1, a substrate (S) binds to a metal complex containing a vacant coordination site, resulting in the substrate coordinated complex (L-M-S).

$$L\text{-}M\text{-} + S = L\text{-}M\text{-}S \tag{1}$$

To obtain substrate binding energies, an additional step to the workflow, outlined in Figure 3, was introduced. Following complex sorting, automated removal of the acetonitrile substrate from the complex structures was performed. Additional information about the substrate removal script is available in SI section B. The substrate binding energy is defined by the following formulae:

$$E_{bind} = E_{DFT,opt,complex} - (E_{DFT,opt,complex-nosubs} + E_{DFT,opt,subs}) \quad (2)$$

In this equation, substrate binding energy is described by the DFT optimized energy differences between the complex, ($E_{DFT,opt,complex}$) minus the sum of substrate-removed complex ($E_{DFT,opt,complex-nosubs}$) and the energy of non-bonded substrate ($E_{DFT,opt,subs}$).

Descriptor calculation

The OBeLiX descriptor calculator[19] was employed to automate the extraction of chemical-physical properties and descriptors of DFT-optimized complexes. This tool determines electronic, steric and geometric descriptors using Morfeus[49] and cclib.[50] A graph based method is employed to locate and label the bidentate donor atoms based on charges calculated by a xTB single-point calculation. Based on these charges, the bidentate donor atoms are labeled as either 'min' or 'max'. Additionally, the descriptor calculator enables descriptors to be obtained on different levels of geometry optimization. In this study, descriptors based on DFT optimized structures were considered. In total, a selection of 28 commonly used DFT-based descriptors were analyzed. A full list of descriptors, alongside their category and a description is available in SI section C.

Results and discussion

The results section is divided into three parts: preferences in ligand configurations, the influence of these configurations on physical-chemical properties and lastly the thermodynamic accessibility of metastable configurations.

Preferences in ligand configuration

To investigate whether a specific configuration is generally more favorable over others, the relative stability and substrate binding energy differences of possible complex configurations are analyzed. We studied whether a minimum on the potential energy surface corresponds to the same ligand configuration across a wide set of bisphosphine bidentate (PP) ligands.

The relative stability of alternative configurations with respect to the pre-assumed reference structure per metal center is presented in Figure 4. At the top of each figure, the reference structure is depicted, with the alternative configurations shown at the bottom. The H-N axial ligand pair structure is used as the reference, being the only common configuration present for all metals. The difference in stability between the reference and alternative configurations is denoted as ΔE_{ref}.

Figure 4a illustrates the ΔE_{ref} data for iridium(III) complexes. The general reference structure of the H-N axial ligand configuration is shown on top, while the structure of the alternative H-H axial ligand pair configuration is depicted at the bottom of the graph. A positive ΔE_{ref} is observed for the majority of bidentate ligands, indicating lower stability of the H-H axial ligand pair structure compared to the reference H-N axial arrangement.

Relative stability data for ruthenium complexes is presented in Figure 4b. The alternative configurations in this case include octahedral complexes featuring C-N, C-H and H-H atoms in the axial position. Compared to iridium, more alternative configurations are observed due to the different auxiliary ligand set. Assuming that hydrides are indistinguishable, iridium(III) complexes can only form two distinctive configurations, whereas four different

configurations of the ruthenium complex can be formed. A significant variety in the data set is observed, as the complex configuration lowest in energy often varies among the different bidentate ligands.

Similar to iridium(III), the H-H axial ligand configuration exhibits a positive ΔE_{ref} for the majority of bidentate ligands, signifying a lower stability compared to the reference H-N configuration. However, a notable difference from iridium(III) complexes is the common observation of a more stable alternative configuration for a many systems. For instance, the C-N axial configuration commonly exhibits a negative ΔE_{ref}, indicating their higher stability than the the H-N reference. Furthermore, our workflow identifies multiple Ru complexes with alternative configurations varying more than 50 kJ/mol compared to the reference case. These 'outliers' are the result of unfavorable conformations imposed by the specific ligand arrangement on the metal center.

Figure 4c visualizes ΔE_{ref} among generated manganese(I) complexes. Stability of the complexes featuring C-H, C-C and C-N axial ligand pair configurations are plotted relative to the stability of the H-N reference. Similar to ruthenium(II) complexes, 4 different ligand configurations are possible in this case. The most stable preferred configuration varies among different bidentate ligands, and no clear preference any of them could be identified for Mn complexes. Distinctive to manganese(I) complexes is the C-C configuration, for which a positive ΔE_{ref} is observed for the majority of the cases, indicating its overall lower stability. Nevertheless, a negative ΔE_{ref} is more frequently observed across all bidentate ligands as opposed to ruthenium(II) and iridium(III) complexes. This implies that the reference H-N axial ligand configuration does not show a thermodynamic preference among the considered manganese(I) complexes.

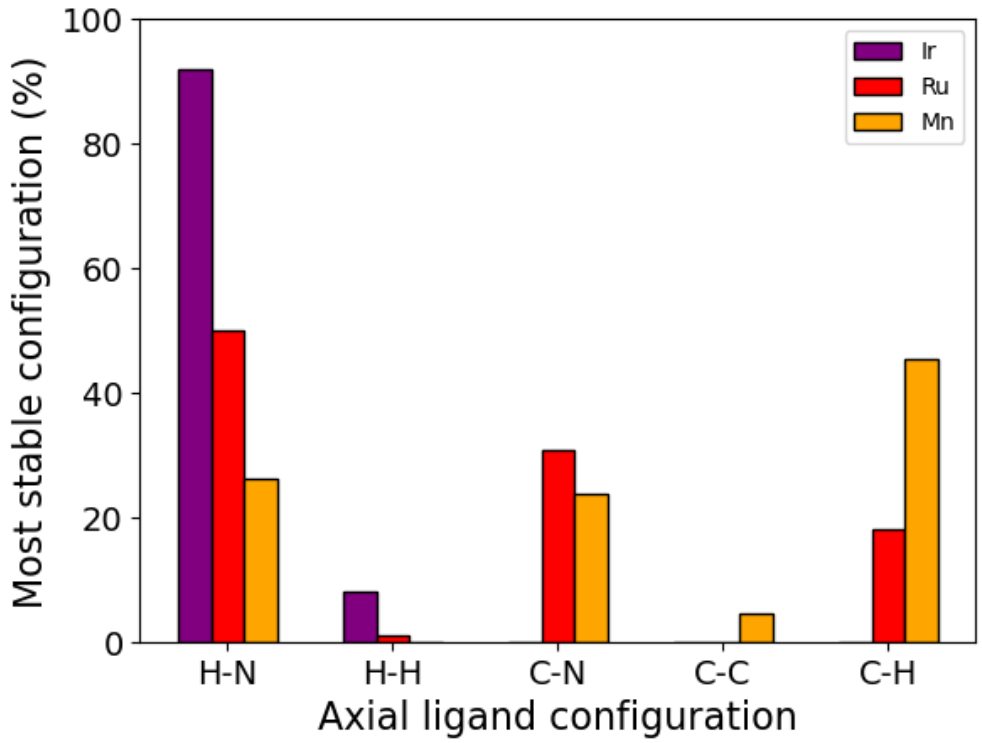


Figure 5: Distribution of most stable ligand configuration over all possible ligand configurations for manganese(I), ruthenium(II) and iridium(III) complexes, alongside 88 bisphosphine (PP) bidentate ligands

The statistical preference for a given ligand configuration in the analyzed dataset is presented in Figure 5. This figure illustrates the possible axial ligand configurations alongside the percentage of bidentate ligands for which those specific configurations are found to be most stable. The iridium(III) complexes clearly favor the H-N arrangement, which is the case for 92% of bidentate ligands. The remaining 8% favor the single alternative H-H configuration. For ruthenium(II) complexes, the H-N axial configuration is also frequently identified as the most stable one, but it now only accounts for 50% of the bidentate ligands. Both the C-N and C-H axial ligand arrangement emerge as the most stable ones for a notable number of bidentate ligands (31% and 18%, respectively). The H-H axial ligand configuration is the most stable one for only a single bidentate ligand. Unlike ruthenium(II) and iridium(III), related manganese(I) complexes do not display a pronounced preference for the H-N configuration. This geometry is preferred for only 26% of manganese(I) complexes, while the

alternative C-H configuration accounts for 45% of the bidentate ligands in this case. The C-N and C-C axial arrangement are preferred by 24% and 5% of the manganese(I) complexes.

Influence of configuration preferences on physical-chemical properties

The previously presented findings imply that ligand configurations are not equally favorable across the set of bidentate ligands and metals. This suggests that ligand configurations possess distinct properties resulting in stability differences. To quantify the physical-chemical differences between the possible complex ligand configurations, a selection of thermodynamic, electronic, steric and geometric descriptors were analyzed. A full list of calculated descriptors, corresponding category and definition are provided in SI section C. In total, 1 thermodynamic, 17 electronic, 7 geometric and 4 steric descriptors are considered in this study.

The analysis similar to that presented in Figure 4 and Figure 5 for stability, have been conducted for the thermodynamic descriptor, that is the acetonitrile substrate binding energy. Details and figures for this analysis are available in the SI, section D. In summary, similar trends are observed for the binding energy as those previously identified for stability. iridium(III) complexes display a clear preference in ligand configuration, whereas manganese(I) and ruthenium(II) complexes lack such a preference. However, our analysis shows that the configurations providing the highest binding energies, that may reflect their ability to activate the substrate, are not the same that show the highest stability. This implies that, in the context of HT virtual screening, the structure selection based on either highest relative stability or lowest substrate binding energy would yield different TM ligand configurations, leading to differences in predictions of catalytic behavior.

The HOMO, LUMO and HOMO-LUMO gap energies are widely used electronic descriptors, as they can be related to the complex reactivity and stability. Firstly, the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of

the different complex configurations are visualized to illustrate the differences in the electronic structure. The visualized orbitals can be found in the supporting information (see SI sectionE). Analysis reveals a similar LUMO shape among the different ligand configurations, primarily localized on the bidentate ligand backbone. However, substantial variations in the nature of HOMO could be observed for different ligand configurations. These electronic differences are further quantified in Figure 6, depicting the variations in energies of the HOMO, LUMO and HOMO-LUMO gap for different ligand configurations of ruthenium(II)-PP complexes. A color map is supplied signifying the relative stability of diverse ligand configurations compared to the most stable one, denoted as ΔE_{lowest}. Furthermore, the most stable ligand configuration is depicted in dark blue. The color of metastable ligand configurations progressively transitions to red the more positive ΔE_{lowest} becomes, indicating lower stability.

The HOMO energies of ligand configurations are presented in Figure 6a. The HOMO energy for a majority of bidentate ligands falls within a similar energy range, spanning from -5.0 to -5.8 eV. Notably, bidentate ligands 46 to 51 systematically display HOMO energies below -6.0 eV, signifying the influence of the specific bidentate ligand on the HOMO energy. Furthermore, The figure shows significant variance between different ligand configurations, as many of the bidentate ligands display differences of HOMO energies of at least 0.3 eV. Nonetheless, no distinctive pattern is observed for the most stable ligand configuration.

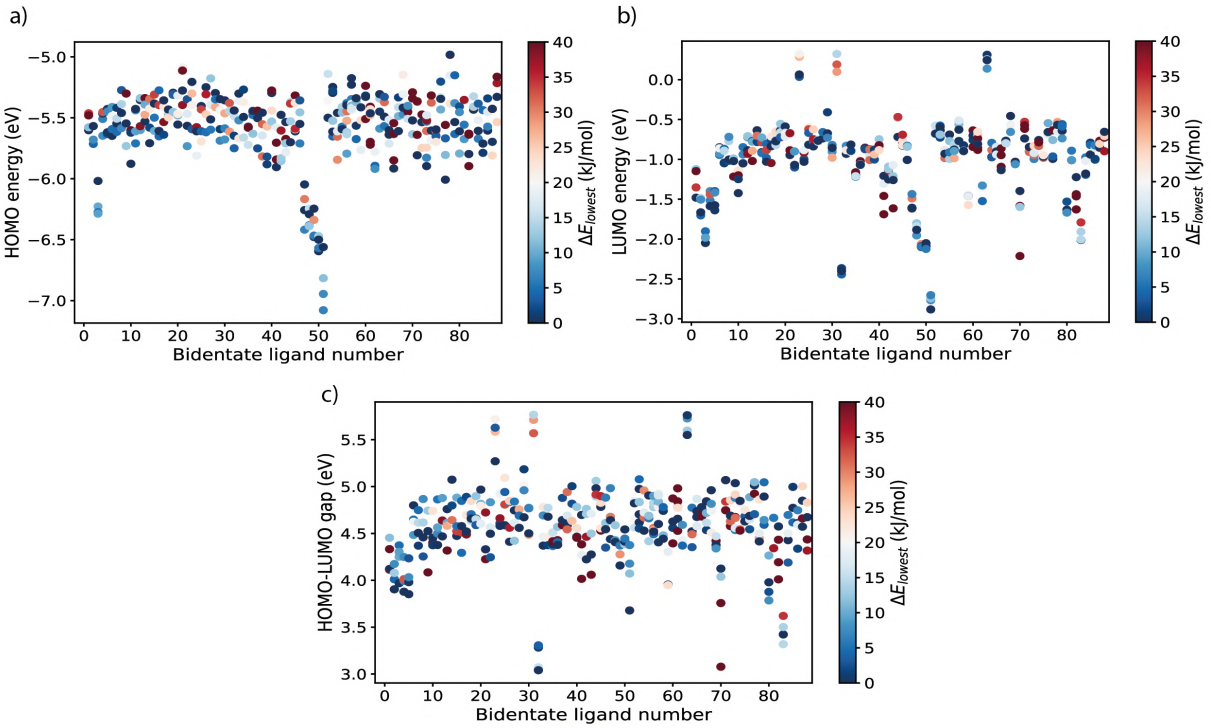


Figure 6: Energies for the HOMO (a), LUMO (b) and HOMO-LUMO gap (c) for ruthenium(II) complexes featuring different bidentate ligands. The color map represents the relative stability of different ligand configurations in comparison to the configuration most stable in energy (ΔE_{lowest}).

The LUMO energies are visualized in Figure 6b. Similar to HOMO energies, no distinctive pattern for the most stable ligand configuration is observed. Moreover, a comparable variance between ligand configurations, as found for HOMO energies, is evident in the LUMO energies. This suggests that both HOMO and LUMO orbitals may display differences in shape and size between ligand configurations. Furthermore, the variations in HOMO energies (Figure 6a) are not correlated with the changes in LUMO energies (Figure 6b) for different ligand configurations. As a result, no trend in HOMO-LUMO gap could be identified for the current dataset (Figure 6c). The majority of bidentate ligands display HOMO-LUMO gaps ranging between 4.0 and 5.0 eV. However, a small part of bidentate ligands shows notably higher or lower HOMO-LUMO gaps, caused by the nature of the specific bidentate ligands. The ligand configuration most favorable in energy displays no distinctive pattern, indicating that the HOMO, LUMO and HOMO-LUMO gap supply no significant insight into stability

preferences for a specific ligand configuration.

For iridium(III) and manganese(I) complexes, similar conclusion are drawn as for the previously shown ruthenium(II) complexes. Visualization of the HOMO, LUMO and the gap are provided in the supporting information (see section E), but they do not provide insight into stability differences among ligand configurations. Furthermore, all other electronic descriptors also show no convincing ability to differentiate between ligand configurations.

Nevertheless, one particular electronic descriptor does exhibit some ability to separate different ligand configurations, namely the NBO charge on the maximum bidentate ligand donor atom. The data of the NBO charge on the maximum bidentate ligand descriptor is depicted in Figure 7. In this figures, values for the NBO charge for the different ligand configurations are visualized, utilizing a similar color map as used in Figure 6. Figure 7a displays NBO charge values for manganese(I) complexes. In this graph, the NBO charge is generally similar among ligand configurations, whereas NBO charge for ruthenium(II) ligand configurations, displayed in Figure 7b, display a broader separation among configurations. No complete separation of ligand configurations is obtained. For iridium(III) complexes, displayed in Figure 7c, however, a significant separation of the most stable ligand configuration is observed. Generally, the NBO charge on the maximum bidentate ligand of the most stable configuration is 0.15 eV lower than the NBO charge of other configurations. Nevertheless, for a small part of bidentate ligands this relation does not hold true, resulting in no absolute separation of ligand configurations.

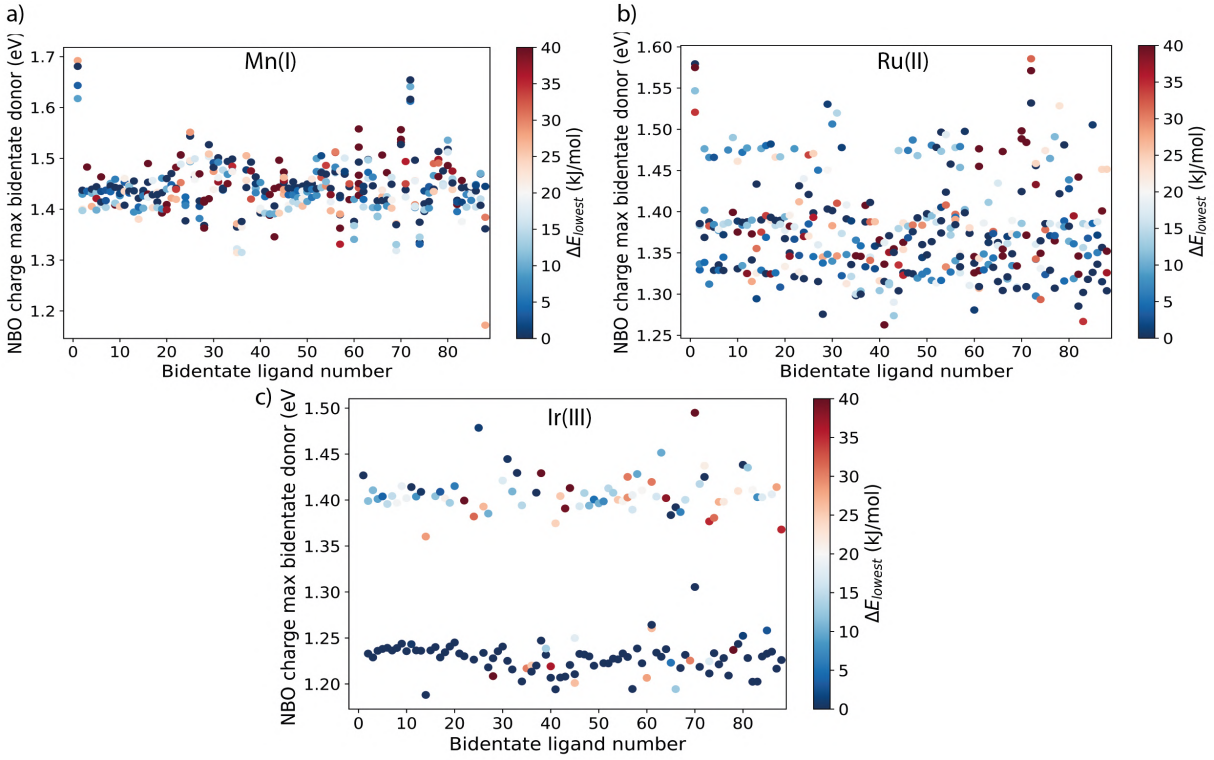


Figure 7: NBO charge of the maximum bidentate donor atoms for manganese(I) (a) ruthenium(II) (b) and iridium(III) (c) complexes for the set of bidentate ligands, alongside a color map indicating the relative stability of ligand configurations compared to the most stable ligand configuration, denoted as ΔE_{lowest}

Apart from electronic descriptors, steric and geometric descriptors are also considered (see SI section F). However, similar to most electronic descriptors, no distinctive differences between most stable complex ligand configuration and metastable configurations are observed.

Thermodynamic accessibility of metastable configurations

The distribution of stability among ligand configuration, depicted in Figure 4, further reveals that multiple isomers can exhibit similar stability. This finding suggests that under reaction conditions, multiple ligand configurations may contribute to the population of the coordination complex, thereby impacting the overall observed catalytic behavior. To assess this factor quantitatively, we have analyzed the proportion of systems for which multiple

ligand configuration were obtained within an energy threshold of 10 kJ/mol from the global minimum state. The choice of the 10 kJ/mol threshold is based on the assumption that a catalyst population follows a Boltzmann average, resulting in at least 5%, and up to 50% of the total population to be in a metastable state. For each complex ensemble, the number of isomers within a 10 kJ/mol range of the lowest energy isomer is obtained. The figures presenting the number of ligand configurations found within the energy threshold for each individual bidentate ligand are available in SI section G.

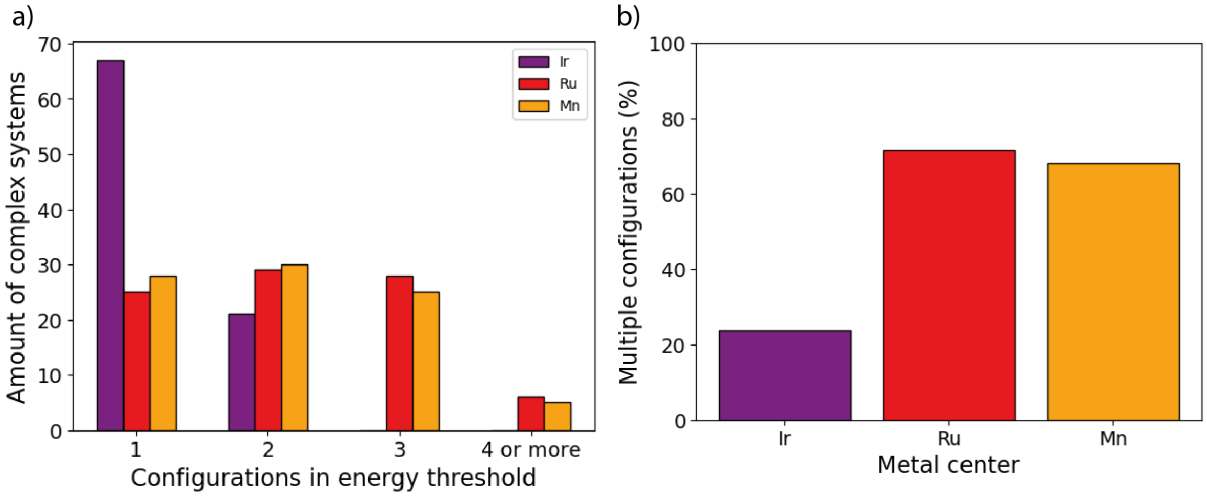


Figure 8: Number of ligand configurations within 10 kJ/mol of the most stable ligand configuration for the researched bidentate ligands (a), and the percentage of bidentate ligands for which multiple ligand configurations are found within the specified 10 kJ/mol energy range (b)

Figure 8a illustrates the number of ligand configurations within the 10 kJ/mol energy range from the most stable configuration for manganese(I), ruthenium(II) and iridium(III) complexes. The percentage of bidentate ligands containing multiple ligand configurations within the energy range, is shown in Figure 8b. The distribution of number of ligand configurations within 10 kJ/mol energy range over the bidentate ligands and metal centers can be found in the supporting information. For the majority of iridium(III) complexes, only a single configuration is observed within the specified energy range. This suggests significant differences in stability between ligand configurations. For 24% of the bidentate ligands,

multiple ligand configurations may coexist under reaction conditions, as energy differences are smaller than 10 kJ/mol. For a notable percentage of manganese(I) and ruthenium(II) complexes, multiple ligand configurations are found within the energy range, 72% and 68% respectively. This implies that the majority of ruthenium(II) and manganese(I) complexes contain multiple ligand configurations similar in stability.

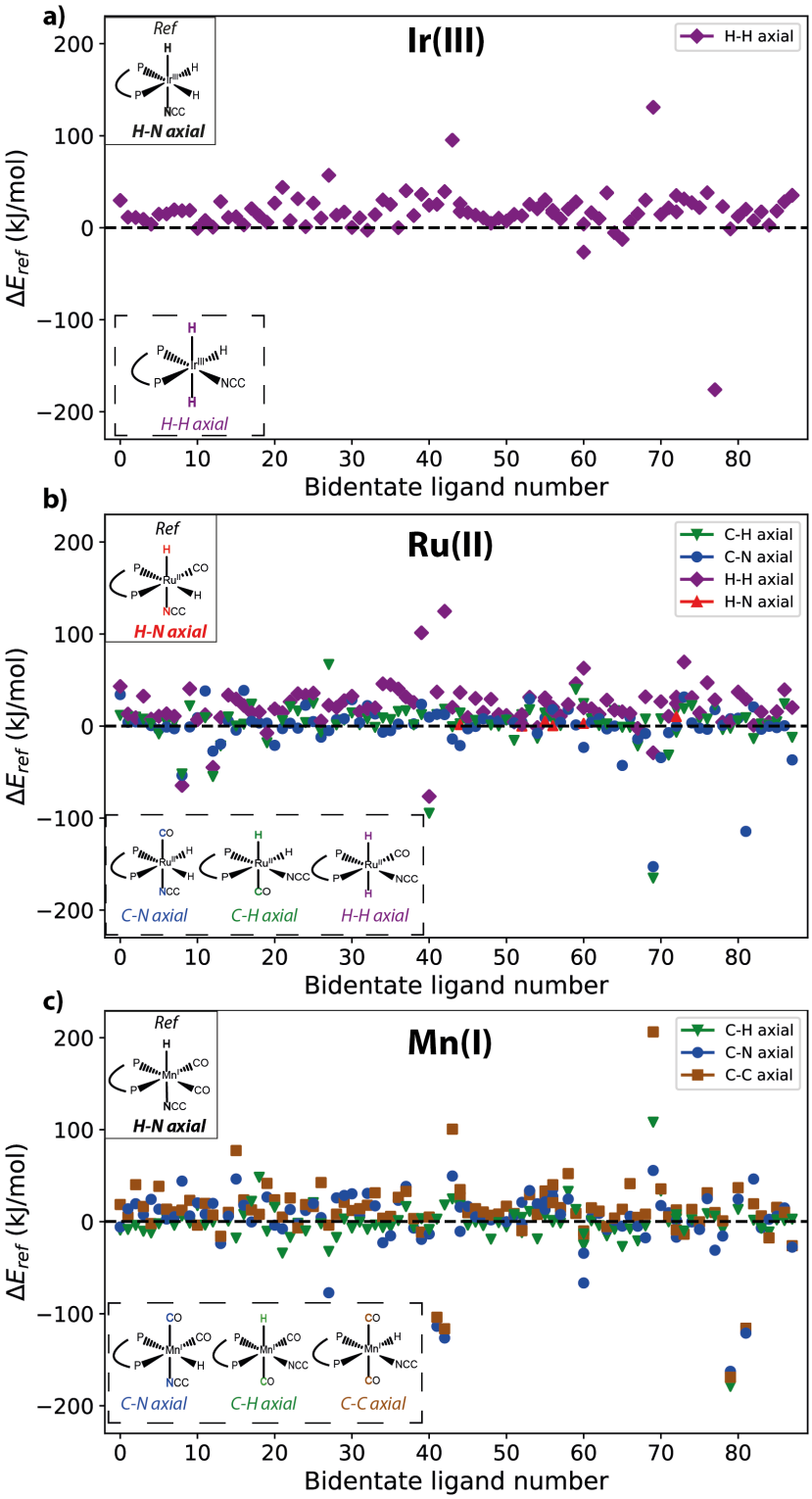


Figure 4: Relative stability of ligand configurations, shown at the bottom of a graph, and a reference structure, shown at the top of a graph, for set of bidentate ligands for iridium(III) (a), ruthenium (b) and manganese(I) (c) complexes

Conclusion

The aim of this study was to investigate differences in stability and physical chemical properties of possible TM-complex ligand configurations. This investigation was performed in three parts. Firstly, the preferences for certain ligand configurations in terms of stability was investigated. Secondly, we focused on explaining the effect of ligand configurations on descriptors of the metal-ligand complex. Finally, we quantified energy differences between ligand configurations, exploring the influence of metastable configurations. To do so, we developed an automated workflow for TM-complex structure generation and descriptor extraction.

For the study of stability based preferences in ligand configuration, our DFT-based findings revealed a clear preference in ligand configuration for iridium complexes, whereas manganese and ruthenium complexes lacked this preference. Nevertheless, the most frequently observed preference in ligand configurations differs between stability and substrate binding energies. This result implies that the most stable ligand configuration may not necessarily provide optimal catalytic performance. Quantifying the stability differences between ligand configurations proved challenging, as analysis of 28 descriptors revealed no descriptor consistently correlated with the most stable ligand configuration. Moreover, for a majority of manganese and ruthenium complexes multiple ligand configurations were found within a 10 kJ/mol energy range from the most favorable one, indicating that multiple ligand configurations may coexist under reaction conditions, both influencing catalyst properties. Our workflows and these findings bear significance in virtual high-throughput screening of catalysts, since a priori selection of a specific ligand configuration can result in insufficient coverage of the chemical space.

Supporting Information Available

All code, data and additional analysis are available in the supporting information. The OBeLiX workflow and the used implementation of MACE can be found on our Github page (`https://github.com/EPiCs-group`)

References

(1) Chiusoli, G. P.; Maitlis, P. M. *Metal-catalysis in industrial organic processes*; Royal Society of Chemistry, 2019.

(2) End, N.; Schöning, K.-U. Immobilized Catalysts in Industrial Research and Application. *Top Curr Chem* **2004**, 241–271.

(3) Duca, G. Homogeneous Catalysis with Metal Complexes in the Chemical Industry and Foodstuffs Chemistry. *Springer Series in Chem. Phys.* **2012**, 423–465.

(4) Reek, J. N.; de Bruin, B.; Pullen, S.; Mooibroek, T. J.; Kluwer, A. M.; Caumes, X. Transition Metal Catalysis Controlled by Hydrogen Bonding in the Second Coordination Sphere. *Chem. Rev.* **2022**, *122*, 12308–12369.

(5) Gensch, T.; Dos Passos Gomes, G.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A. A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *J. Am. Ch. Soc.* **2022**, *144*, 1205–1217.

(6) Senkan, S. M. High-throughput screening of solid-state catalyst libraries. *Nat.* **1998**, *394*, 350–353.

(7) Renom-Carrasco, M.; Lefort, L. Ligand libraries for high throughput screening of homogeneous catalysts. *Chem. Soc. Rev.* **2018**, *47*, 5038–5060.

(8) Liu, X.; Zou, P.; Song, L.; Zang, B.; Yao, B.; Xu, W.; Li, F.; Schroers, J.; Huo, J.; Wang, J.-Q. Combinatorial high-throughput methods for designing hydrogen evolution reaction catalysts. *ACS Catal.* **2022**, *12*, 3789–3796.

(9) Ahn, S.; Hong, M.; Sundararajan, M.; Ess, D. H.; Baik, M.-H. Design and optimization of catalysts based on mechanistic insights derived from quantum chemical reaction modeling. *Chem. Rev.* **2019**, *119*, 6509–6560.

(10) Cordova, M.; Wodrich, M. D.; Meyer, B.; Sawatlon, B.; Corminboeuf, C. C. Data-Driven Advancement of Homogeneous Nickel Catalyst Activity for Aryl Ether Cleavage. **2020**,

(11) Reetz, M. T. Combinatorial and Evolution-Based Methods in the Creation of Enantioselective Catalysts. *Angew. Chem. Int. Ed.* **2001**, *40*, 284–310.

(12) Jandeleit, B.; Schaefer, D. J.; Powers, T. S.; Turner, H. W.; Weinberg, W. H. Combinatorial materials science and catalysis. *Angew. Chem. Int. Ed.* **1999**, *38*, 2494–2532.

(13) Dotson, J. J.; van Dijk, L.; Timmerman, J. C.; Grosslight, S.; Walroth, R. C.; Gosselin, F.; Püntener, K.; Mack, K. A.; Sigman, M. S. Data-Driven Multi-Objective Optimization Tactics for Catalytic Asymmetric Reactions Using Bisphosphine Ligands. *J. Am. Chem. Soc.* **2023**, *145*, 110–121.

(14) Matsuoka, W.; Harabuchi, Y.; Maeda, S. Virtual Ligand-Assisted Screening Strategy to Discover Enabling Ligands for Transition Metal Catalysis. *ACS Catal.* **2022**, *12*, 3752–3766.

(15) Laplaza, R.; Gallarati, S.; Corminboeuf, C. Genetic Optimization of Homogeneous Catalysts. *Chem. Methods* **2022**, *2*, e202100107.

(16) Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting reaction performance in C-N cross-coupling using machine learning. *Science* **2018**, *360*, 186–190.

(17) Janet, J. P.; Chan, L.; Kulik, H. J. Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network. *J. Phys. Chem. Lett* **2018**, *9*, 57.

(18) Foscato, M.; Jensen, V. R. Automated in Silico Design of Homogeneous Catalysts. *ACS Catal.* **2020**,

(19) Kalikadien, A. V.; Mirza, A.; Hossaini, A. N.; Sreenithya, A.; Pidko, E. A. Paving the road towards automated homogeneous catalyst design. *ChemPlusChem* **2024**, *n/a*, e202300702.

(20) Rosales, A. R.; Wahlers, J.; Limé, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; Wiest, O.; Norrby, P.-O. Rapid virtual screening of enantioselective catalysts using CatVS. *Nat. Cat.* **2018**, *2*, 41–45.

(21) Pracht, P.; Bauer, C. A.; Grimme, S. Automated and efficient quantum chemical determination and energetic ranking of molecular protonation sites. *J. Comput. Chem.* **2017**, *38*, 2618–2631.

(22) Coe, B. J.; Glenwright, S. J. Trans-effects in octahedral transition metal complexes. *Coord. Chem. Rev.* **2000**, *203*, 5–80.

(23) Kumari, S.; Alexandrova, A. N.; Sautet, P. Nature of Zirconia on a Copper Inverse Catalyst Under CO2 Hydrogenation Conditions. *J. Am. Chem. Soc.* **2023**, *145*, 26350–26362.

(24) Zhang, Z.; Zandkarimi, B.; Alexandrova, A. N. Ensembles of Metastable States Govern Heterogeneous Catalysis on Dynamic Interfaces. *Acc. Chem. Res.* **2020**, *53*, 447–458.

(25) Brown, J. M.; Chaloner, P. A. Asymmetric homogeneous hydrogenation catalysed by rhodium complexes; the binding modes of enamides defined by 13C n.m.r. spectroscopy. *J. Chem. Soc., Chem. Commun.* **1979**, 613–615.

(26) Brown, J. M.; Chaloner, P. A. Structural characterisation of a transient intermediate in rhodium-catalysed asymmetric homogeneous hydrogenation. *J. Chem. Soc., Chem. Commun.* **1980**, 344.

(27) brown, J. M.; chaloner, P. A.; glaser, R.; geresh, S. Intermediates in asymmetric hydrogenation. *Tetrahedron* **1980**, *36*, 815–825.

(28) Chan, A. S.; Halpern, J. Interception and characterization of a hydridoalkylrhodium intermediate in a homogeneous catalytic hydrogenation reaction. *J. Am. Chem. Soc.* **1980**, *102*, 838–840.

(29) Gridnev, I. D.; Imamoto, T. Mechanism of enantioselection in Rh-catalyzed asymmetric hydrogenation. The origin of utmost catalytic performance. *Chem. Commun.* **2009**, 7447–7464.

(30) Brethomé, A. V.; Fletcher, S. P.; Paton, R. S. Conformational Effects on Physical-Organic Descriptors: The Case of Sterimol Steric Parameters. *ACS Catal.* **2019**, *9*, 2313–2323.

(31) Burai Patrascu, M.; Pottel, J.; Pinus, S.; Bezanson, M.; Norrby, P.-O.; Moitessier, N. From desktop to benchtop with automated computational workflows for computer-aided design in asymmetric catalysis. *Nat. Catal* **2020**, *3*, 574–584.

(32) Harden, I.; Neese, F.; Bistoni, G. An induced-fit model for asymmetric organocatalytic reactions: a case study of the activation of olefins via chiral Brønsted acid catalysts. *Chem. Sci.* **2022**, *13*, 8848–8859.

(33) Wen, J.; Wang, F.; Zhang, X. Asymmetric hydrogenation catalyzed by first-row transition metal complexes. *Chem. Soc. Rev.* **2021**, *50*, 3211–3237.

(34) Coetzee, J.; Dodds, D. L.; Klankermayer, J.; Brosinski, S.; Leitner, W.; Slawin, A. M.;

Cole-Hamilton, D. J. Homogeneous Catalytic Hydrogenation of Amides to Amines. *Chem. - Eur. J.* **2013**, *19*, 11039–11050.

(35) Lawrence, S. A. *Amines: Synthesis, properties and applications*; Cambridge University Press, 2006.

(36) Chernyshov, Y., I; Pidko, E. MACE: Automated Assessment of Stereochemistry of Transition Metal Complexes and its Applications in Computational CatalysisMACE: Automated Assessment of Stereochemistry of Transition Metal Complexes and its Applications in Computational Catalysis. *J. Chem. Theory Comput.* **2024, accepted**,

(37) Chernyshov, I. MACE: MetAl Complexes Embedding. `https://github.com/EPiCs-group/epic-mace`, 2020.

(38) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model* **1988**, *28*, 31–36.

(39) Frisch, M. J. et al. Gaussian˜16 Revision C.02. 2016; Gaussian Inc. Wallingford CT.

(40) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *Chem. Phys.* **1999**, *110*, 6158–6170.

(41) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *Chem. Phys.* **2019**, *150*, 154122.

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

(43) Sinha, V.; Laan, J. J.; Pidko, E. A. Accurate and rapid prediction of pKa of transition metal complexes: semiempirical quantum chemistry with a data-augmented approach. *Phys. Chem. Chem. Phys.* **2021**, *23*, 2557–2567.

(44) Kalikadien, A. V.; Pidko, E. A.; Sinha, V. ChemSpaX: exploration of chemical space by automated functionalization of molecular scaffold. *Digital Discovery* **2022**, *1*, 8–25.

(45) Silva, M. A.; Goodman, J. M. Aziridinium ring opening: a simple ionic reaction pathway with sequential transition states. *Tetrahedron Lett.* **2005**, *46*, 2067–2069.

(46) Goodman, J. M.; Silva, M. A. QRC: a rapid method for connecting transition structures to reactants in the computational analysis of organic reactivity. *Tetrahedron Lett.* **2003**, *44*, 8233–8236.

(47) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.

(48) Meyer, B.; Sawatlon, B.; Heinen, S.; von Lilienfeld, O. A.; Corminboeuf, C. Machine learning meets volcano plots: computational discovery of cross-coupling catalysts. *Chem. Sci.* **2018**, *9*, 7069–7077.

(49) Jorner, K. Morfeus: molecular features for machine learning. `https://github.com/digital-chemistry-laboratory/morfeus?tab=readme-ov-file`, 2022.

(50) O'Boyle, N. M.; Tenderholt, A. L.; Langner, K. M. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845.

4

Summary & Outlook

4.1 Summary

HT Workflows for computational catalyst screening have the potential to discover novel earth abundant TM-metal complexes. However, these workflows commonly work with an a priori assumed ligand configuration. Doing so, the potential influence of metastable ligand configurations is neglected for chemical space exploration. The aim of this project was to investigate differences and influence of ligand configurations of TM-complexes on catalyst systems. This was done in 3 parts. First, the preference for a certain ligand configuration in terms of stability was investigated. Secondly, differences in physical-chemical properties between ligand configurations were studied. At last, the effect of metastable configurations on the was investigated.

In terms of stability, iridium complexes displayed a clear preference for the H-N axial ligand configuration. For ruthenium, this was also the most preferred ligand configuration, although only for half of the bidentate ligands. This ligand configuration was not favored for manganese complexes and no other ligand configuration was preferred for majority of manganese complexes. For each metal center, another ligand configuration was preferred in terms of activity compared to stability, indicating that geometry selection based on stability or activity would yield different complex structures.

Explaining the preference for certain ligand configurations proved challenging. Analysis of the HOMO, LUMO and HOMO-LUMO gap revealed no trend to consistently identify the most stable ligand configuration. Expanding descriptor analysis to 25 other electronic, steric and geometric descriptors revealed that none of the descriptors exhibit a trend for identifying the most stable ligand configuration.

To study the influence of metastable ligand configurations on the catalytic systems, the number of ligand configurations within a 10 kJ/mol range from the most stable configuration were obtained. This analysis revealed that for most iridium complexes, the stability of ligand configurations are significantly different. However, for ruthenium and manganese complexes, multiple ligand configurations are observed within the specified 10 kJ/mol threshold for a majority of complexes. This implies that various ligand configurations of TM-complexes may coexist under reaction conditions.

These findings indicate that a priori selection of ligand configuration may result in insufficient coverage of the chemical space and insufficient representations of the key catalyst features for predictive *in-silico* catalyst screening campaigns.

4.2 Outlook

The conclusion of this research states that metastable ligand configurations should be considered in chemical space exploration, as these configurations could have major impact on catalyst behavior. Recent research in the field shows that computational predictions do not correlate well with experimental results. The neglect of metastable ligand configurations during chemical space exploration may contribute to these inaccuracies. In order to improve catalyst predictions. an additional step is added in the typical workflow, shown in Figure 1.1, resulting in the proposed workflow depicted in Figure 4.1. In this proposed workflow, the additional step is added after geometry optimization. During this step, relative stability data of ligand configurations is evaluated, and a distribution of ligand configurations in the catalyst system is estimated based on a Boltzmann distribution. In this approach, the influence of metastable ligand configurations is considered in descriptor calculations, which could potentially result in improved accuracy of predictions of catalytic behavior.

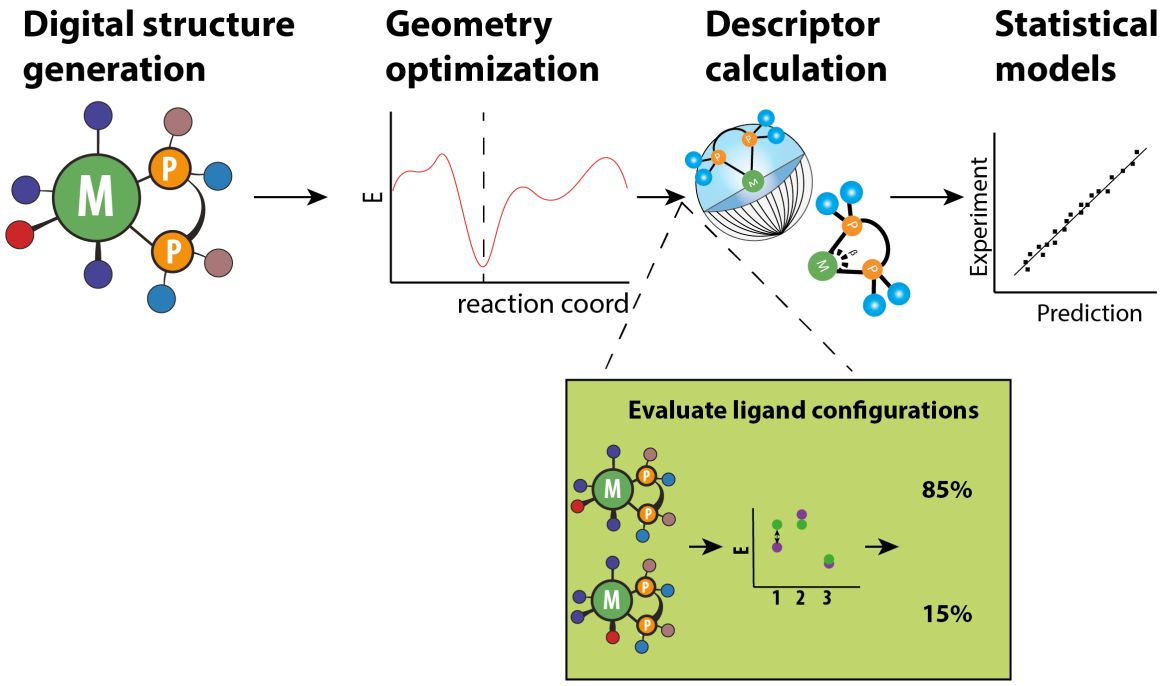


Figure 4.1: Proposed workflow, considering ligand configurations. Descriptors are determined based on the Boltzmann average ligand configuration distribution.

Throughout this study, geometries of all ligand configurations are DFT optimized, resulting in a large number of performed DFT calculations. In order to perform to the proposed workflow depicted in Figure 4.1, computational resources have to be efficiently utilized, by optimizing DFT efficiency. By employing a lower order exchange correlation functional, such as PBE, for initial geometry optimization, coupled with a PBE0 single-point calculation may result in a more resource efficient approach, while retaining similar accuracy to full PBE0 optimization.

Acknowledgements

The last 5 months have been an amazing journey. If someone would have told me that I would be writing a paper during my thesis, I would have found it hard to believe. This experience was truly wonderful and I am grateful to several individuals for their support and guidance throughout the project.

First, I would like to thank my supervisor Professor Evgeny Pidko for providing me with the opportunity to contribute to this research project. Thanks to your guidance I was able to greatly improve my research skills. Furthermore, I am extremely grateful for your help with finding an internship.

I want to thank my daily supervisor Adarsh Kalikadien for providing guidance during the project. Your involvement in the project and willingness to offer assistance have massively contributed to the great experience I had these past few months. Furthermore, I am thankful to get to experience different scientific events. Attending your (poster) presentations on your research was truly inspiring, and I am certain you will achieve first place in the future. It has been a pleasure working with you and I wish you all the best with the remaining stages of your PhD.

A special thanks to all other students and postdocs whom I got to know in the ISE group. Thanks for providing feedback and engaging in discussions. I really appreciate all of you.

Finally, I would like to thank my friends and family. I appreciate the time I shared with friends playing cards during the lunch break. This was definitely a highlight of my everyday thesis life. Last, but definitely not least I would like to thank my family for always supporting me and believing in me.

- Niels van der Lem
Delft, February 2024

Bibliography

[1] Zhen Ma and Francisco Zaera. "Heterogeneous Catalysis by Metals". In: *Encyclopedia of Inorganic and Bioinorganic Chemistry*. John Wiley Sons, Ltd, 2014, pp. 1–16. ISBN: 9781119951438. DOI: https://doi.org/10.1002/9781119951438.eibc0079.pub2.

[2] Christopher Blum, Dirk Bunke, Maximilian Hungsberg, Elsbeth Roelofs, Anke Joas, Reinhard Joas, Markus Blepp, and Hans-Christian Stolzenberg. "The concept of sustainable chemistry: Key drivers for the transition towards sustainable development". In: *Sustain. Chem. Pharm.* 5 (2017), pp. 94–104. ISSN: 2352-5541. DOI: https://doi.org/10.1016/j.scp.2017.01.001.

[3] R.A Sheldon and R.S Downing. "Heterogeneous catalytic transformations for environmentally friendly production". In: *App. Catal.* 189.2 (1999), pp. 163–183. ISSN: 0926-860X. DOI: https://doi.org/10.1016/S0926-860X(99)00274-4.

[4] Yogesh G. Shelke, Afsana Yashmeen, Aniket V. A. Gholap, Santosh J. Gharpure, and Anant R. Kapdi. "Homogeneous Catalysis: A Powerful Technology for the Modification of Important Biomolecules". In: *Chem. Asian J.* 13.20 (2018), pp. 2991–3013. DOI: https://doi.org/10.1002/asia.201801020.

[5] Indrajit Ghosh, Nikita Shlapakov, Tobias A Karl, Jonas Düker, Maksim Nikitin, Julia V Burykina, Valentine P Ananikov, and Burkhard König. "General cross-coupling reactions with adaptive dynamic homogeneous catalysis". In: *Nat.* 619 (2023). DOI: 10.1038/s41586-023-06087-4. URL: https://doi.org/10.1038/s41586-023-06087-4.

[6] Lucero González-Sebastián, Adan Reyes-Sanchez, and David Morales-Morales. "Hydrogenation and Cross-Coupling Reactions Catalyzed by Mn, Fe, and Co Aromatic Pincer Complexes". In: *Organometallics* 42.18 (2023), pp. 2426–2446. DOI: 10.1021/acs.organomet.3c00261.

[7] Elizabeth L. Bell, William Finnigan, Scott P. France, Anthony P. Green, Martin A. Hayes, Lorna J. Hepworth, Sarah L. Lovelock, Haruka Niikura, Sílvia Osuna, Elvira Romero, Katherine S. Ryan, Nicholas J. Turner, and Sabine L. Flitsch. "Biocatalysis". In: *Nat. Rev. Methods. Primers.* 1 (1 Dec. 2021). ISSN: 26628449. DOI: 10.1038/s43586-021-00044-z.

[8] Matthew R. Elsby and R. Tom Baker. "Strategies and mechanisms of metal–ligand cooperativity in first-row transition metal complex catalysts". In: *Chem. Soc. Rev.* 49 (24 2020), pp. 8933–8987. DOI: 10.1039/D0CS00509F. URL: http://dx.doi.org/10.1039/D0CS00509F.

[9] Andrew L. Clevenger, Ryan M. Stolley, Justis Aderibigbe, and Janis Louie. "Trends in the Usage of Bidentate Phosphines as Ligands in Nickel Catalysis". In: *Chem. Rev.* 120.13 (2020), pp. 6124–6196. DOI: 10.1021/acs.chemrev.9b00682.

[10] James Keeler and Peter Wothers. "Chapter 15: Transition metals". In: *Chemical structure and reactivity: An integrated approach*. Oxford University Press, 2014.

[11] Cheng Hou, Yinwu Li, and Zhuofeng Ke. "Transition metal center effect on the mechanism of homogenous hydrogenation and dehydrogenation". In: *Inorganica Chim. Acta* 511 (Oct. 2020), p. 119808. DOI: https://doi.org/10.1016/j.ica.2020.119808.

[12] James Pritchard, Georgy A. Filonenko, Robbert van Putten, Emiel J. M. Hensen, and Evgeny A. Pidko. "Heterogeneous and homogeneous catalysis for the hydrogenation of carboxylic acid derivatives: history, advances and future directions". In: *Chem. Soc. Rev.* 44 (11 2015), pp. 3808–3833. DOI: 10.1039/C5CS00038F.

[13] Amit Kumar, Prosenjit Daw, and David Milstein. "Homogeneous Catalysis for Sustainable Energy: Hydrogen and Methanol Economies, Fuels from Biomass, and Related Topics". In: *Chem. Rev.* 122.1 (2022), pp. 385–441. DOI: 10.1021/acs.chemrev.1c00412.

[14] Stephen A. Lawrence. *Amines: Synthesis, properties and applications*. Cambridge University Press, 2006.

[15] Weicheng Zhang, Yongxiang Chi, and Xumu Zhang. "Developing Chiral Ligands for Asymmetric Hydrogenation". In: *Acc. Chem. Res.* 40.12 (2007), pp. 1278–1290. DOI: 10.1021/ar7000028.

[16] Dattatraya B Bagal and Bhalchandra M Bhanage. "Recent Advances in Transition Metal-Catalyzed Hydrogenation of Nitriles". In: *ASC Catal.* 357 (5 2015), pp. 883–900. URL: https://onlinelibrary.wiley.com/doi/10.1002/adsc.201400940.

[17] Andrew K. King, Aneelman Brar, Guigen Li, and Michael Findlater. "Homogeneous and Recyclable Palladium Catalysts: Application in Suzuki–Miyaura Cross-Coupling Reactions". In: *Organometallics* 42.17 (2023), pp. 2353–2358. DOI: 10.1021/acs.organomet.3c00231.

[18] Yujie Wang, Mingyang Wang, Yibiao Li, and Qiang Liu. "Homogeneous manganese-catalyzed hydrogenation and dehydrogenation reactions". In: *Chem* 7.5 (May 2021), pp. 1180–1223. ISSN: 24519294.

[19] Kuhali Das, Satyadeep Waiba, Akash Jana, and Biplab Maji. "Manganese-catalyzed hydrogenation, dehydrogenation, and hydroelementation reactions". In: *Chem. Soc. Rev.* 51 (11 June 2022), pp. 4386–4464. ISSN: 1460-4744. URL: https://pubs.rsc.org/en/content/articlehtml/2022/cs/d2cs00093h%20https://pubs.rsc.org/en/content/articlelanding/2022/cs/d2cs00093h.

[20] Tobias Gensch, Gabriel Dos Passos Gomes, Pascal Friederich, Ellyn Peters, Théophile Gaudin, Robert Pollice, Kjell Jorner, Akshatkumar Nigam, Michael Lindner-D'Addario, Matthew S. Sigman, and Alán Aspuru-Guzik. "A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis". In: *J. Am. Chem. Soc.* 144 (3 Jan. 2022), pp. 1205–1217. ISSN: 15205126. DOI: 10.1021/jacs.1c09718.

[21] Steven M. Mennen, Carolina Alhambra, C. Liana Allen, Mario Barberis, Simon Berritt, Thomas A. Brandt, Andrew D. Campbell, Jesús Castañón, Alan H. Cherney, Melodie Christensen, David B. Damon, J. Eugenio de Diego, Susana García-Cerrada, Pablo García-Losada, Rubén Haro, Jacob Janey, David C. Leitch, Ling Li, Fangfang Liu, Paul C. Lobben, David W. C. MacMillan, Javier Magano, Emma McInturff, Sebastien Monfette, Ronald J. Post, Danielle Schultz, Barbara J. Sitter, Jason M. Stevens, Iulia I. Strambeanu, Jack Twilton, Ke Wang, and Matthew A. Zajac. "The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future". In: *Org Process Res Dev.* 23.6 (2019), pp. 1213–1242. DOI: 10.1021/acs.oprd.9b00140.

[22] Ruben Laplaza, Simone Gallarati, and Clemence Corminboeuf. "Genetic Optimization of Homogeneous Catalysts". In: *Chem. Methods* 2 (6 June 2022), e202100107. ISSN: 2628-9725.

[23] Jordan J. Dotson, Lucy van Dijk, Jacob C. Timmerman, Samantha Grosslight, Richard C. Walroth, Francis Gosselin, Kurt Püntener, Kyle A. Mack, and Matthew S. Sigman. "Data-Driven Multi-Objective Optimization Tactics for Catalytic Asymmetric Reactions Using Bisphosphine Ligands". In: *J. Am. Chem. Soc.* 145 (1 Jan. 2023), pp. 110–121. ISSN: 15205126. URL: https://pubs.acs.org/doi/full/10.1021/jacs.2c08513.

[24] Wataru Matsuoka, Yu Harabuchi, and Satoshi Maeda. "Virtual Ligand-Assisted Screening Strategy to Discover Enabling Ligands for Transition Metal Catalysis". In: *ACS Cat.* 12 (7 Apr. 2022), pp. 3752–3766. ISSN: 21555435. URL: https://pubs.acs.org/doi/full/10.1021/acscatal.2c00267.

[25] Marco Foscato and Vidar R Jensen. "Automated in Silico Design of Homogeneous Catalysts". In: *ACS Cat.* (2020). DOI: 10.1021/acscatal.9b04952. URL: https://dx.doi.org/10.1021/acscatal.9b04952.

[26] Derek T. Ahneman, Jesús G. Estrada, Shishi Lin, Spencer D. Dreher, and Abigail G. Doyle. "Predicting reaction performance in C–N cross-coupling using machine learning". In: *Science* 360 (6385 Apr. 2018), pp. 186–190. ISSN: 0036-8075. DOI: 10.1126/science.aar5169.

[27] Fred Basolo and Ralph G. Pearson. "The trans effect in metal complexes". In: *Prog. Inorg. Chem.* (1962), pp. 381–453. DOI: 10.1002/9780470166055.ch6.

[28] Benjamin J Coe and Susan J Glenwright. "Trans-effects in octahedral transition metal complexes". In: *Coord. Chem. Rev.* 203.1 (2000), pp. 5–80. DOI: https://doi.org/10.1016/S0010-8545(99)00184-8.

[29] Simran Kumari, Anastassia N. Alexandrova, and Philippe Sautet. "Nature of Zirconia on a Copper Inverse Catalyst Under CO2 Hydrogenation Conditions". In: *J. Am. Chem. Soc.* 145.48 (2023), pp. 26350–26362. DOI: 10.1021/jacs.3c09947.

[30] Zisheng Zhang, Borna Zandkarimi, and Anastassia N. Alexandrova. "Ensembles of Metastable States Govern Heterogeneous Catalysis on Dynamic Interfaces". In: *Acc. Chem. Res.* 53.2 (2020), pp. 447–458. DOI: 10.1021/acs.accounts.9b00531.

[31] John M. Brown and Penny A. Chaloner. "Asymmetric homogeneous hydrogenation catalysed by rhodium complexes; the binding modes of enamides defined by 13C n.m.r. spectroscopy". In: *J. Chem. Soc., Chem. Commun.* (14 1979), pp. 613–615. DOI: 10.1039/C39790000613.

[32] John M. Brown and Penny A. Chaloner. "Structural characterisation of a transient intermediate in rhodium-catalysed asymmetric homogeneous hydrogenation". In: *J. Chem. Soc., Chem. Commun.* 8 (1980), p. 344. DOI: 10.1039/c39800000344.

[33] John M. brown, Penny A. chaloner, Robert glaser, and Shimona geresh. "Intermediates in asymmetric hydrogenation". In: *Tetrahedron* 36.6 (Jan. 1980), pp. 815–825. DOI: 10.1016/s0040-4020(01)93700-7.

[34] A. S. Chan and Jack Halpern. "Interception and characterization of a hydridoalkylrhodium intermediate in a homogeneous catalytic hydrogenation reaction". In: *J. Am. Chem. Soc.* 102.2 (Jan. 1980), pp. 838–840. DOI: 10.1021/ja00522a067.

[35] Ilya D. Gridnev and Tsuneo Imamoto. "Mechanism of enantioselection in Rh-catalyzed asymmetric hydrogenation. The origin of utmost catalytic performance". In: *Chem. Commun.* (48 2009), pp. 7447–7464. DOI: 10.1039/B912440C.

[36] Errol G. Lewars. "Density Functional Calculations". In: *Computational chemistry: Introduction to the theory and applications of molecular and Quantum Mechanics*. Springer, 2016.

[37] Konstantinos D. Vogiatzis, Mikhail V. Polynski, Justin K. Kirkland, Jacob Townsend, Ali Hashemi, Chong Liu, and Evgeny A. Pidko. "Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities". In: *Chemical Reviews* 119 (4 Feb. 2019), pp. 2453–2523. ISSN: 15206890. DOI: 10.1021/acs.chemrev.8b00361.

[38] Yu-hong Lam, Matthew N. Grayson, Mareike C. Holland, Adam Simon, and K. N. Houk. "Theory and Modeling of Asymmetric Catalytic Reactions". In: *Acc. Chem. Res.* 49.4 (2016), pp. 750–762. DOI: 10.1021/acs.accounts.6b00006.

[39] Christopher J. Cramer and Donald G. Truhlar. "Density functional theory for transition metals and transition metal chemistry". In: *Phys. Chem. Chem. Phys.* 11 (46 2009), pp. 10757–10816. DOI: 10.1039/B907148B. URL: http://dx.doi.org/10.1039/B907148B.

[40] Theresa Sperger, Italo A. Sanhueza, Indrek Kalvet, and Franziska Schoenebeck. "Computational Studies of Synthetically Relevant Homogeneous Organometallic Catalysis Involving Ni, Pd, Ir, and Rh: An Overview of Commonly Employed DFT Methods and Mechanistic Insights". In: *Chem. Rev.* 115.17 (2015), pp. 9532–9586. DOI: 10.1021/acs.chemrev.5b00163.

[41] Wolfram Koch and Max C. Holthausen. *A chemist's guide to density functional theory*. Wiley-VCH, 2001.

[42] Atilla Szabo and Neil S Ostlund. *Modern Quantum Chemistry - introduction to advanced electronic structure theory*. Dover Publications Inc., 1996.

[43] L. S. Cederbaum. "Born–oppenheimer approximation and beyond for time-dependent electronic processes". In: *J. Chem. Phys.* 128.12 (2008). DOI: 10.1063/1.2895043.

[44] Yutaka Shikano, Hiroshi C Watanabe, Ken M Nakanishi, and Yu-Ya Ohnishi. "Post-Hartree-Fock method in quantum chemistry for quantum computer". In: *Eur. Phys. J. Spec. Top* 230 (Apr. 2021), pp. 1037–1051. DOI: 10.1140/epjs/s11734-021-00087-z. URL: https://doi.org/10.1140/epjs/s11734-021-00087-z.

[45] Pragya Verma and Donald G. Truhlar. "Status and Challenges of Density Functional Theory". In: *Trends Chem.* 2.4 (2020). Special Issue - Laying Groundwork for the Future, pp. 302–318. ISSN: 2589-5974. DOI: https://doi.org/10.1016/j.trechm.2020.02.005. URL: https://www.sciencedirect.com/science/article/pii/S2589597420300411.

[46] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871. DOI: 10.1103/PhysRev.136.B864. URL: https://link.aps.org/doi/10.1103/PhysRev.136.B864.

[47] Frank Neese. "Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling". In: *Coord. Chem. Rev.* 253.5 (2009). Theory and Computing in Contemporary Coordination Chemistry, pp. 526–563. ISSN: 0010-8545. DOI: https://doi.org/10.1016/j.ccr.2008.05.014. URL: https://www.sciencedirect.com/science/article/pii/S0010854508001197.

[48] John P. Perdew and Karla Schmidt. "Jacob's ladder of density functional approximations for the exchange-correlation energy". In: *AIP Conf. Proc.* 577.1 (July 2001), pp. 1–20. ISSN: 0094-243X. DOI: 10.1063/1.1390175. eprint: https://pubs.aip.org/aip/acp/article-pdf/577/1/1/12108089/1_1_online.pdf. URL: https://doi.org/10.1063/1.1390175.

[49] Carlo Adamo and Vincenzo Barone. "Toward reliable density functional methods without adjustable parameters: The PBE0 model". In: *Journal of Chemical Physics* 110 (13 Apr. 1999), pp. 6158–6170. ISSN: 00219606. DOI: 10.1063/1.478522.

[50] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Physical Review Letters* 77 (18 Oct. 1996), pp. 3865–3868. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.77.3865.

[51] Jingjing Zheng, Xuefei Xu, and Donald G Truhlar. "Minimally augmented Karlsruhe basis sets". In: *Theor. Chem. Acc.* 128 (2011), pp. 295–305. DOI: 10.1007/s00214-010-0846-z.

[52] 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". In: *Phys. Chem. Chem. Phys.* 7 (18 2005), p. 3297. ISSN: 1463-9076. DOI: 10.1039/b508541a.

[53] Stefan Grimme. "Density functional theory with London dispersion corrections". In: *Ltd. WIREs Comput Mol Sci* 1 (2011), pp. 211–228. DOI: 10.1002/wcms.30.

[54] 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". In: *J. Chem. Phys.* 132 (15 Apr. 2010), p. 154104. ISSN: 0021-9606. DOI: 10.1063/1.3382344.

[55] Francesco Fracchia, Gianluca Del Frate, Giordano Mancini, Walter Rocchia, and Vincenzo Barone. "Force Field Parametrization of Metal Ions from Statistical Learning Techniques". In: *J. Chem. Theory Comput.* 14.1 (2018), pp. 255–273. DOI: 10.1021/acs.jctc.7b00779.

[56] A. K. Rappe, K. S. Colwell, and C. J. Casewit. "Application of a universal force field to metal complexes". In: *Inorg. Chem.* 32.16 (Aug. 1993), pp. 3438–3450. DOI: 10.1021/ic00068a012.

[57] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. III Goddard, and W. M. Skiff. "UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations". In: *J. Am. Chem. Soc.* 114.25 (1992), pp. 10024–10035. DOI: 10.1021/ja00051a040.

[58] Ramakrishnan Parthasarathi and Alok Dhawan. "Chapter 5 - In Silico Approaches for Predictive Toxicology". In: *In Vitro Toxicology*. Ed. by Alok Dhawan and Seok Kwon. Academic Press, 2018, pp. 91–109. ISBN: 978-0-12-804667-8. DOI: https://doi.org/10.1016/B978-0-12-804667-8.00005-5.

[59] Marzieh Miar, Abolfazl Shiroudi, Khalil Pourshamsian, Ahmad Reza Oliaey, and Farhad Hatamjafari. "Theoretical investigations on the HOMO–LUMO gap and global reactivity descriptor studies, natural bond orbital, and nucleus-independent chemical shifts analyses of 3-phenylbenzo[d]thiazole-2(3H)-imine and its para-substituted derivatives: Solvent and substituent effects". In: *J. Chem. Res.* 45.1-2 (2021), pp. 147–158. DOI: 10.1177/1747519820932091.

[60] Alexandre V. Brethomé, Stephen P. Fletcher, and Robert S. Paton. "Conformational Effects on Physical-Organic Descriptors: The Case of Sterimol Steric Parameters". In: *ACS Catal.* 9.3 (2019), pp. 2313–2323. DOI: 10.1021/acscatal.8b04043.

[61] Adrián Gómez-Suárez, David J. Nelson, and Steven P. Nolan. "Quantifying and understanding the steric properties of N-heterocyclic carbenes". In: *Chem. Commun.* 53 (18 2017), pp. 2650–2660. DOI: 10.1039/C7CC00255F.

[62] Peter Dierkes and Piet W. N. M. van Leeuwen. "The bite angle makes the difference: a practical ligand parameter for diphosphine ligands". In: *J. Chem. Soc., Dalton Trans.* (10 1999), pp. 1519–1530. DOI: 10.1039/A807799A.

[63] Jesús Jover and Jordi Cirera. "Computational assessment on the Tolman cone angles for P-ligands". In: *Dalton Transactions* 48 (2019), p. 15036. DOI: 10.1039/c9dt02876e.

[64] Jenna A. Bilbrey, Arianna H. Kazez, Jason Locklin, and Wesley D. Allen. "Exact ligand cone angles". In: *J. Comput. Chem.* 34.14 (2013), pp. 1189–1197. DOI: https://doi.org/10.1002/jcc.23217.

A

Supporting information

This appendix consists of the supporting information of the paper. References to the SI made in the paper will be in this supporting information.

A Biphosphine bidentate ligands

Table 1: Studied biphosphine bidentate ligands

Ligand number	Ligand alias	CAS	Formula
1	BINAM-P	74974-14-4	C44H34N2P2
2	BINAP	76189-55-4	C44H32P2
3	BTFM-GarPhos	1365531-84-5	C48H28F24O4P2
4	Tol-BINAP	99646-28-3	C48H40P2
5	Xyl-BINAP	137219-86-4	C52H48P2
6	H8-BINAP	139139-86-9	C44H40P2
7	SegPhos	210169-54-3	C38H28O4P2
8	Xyl-SegPhos	210169-57-6	C46H44O4P2
9	DTBM-SegPhos	210169-40-7	C74H100O8P2
10	Cl-MeO-BIPHEP	185913-97-7	C38H30Cl2O2P2
11	SL-A109-1	352655-61-9	C74H104O6P2
12	SL-A120-1	394248-45-4	C46H48O2P2
13	SL-A107-1	352655-40-4	C70H100N4O2P2
14	SL-A108-2	145214-59-1	C30H24O6P2
15	SL-A102-2	133545-25-2	C42H40O2P2
16	SL-A121-1	192138-05-9	C70H96O2P2
17	SL-A104-1	256390-47-3	C50H56O14P2
18	GarPhos	1365531-75-4	C40H36O4P2
19	Xyl-GarPhos	1365531-89-0	C48H52O4P2
20	DTBM-GarPhos	1365531-98-1	C76H108O8P2
21	iPr-BIPHEP	150971-43-0	C26H40O2P2
22	C3-TunePhos	301847-89-2	C39H32O2P2
23	iPr-BPE	528854-34-4	C22H44P2
24	SPIRAP	NA	C43H38O2P2
25	DuanPhos	528814-26-8	C24H32P2
26	DiPamp	55739-58-7	C28H28O2P2
27	Xyl-SDP	917377-75-4	C49H50P2
28	Xyl-SKP	1429939-35-4	C52H54O2P2
29	Ph-BPE	528565-79-9	C34H36P2
30	ChiraPhos	64896-28-2	C28H28P2
31	Et-BPE	136705-62-9	C18H36P2
32	QuinoxP	866081-62-1	C18H28N2P2
33	Et-DuPhos	136705-64-1	C22H36P2
34	Me-DuPhos	147253-67-6	C18H28P2
35	PhanePhos	192463-40-4	C40H34P2
36	Xyl-PhanePhos	325168-89-6	C48H50P2
37	BDPP	96183-46-9	C29H30P2
38	NorPhos	71042-55-2	C31H28P2
39	DIOP	32305-98-9	C31H32O2P2
40	DPE-Phos	2119686-55-2	C38H32O3P2
41	BABIBOP	2207601-04-3	C22H28O2P2
42	Me-BABIBOP	2207601-10-1	C24H32O2P2
43	iPr-BABIBOP	2207601-12-3	C28H40O2P2
44	Me-BIBOP	1884680-48-1	C38H44O6P2
45	PPM	77450-05-6	C29H29NP2
46	SL-A101-2	133545-16-1	C38H32O2P2
47	MeO-F12-BIPHEP	116008-37-6	C38H20F12O2P2
48	MeO-F16-BIPHEP	NA	C42H24F16O2P2
49	MeO-py-F12-BIPHEP	NA	C38H24F12N4O2P2
50	MeO-F20-BIPHEP	NA	C42H20F20O2P2
51	MeO-BFPy-BIPHEP	NA	C42H20F24N4O2P2
52	XylSKEWPhos	551950-92-6	C37H46P2
53	DIPSKEWPhos	NA	C53H78P2

54	catASium D(R)	99135-95-2	C35H33NP2
55	Xyl-P-Phos	442905-33-1	C46H50N2O4P2
56	ProPhos	67884-32-6	C27H26P2
57	[Bis(1,1-dimethylethyl)phosphino]-3-(1,1-dimethylethyl) -2,3-dihydro-4-methoxy-1,3-benzoxaphosphole	1215081-28-9	C20H34O2P2
58	BenzP*	919778-41-9	C16H28P2
59	1-Naphthyl-DiPamp	256469-70-2	C34H28P2
60	Tol-SKP	1429939-32-1	C48H46O2P2
61	BCPM	114751-47-2	C34H49NO2P2
62	DiFluorPhos	503538-69-0	C38H24F4O4P2
63	Me-BPE	129648-07-3	C14H28P2
64	SynPhos	445467-61-8	C40H32O4P2
65	SDP	917377-74-3	C41H34P2
66	Ph-BIBOP	2301856-53-9	C34H36O2P2
67	Tol-SDP	528521-87-1	C45H42P2
68	DMM-GarPhos	1365531-93-6	C52H60O8P2
69	An-PhanePhos	364732-86-5	C44H42O4P2
70	BINAPINE	528854-26-4	C52H48P2
71	MeO-BIBOP	1202033-19-9	C24H32O4P2
72	CTH-BINAM	208248-67-3	C44H42N2P2
73	BPPM	72598-03-9	C34H37NO2P2
74	Ph-SKP	1360823-43-3	C44H38O2P2
75	CTH-P-Phos	221012-82-4	C38H34N2O4P2
76	Tol-GarPhos	1365531-81-2	C44H44O4P2
77	BIBOP	1202033-17-7	C22H28O2P2
78	Cy-GarPhos	2829282-18-8	C40H60O4P2
79	tBuPh-SKEWPhos	911415-22-0	C45H62P2
80	BINAPhane	253311-88-5	C50H36P2
81	iPr-DuPhos	136705-65-2	C26H44P2
82	Et-BABIBOP	2415751-83-4	C26H36O2P2
83	WingPhos	1884680-45-8	C50H44O2P2
84	C2-TunePhos	301847-88-1	C38H30O2P2
85	CyPP	70774-28-6	C32H34P2
86	SunPhos	765312-54-7	C42H36O4P2
87	1,4-dimethyl-1,4-butanediylbis(diphenylphosphine)	142494-67-5	C30H32P2
88	XantPhos	2119686-35-8	C41H36O3P2

B Automated substrate removal

Structures of TM-complexes without the substrate are essential for determining substrate binding energy. The automation of substrate removal has the potential to save time and reduce errors in structures. The general workflow of the substrate remover is shown in Figure 1

The substrate is consistently found at a fixed location for MACE generated xyz files. A MACE

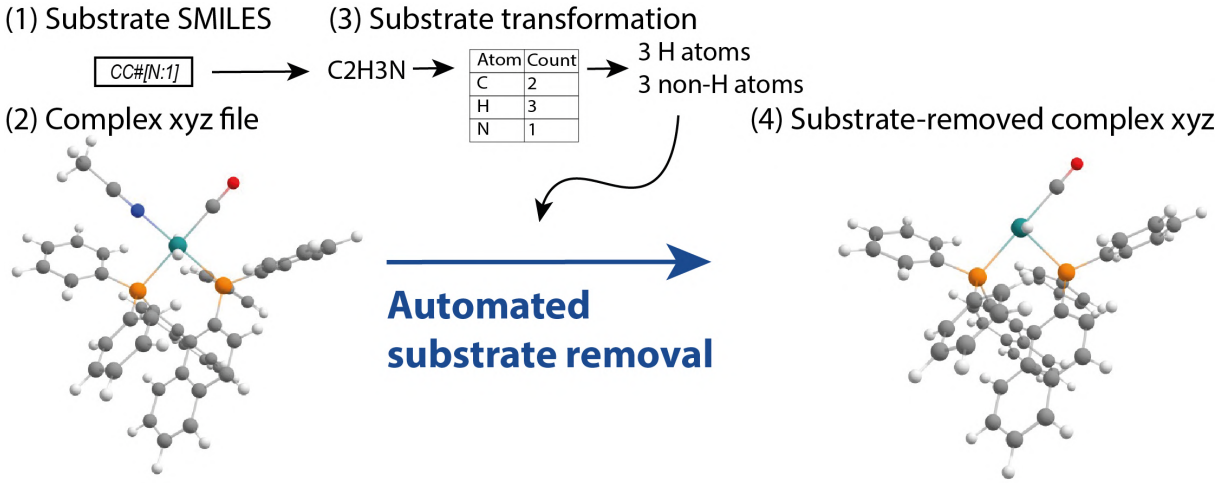


Figure 1: Automated substrate removal workflow, showing inputes (1) and (2), substrate transformation (3) and workflow output (4)

generated .xyz file consists of individual rows for each atom present in the TM-complex. All hydrogen atoms present are located at the bottom rows of the file. Substrate hydrogen atoms are consistently found at the initial hydrogen rows, while all non-hydrogen substrate atoms are found at the foremost rows. This arrangement of atoms makes substrate removal straightforward if the number of hydrogen and non-hydrogen atoms of the substrate is known.

The following input has to be provided to the substrate remover: TM-complex directory, bidentate ligand number and substrate information. Substrate SMILES have already been supplied for structure generation input. For easy and integrated use of the substrate remover, these SMILES should also be used as input for the substrate remover. However, SMILES only represent the non-hydrogen atoms of the substrate. Fortunately, SMILES can be converted into chemical formulas through RDKit tools, as shown in Figure 1-3. With the chemical formula known, it becomes feasible to iterate over this formula to obtain the specific types and quantities of atoms present in the substrate. From this information, the number of hydrogen and non-hydrogen atoms can be determined.

Once substrate information is known, the substrate is removed from the MACE generated TM-complex. A new directory called *substrate_removed* is created to save the newly created complexes without substrate. The suffix *no_substrate* is added to the end of each individual file to denote the difference from the original file.

The substrate of interest for this research is acetonitrile, known as the simplest organic nitrile molecule. To evaluate the flexibility of the substrate removal script, complexes containing three larger organic substrates are generated and the performance of the substrate remover is evaluated. Structures of the three organic substrates are shown in Figure 2. In addition to the size of the molecules, these substrates contain ester and amide groups, sm1 and sm 3 also contain a phenyl ring. To assess the accuracy of the automated substrate removal script, a comparison is made between manually removed substrate and automatically removed substrate data. The results confirm that both files contain the exact same 3D structure of the substrate removed TM-complexes, indicating that the automated substrate removal script functions as intended.

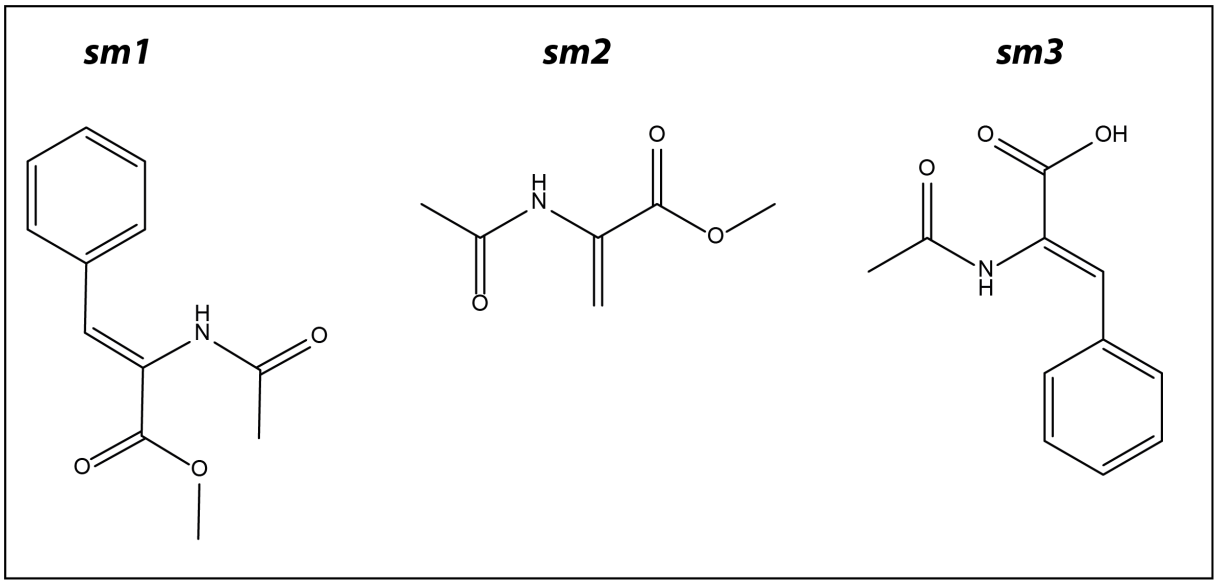


Figure 2: Other substrates used as test cases for the substrate removed

C Descriptor list

The following table lists all studied descriptors to identify differences between ligand configurations. These descriptors are categorized and a definition is provided.

Table 2: List of analysed descriptors

Descriptor	Category	Definition
Hardness	electronic	Measure of convexity of the energy function, quantifies stability
Softness	electronic	Measure of convexity of the energy function, quantifies stability
HOMO energy	electronic	Energy of the highest occupied molecular orbital
LUMO energy	electronic	Energy of the lowest unoccupied molecular orbital
HOMO-LUMO gap	electronic	Difference between HOMO and LUMO energy
mulliken charge max donor	electronic	Highest mulliken charge of the 2 bidentate ligand donor atoms
mulliken charge min donor	electronic	Lowest mulliken charge of the 2 bidentate ligand donor atoms
mulliken charge metal center	electronic	Mulliken charge on the metal center
nbo charge max donor	electronic	Highest NBO charge of the 2 bidentate ligand donor atoms
nbo charge min donor	electronic	Lowest NBO charge of the 2 bidentate ligand donor atoms
nbo charge metal center	electronic	NBO charge on the metal center
Dispersion energy	electronic	Dispersion energy of atomic vibrations
Lone pair occupancy max donor	electronic	Measure of lone pair filling maximum bidentate donor
Lone pair occupancy min donor	electronic	Measure of lone pair filling minimum bidentate donor
dipole moment	electronic	Debye dipole moment
electrophilicity	electronic	Measure of ability to take up electrons
electronegativity	electronic	Measure to attract bonding electrons
Buried volume 7A	steric	Percentage of a 7A sphere around metal occupied by ligands
Buried volume 6A	steric	Percentage of a 6A sphere around metal occupied by ligands
Buried volume 5A	steric	Percentage of a 5A sphere around metal occupied by ligands
Buried volume 4A	steric	Percentage of a 4A sphere around metal occupied by ligands
buried volume donor min	steric	Percentage of a 7A sphere around metal occupied by ligands
buried volume donor max	steric	Percentage of a 7A sphere around metal occupied by ligands
buried volume 3.5A	steric	Percentage of a 7A sphere around metal occupied by ligands
distance metal center min donor	geometric	Distance between metal center and minimum bidentate ligand donor
distance metal center max donor	geometric	Distance between metal center and maximum bidentate ligand donor
cone angle	geometric	Measure of ligand bulk, angle between metal and ligand corners
bite angle	geometric	Angle between metal and 2 bidentate ligand donors

D Substrate binding energy results

Performed analysis for substrate binding energy is the same as for stability. First, energies between ligand configurations are compared to investigate trends in configuration most favorable in energy.

D.1 Energy differences between ligand configurations

Binding energy results are visualised in Figure 3a, On the left side, interaction ΔE_{bind} between ligand configurations are plotted for each metal center. Binding energy differences are depicted in the right graphs. Figure 3a illustrates the interaction ΔE_{bind} among iridium ligand configurations. At the top of this figure, the general reference structures of the ligand configurations are displayed, whereas the structures for the other ligand configurations are depicted at the bottom of the graph. In this figure, energy of the H-N axial ligand pair complexes are compared with with H-H axial ligand pair complexes for the tested bidentate ligands. A negative ΔE_{bind} is observed for the majority of bidentate ligands. This indicates that energy of the H-H axial ligand pair configuration is more favorable in energy over the H-N axial ligand pair ligand configurations for most bidentate ligands.

Figure 4b shows the binding energy differences between ligand configurations of ruthenium complexes. In this figure, the energy of complexes with axial ligands C-H, C-N and H-H are compared with that of complexes featuring the H-N axial ligand pair. General structure for the reference H-N axial ligand pair is shown at the top of the figure, and compared structures are shown at the bottom. The lowest energy configurational configuration often varies among the different bidentate ligands. However, a positive ΔE_{bind} is observed for a majority of ligand configurations, indicating that the H-N axial configuration is clearly not favorable in terms of substrate binding energy in comparison to the other ligand configurations.

ΔE_{bind} data for manganese ligand configurations is shown in Figure 3c. In this figure, substrate binding energies of the C-N, C-H and C-C ligand configurations are compared to the substrate binding energy of the H-N isomer. A significant number of ligand configurations show a positive ΔE_{bind}. However, no clear majority of most energy favorable configuration is observed.

Figure 4 illustrates the distribution of lowest binding energy ligand configurations among the axial ligand pairs. However, the preference of iridium complexes is less overwhelming for binding energy as opposed to interaction energy. For binding energies, the H-H axial ligand pair is lowest in energy for 64.7% of bidentate ligands, which is significantly lower than the 92% observed for interaction energies. For ruthenium and manganese complexes, most observed lowest energy configuration also changed for binding energies compared to interaction energies. Nevertheless, the determination of the lowest energy configurational configuration varies among the tested bidentate ligands. In most cases, no specific axial ligand pair is significantly more abundantly favorable over the other ligand configurations.

D.2 Effect of metstable ligand configurations

Figure 5 visualizes configuration statistics for binding energy data of the generated ligand configurations. Figure 5a shows the number of ligand configurations present within 10 kJ/mol range from lowest binding energy for each bidentate ligand. In Figure 5b, the percentage of bidentate ligands for which multiple ligand configurations have binding energies within 10 kJ/mol from the minimum is visualised. Similarly to interaction energy, 18.2% of iridium complexes contain multiple ligand configurations present within the energy range. Moreover, a slight increase in the percentage of bidentate ligands containing multiple ligand configurations within the 10 kJ/mol energy range is observed. Namely, 78.4% of ruthenium complexes and 72.7% of manganese complexes.

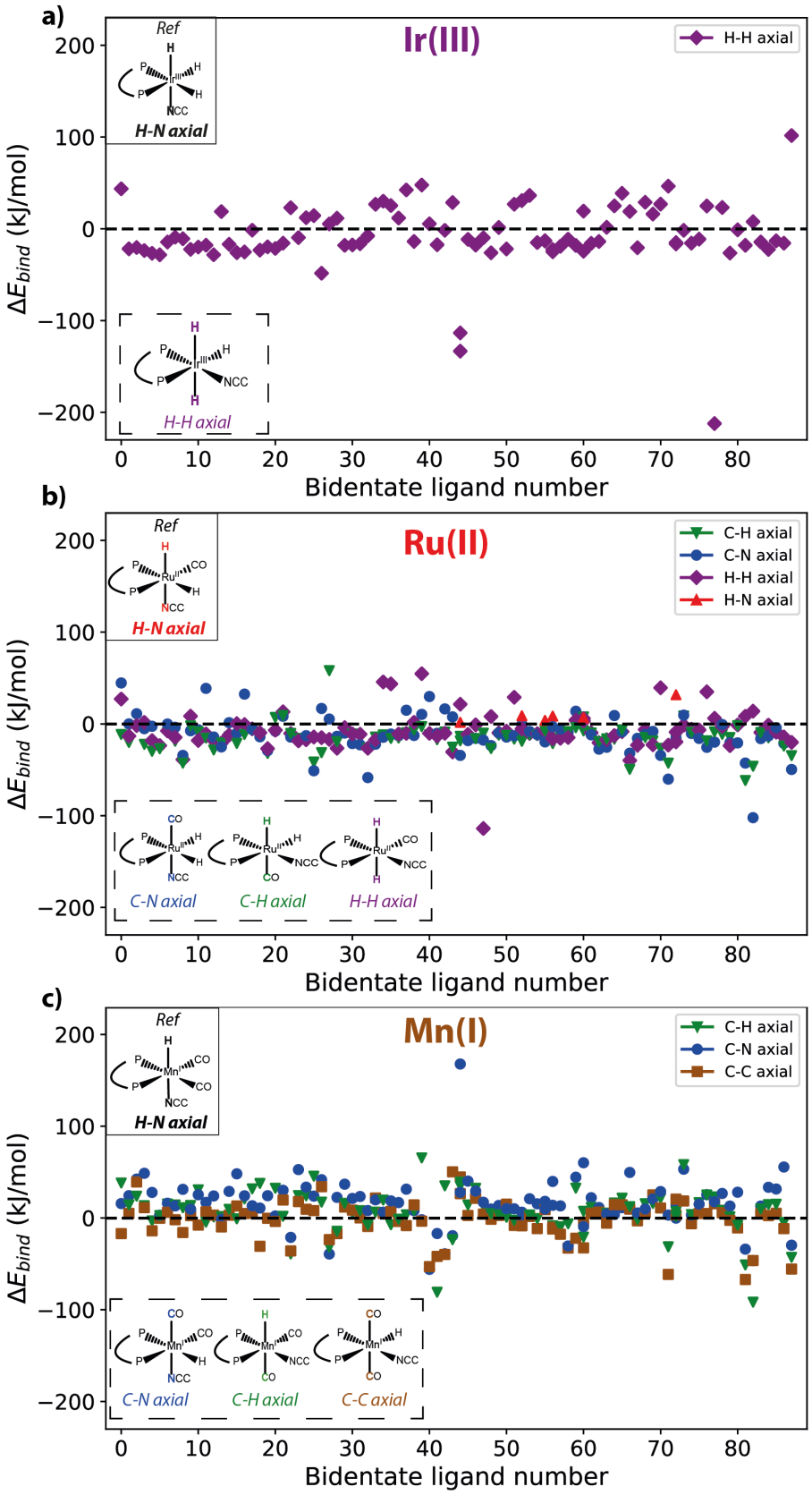


Figure 3: Substrate binding energy differences between a reference configurational isomer, shown at the top of the graphs, and other configurational ligand configurations, shown in the bottom of the graphs, for set of bidentate ligands for iridium (a), ruthenium (b) and manganese (c) complexes

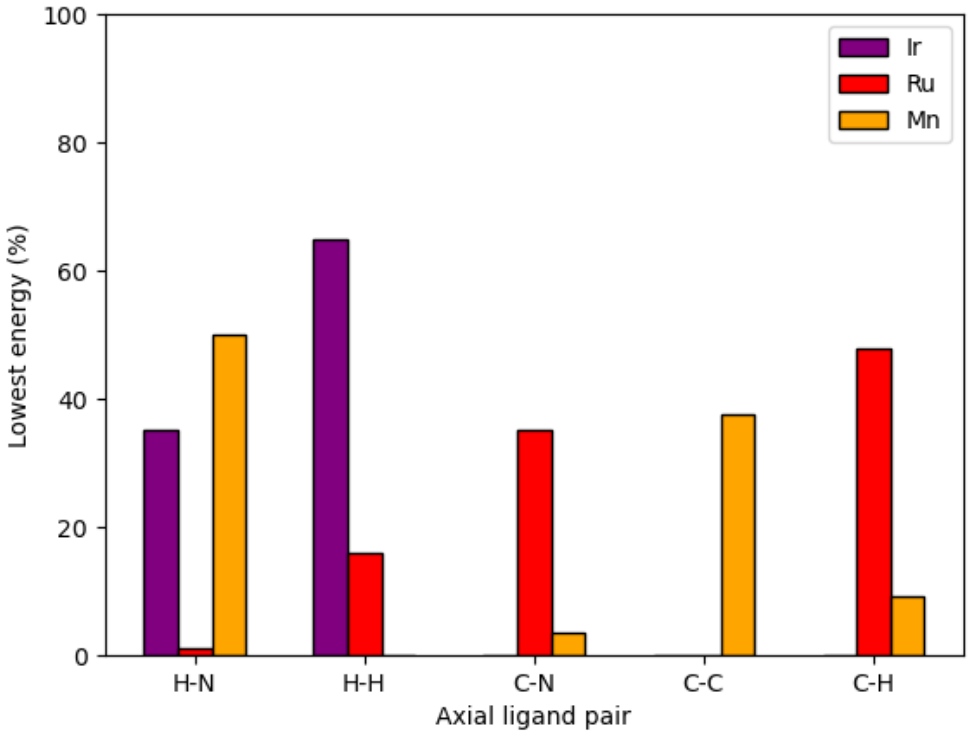


Figure 4: Distribution of the lowest binding energy configuration among axial ligand pairs for the 88 bidentate ligands in combination with 3 metal centers

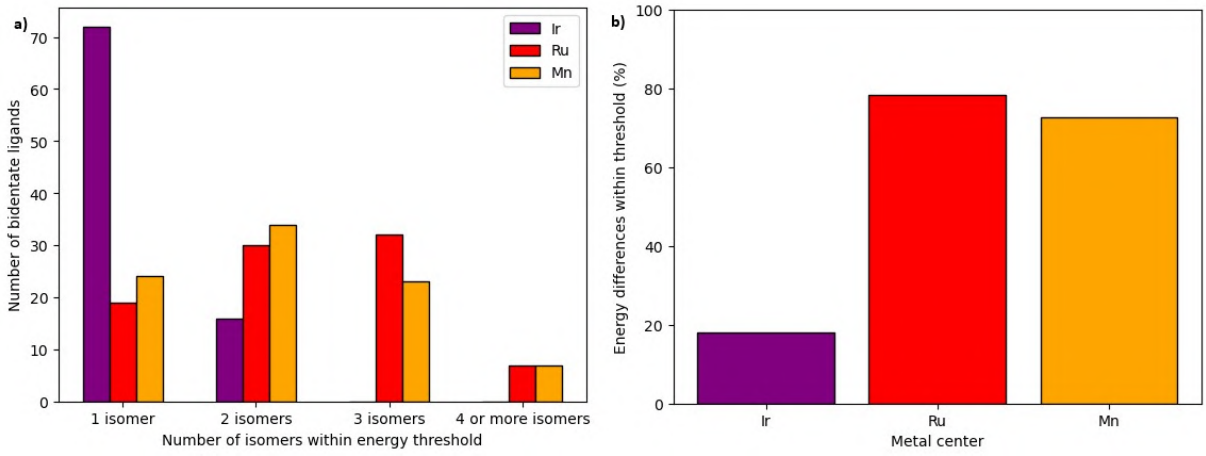


Figure 5: Number of ligand configurations present within a 10 kJ/mol energy range from the lowest binding energy configuration for the researched bidentate ligands (a) and the percentage of bidentate ligands for which multiple ligand configurations are found within the 10 kJ/mol energy range (b)

E Orbital visualization

In this section, visualized orbitals for two bisphosphine bidentate ligands are shown. Figure 6 shows the HOMO and LUMO orbitals for iridium complexes, Figure 7 shows the HOMO and LUMO orbitals for ruthenium complexes, and Figure 8 shows the HOMO and LUMO orbitals for manganese complexes.

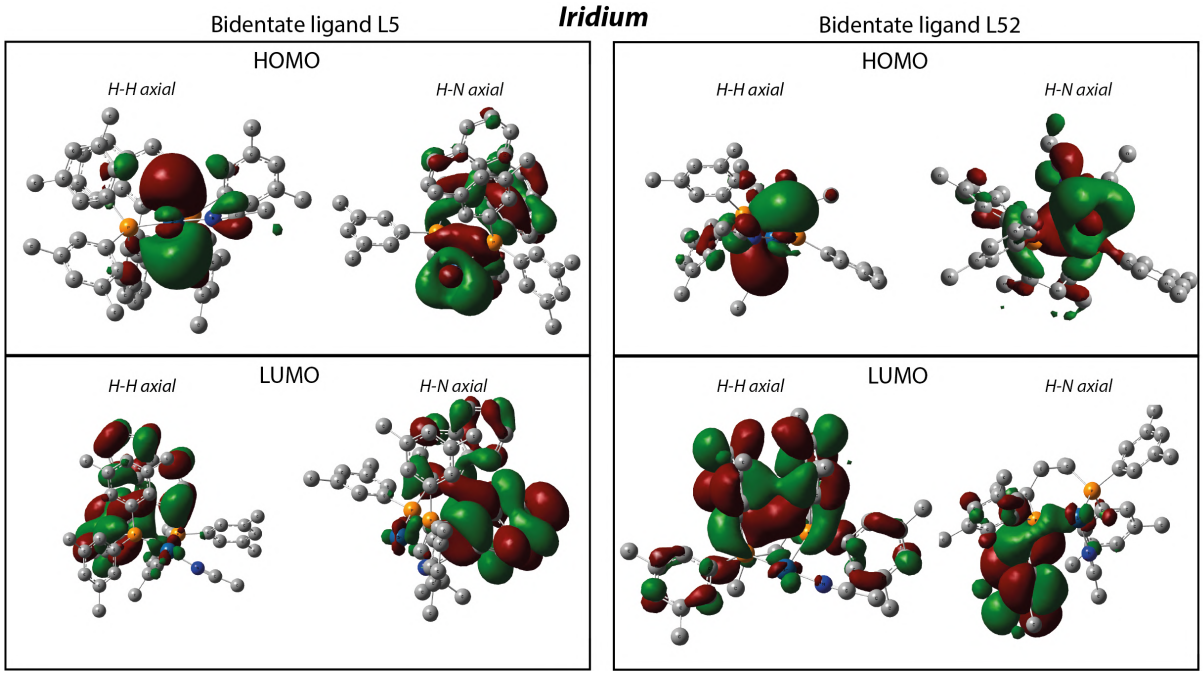


Figure 6: HOMO and LUMO orbital visualizations for different ligand configurations for iridium complexes

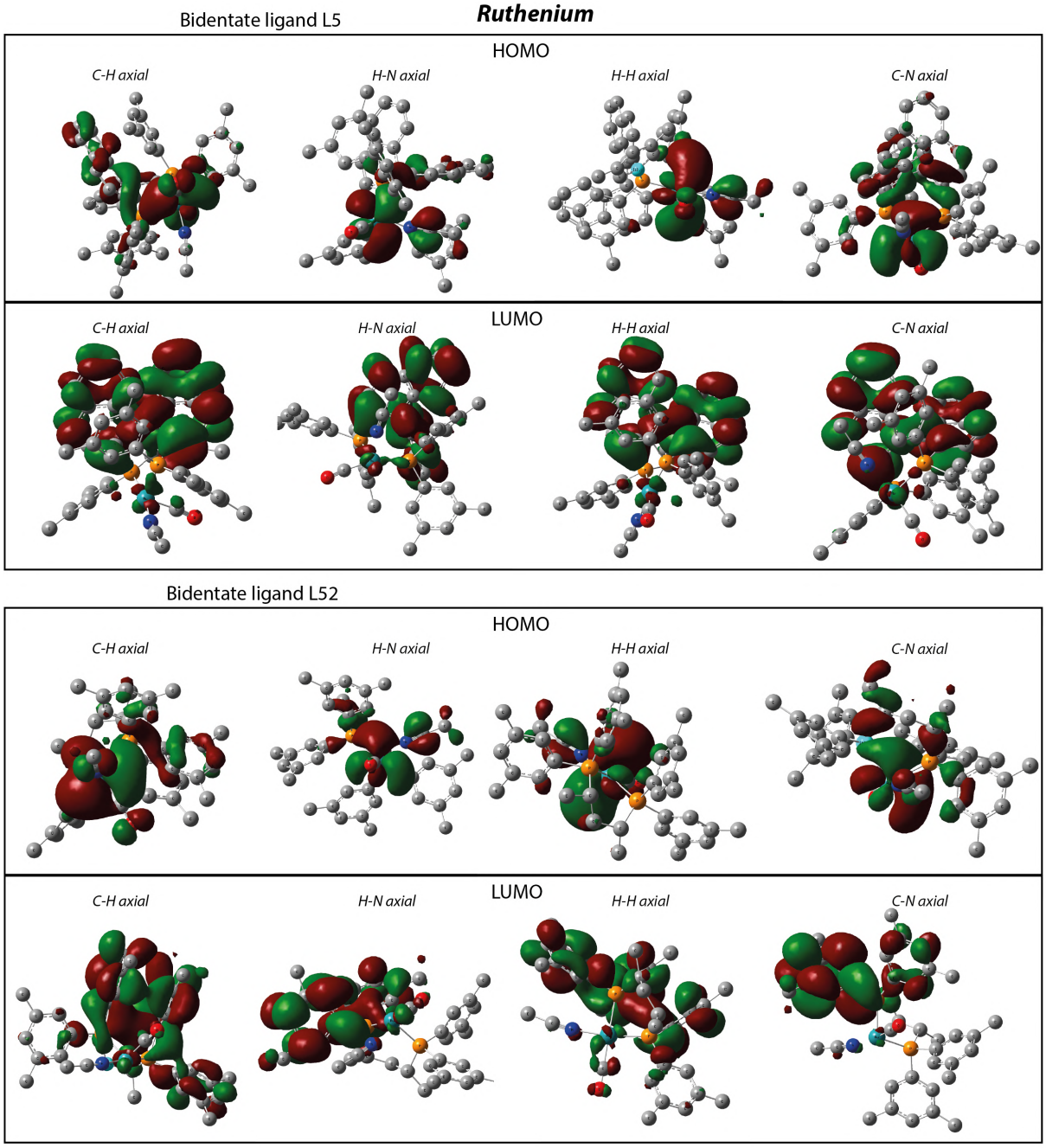


Figure 7: HOMO and LUMO orbital visualizations for different ligand configurations for ruthenium complexes

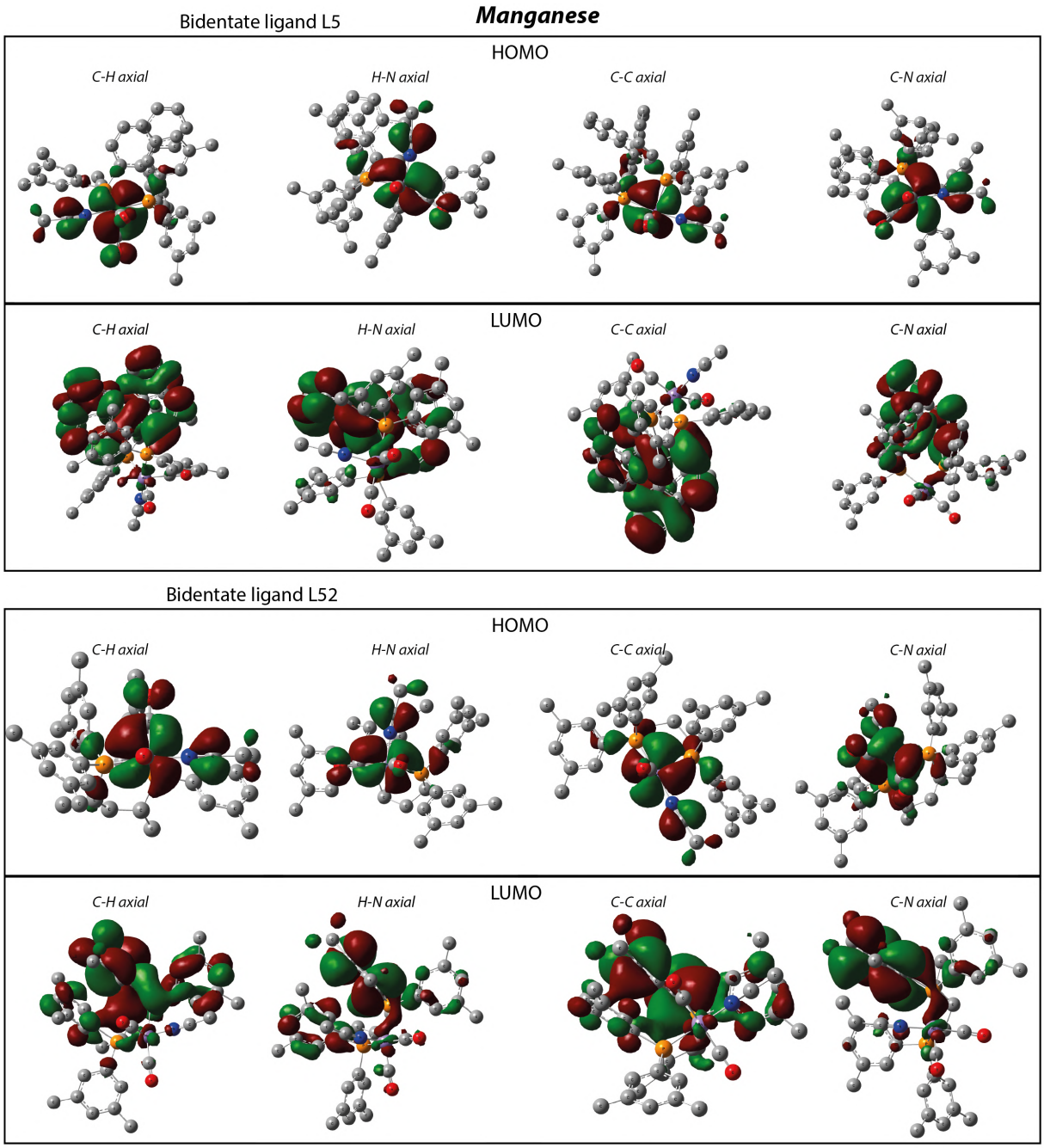
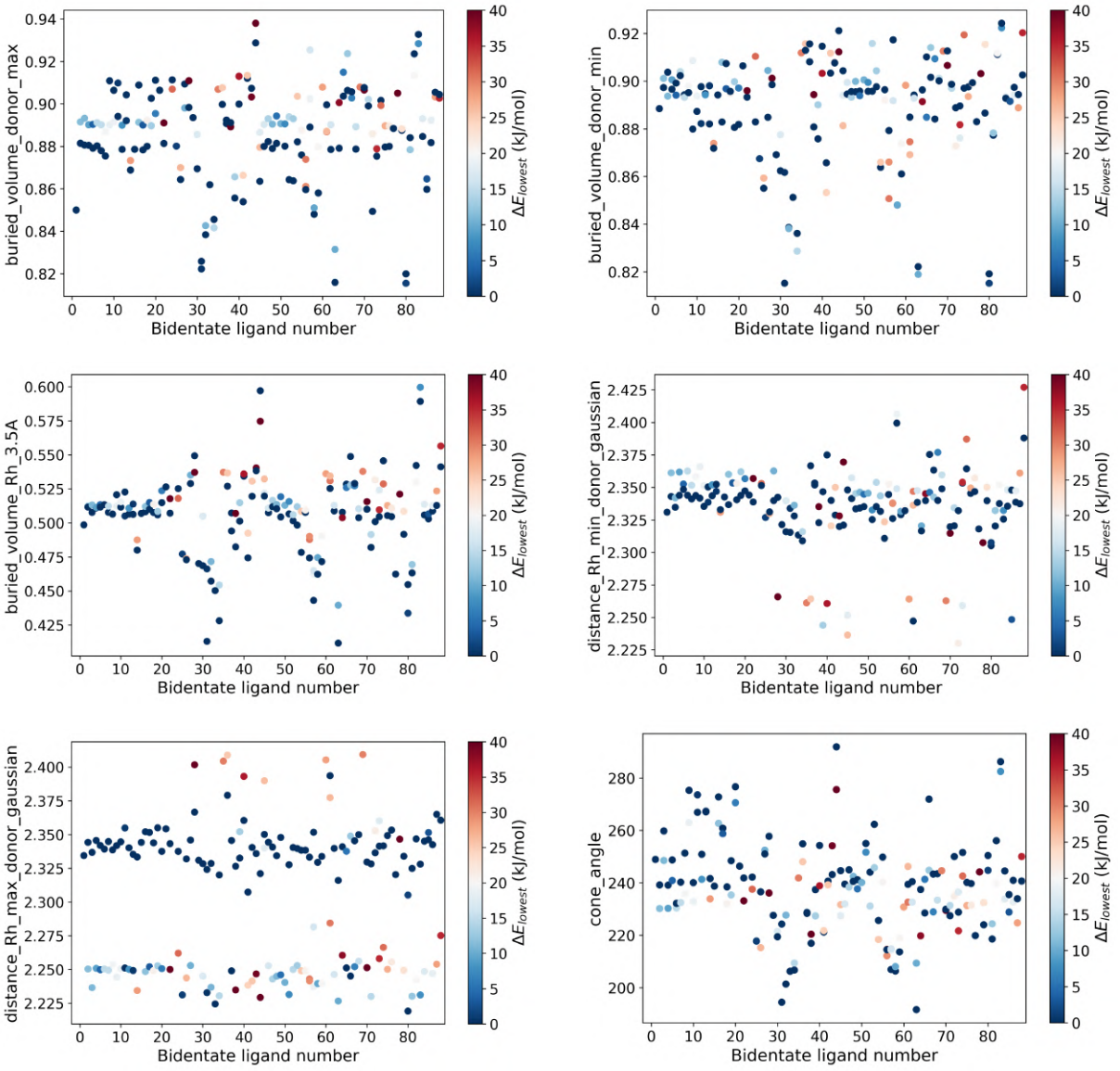
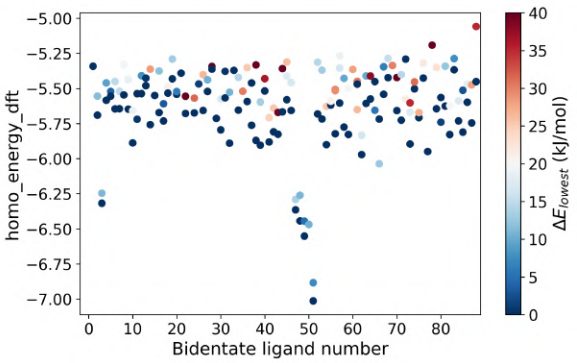


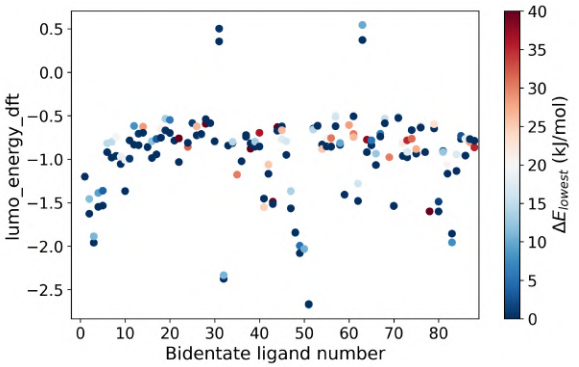
Figure 8: HOMO and LUMO orbital visualizations for different ligand configurations for manganese complexes

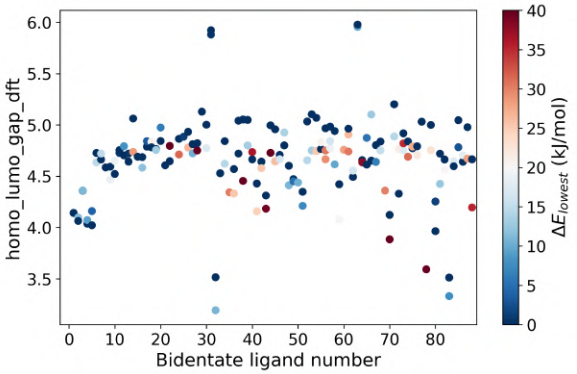
F Descriptor plots

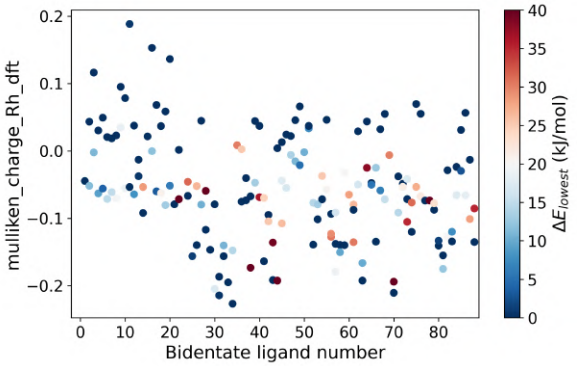
F.1 IrH3(CH3CN) plots

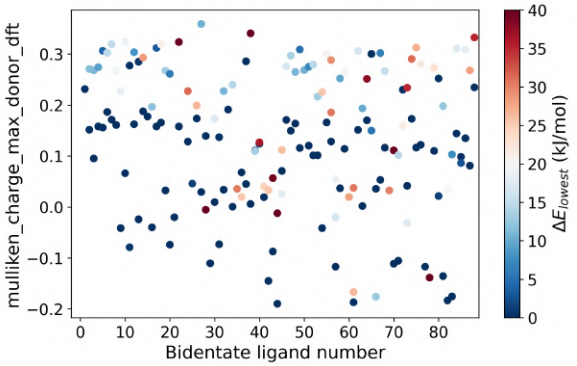


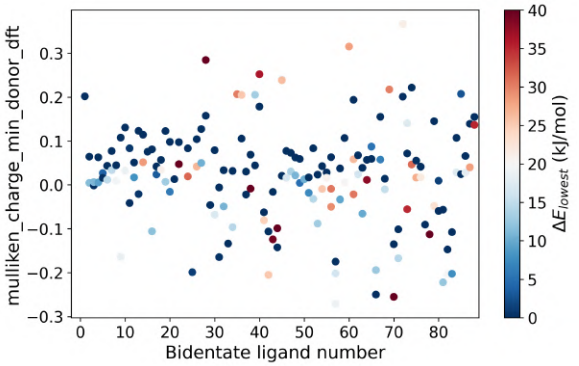

homo_energy_dft
−5.00
−5.25
−5.50
−5.75
−6.00
−6.25
−6.50
−6.75
−7.00
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

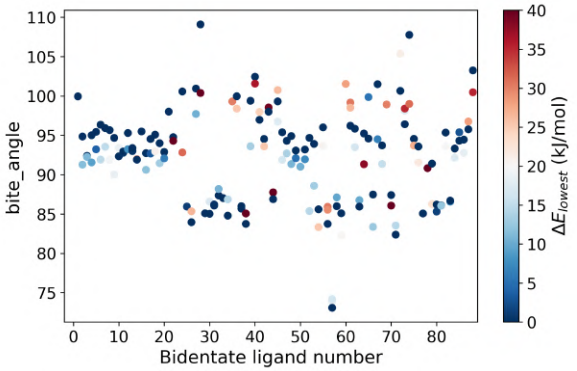

lumo_energy_dft
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

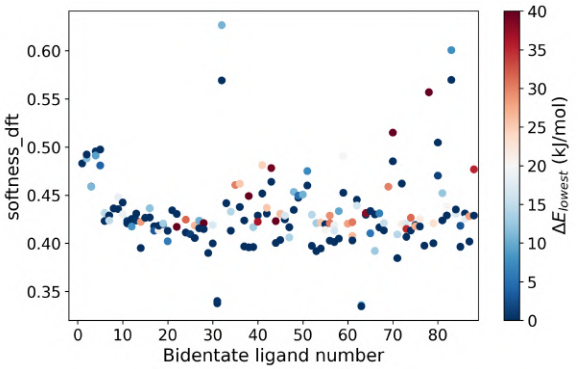

homo_lumo_gap_dft
6.0
5.5
5.0
4.5
4.0
3.5
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

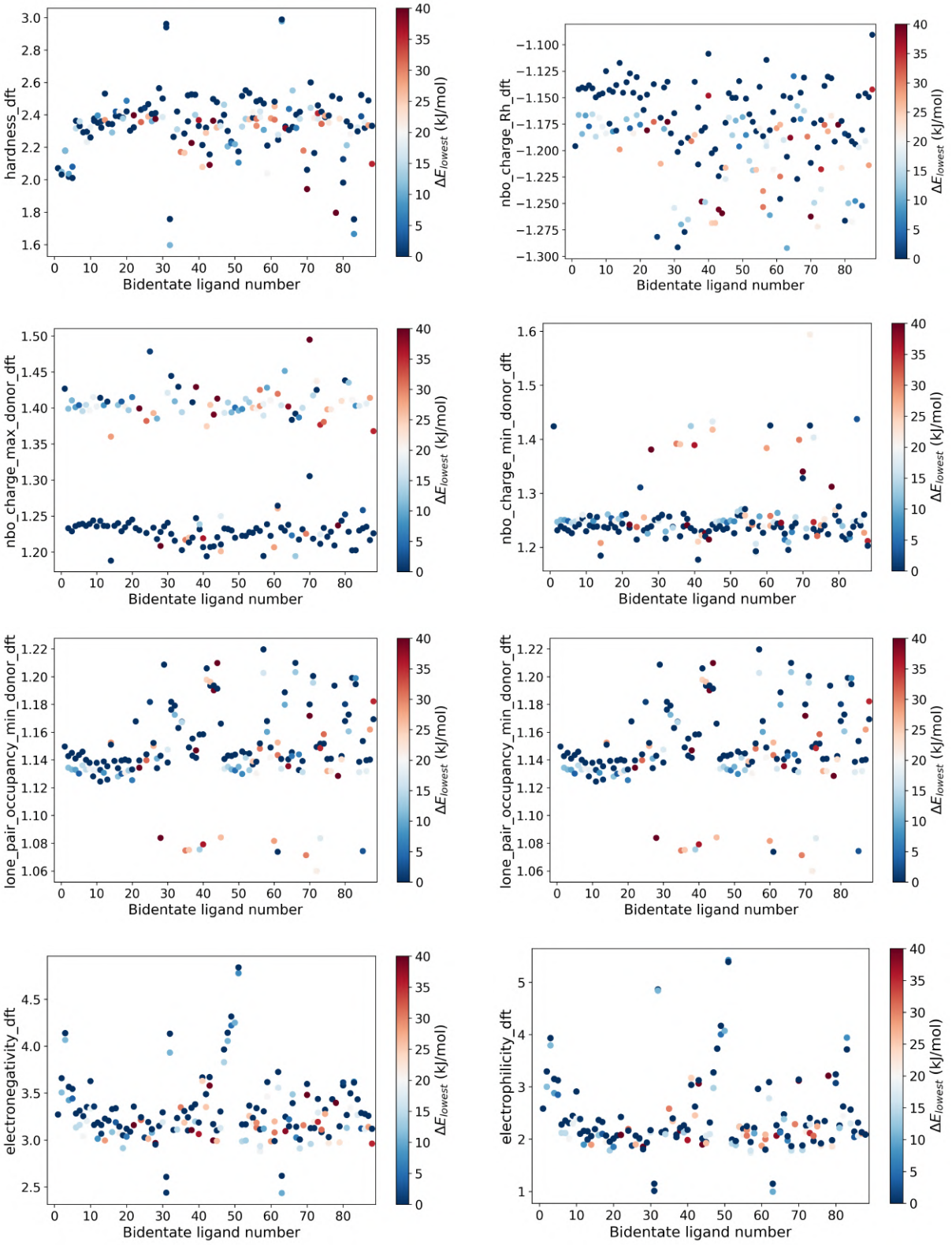

mulliken_charge_Rh_dft
0.2
0.1
0.0
−0.1
−0.2
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

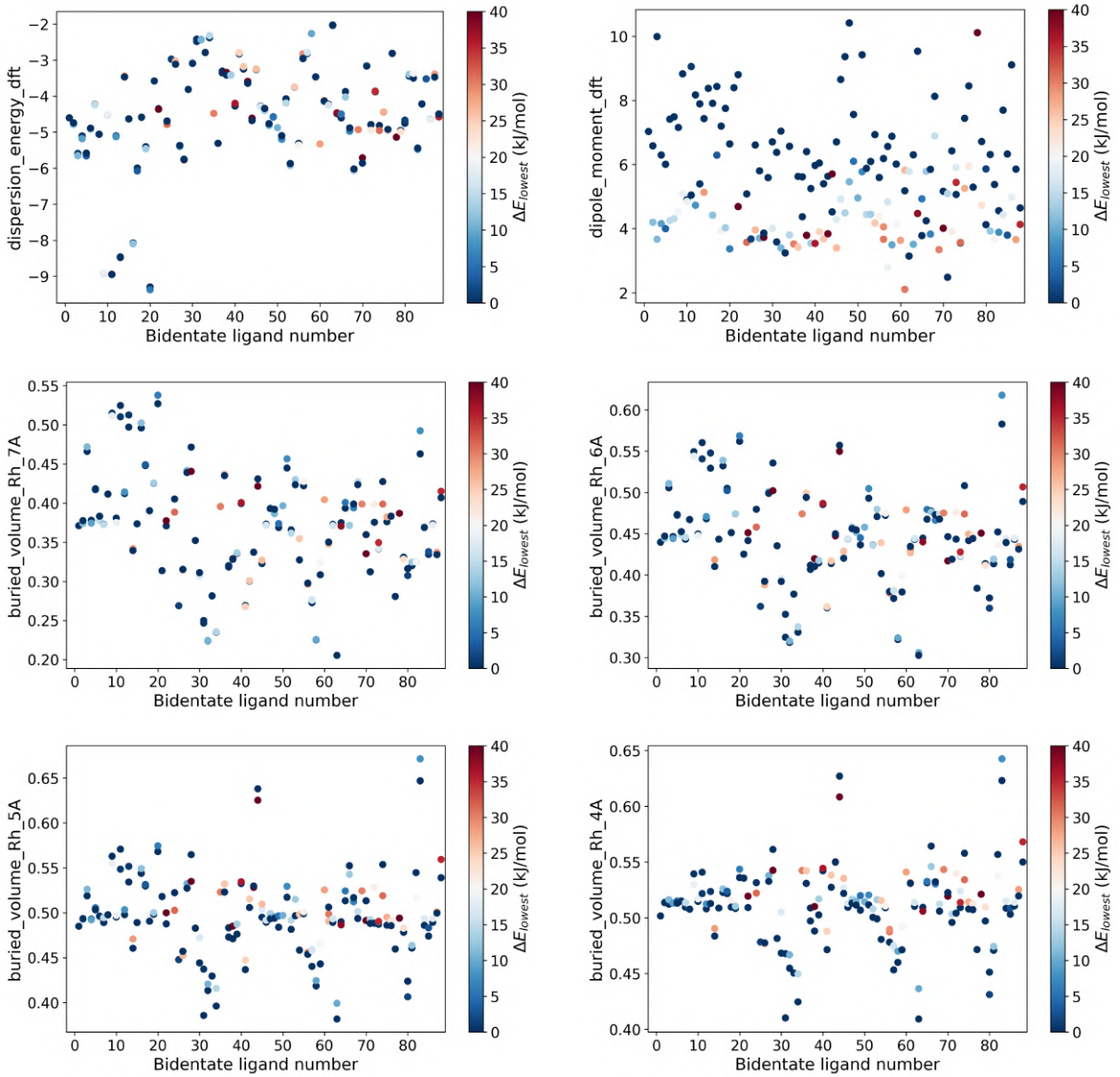

mulliken_charge_max_donor_dft
0.3
0.2
0.1
0.0
−0.1
−0.2
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40


mulliken_charge_min_donor_dft
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

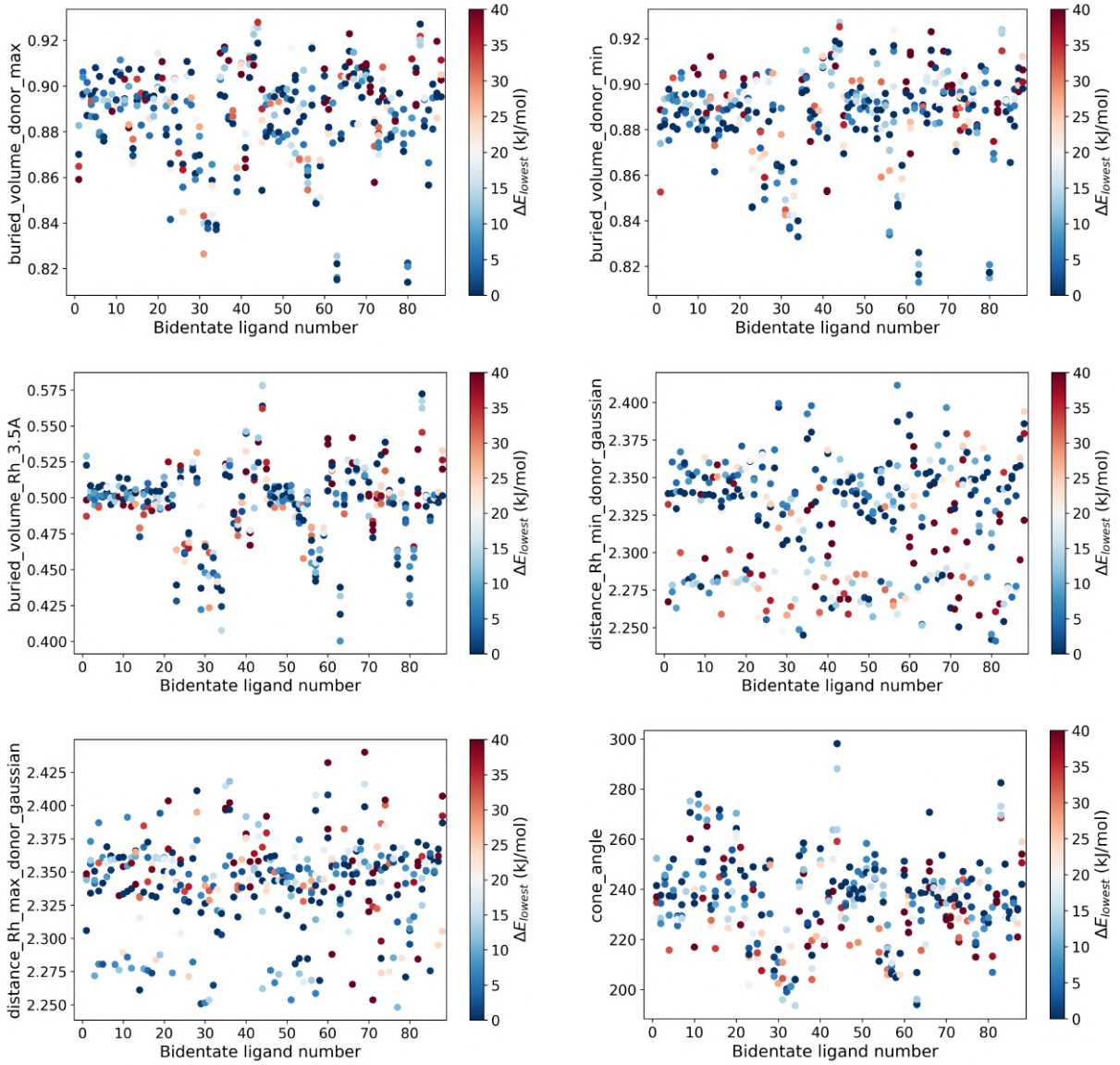

bite_angle
110
105
100
95
90
85
80
75
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

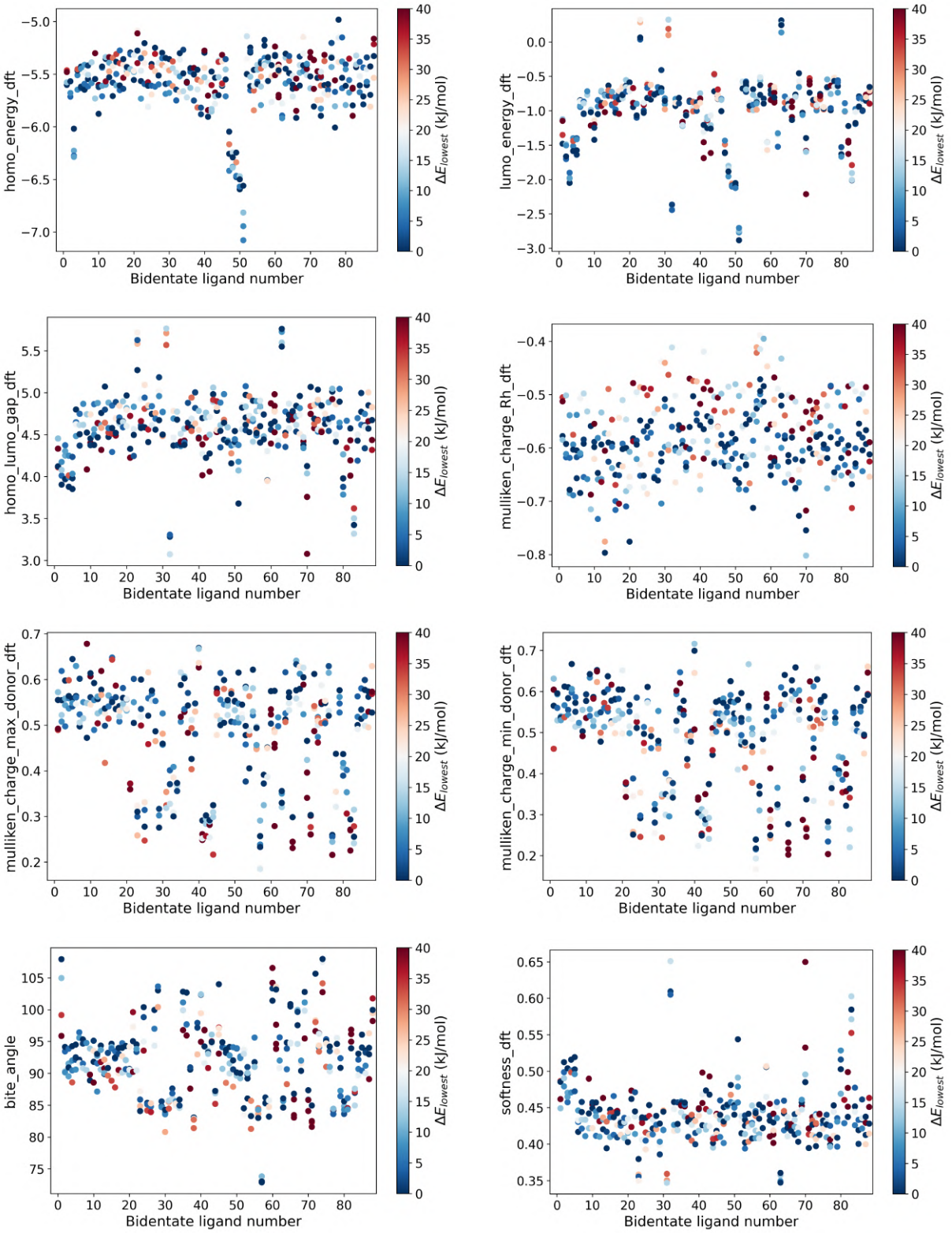

softness_dft
0.60
0.55
0.50
0.45
0.40
0.35
0 10 20 30 40 50 60 70 80
Bidentate ligand number
ΔE_lowest (kJ/mol)
0 5 10 15 20 25 30 35 40

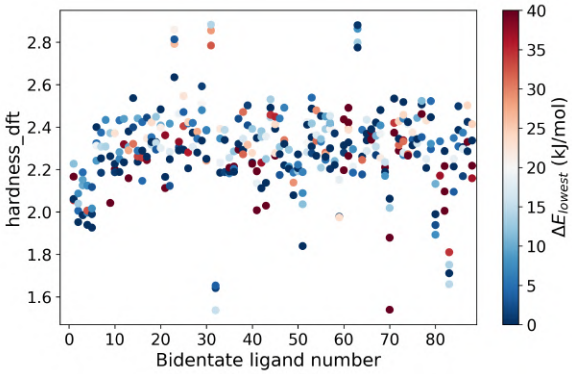

hardness_dft
nbo_charge_Rh_dft
nbo_charge_max_donor_dft
nbo_charge_min_donor_dft
lone_pair_occupancy_min_donor_dft
lone_pair_occupancy_min_donor_dft
electronegativity_dft
electrophilicity_dft
Bidentate ligand number
ΔE_lowest (kJ/mol)

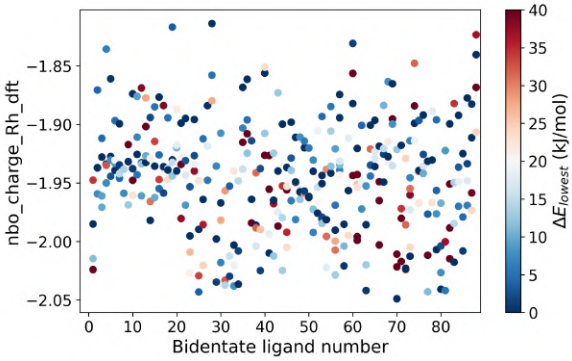

dispersion_energy_dft
dipole_moment_dft
buried_volume_Rh_7A
buried_volume_Rh_6A
buried_volume_Rh_5A
buried_volume_Rh_4A
Bidentate ligand number
ΔE_lowest (kJ/mol)

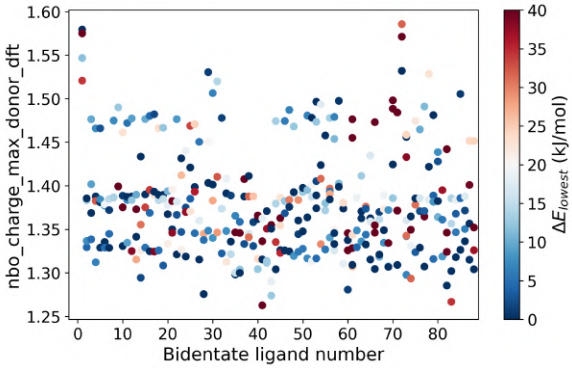
F.2 RuH2(CO)CH3CN plots

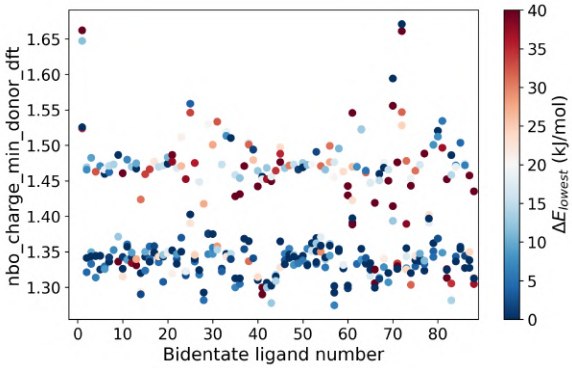


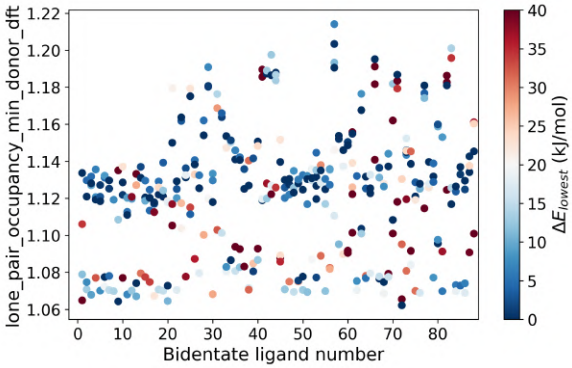

homo_energy_dft
lumo_energy_dft
homo_lumo_gap_dft
mulliken_charge_Rh_dft
mulliken_charge_max_donor_dft
mulliken_charge_min_donor_dft
bite_angle
softness_dft
Bidentate ligand number
ΔE_lowest (kJ/mol)

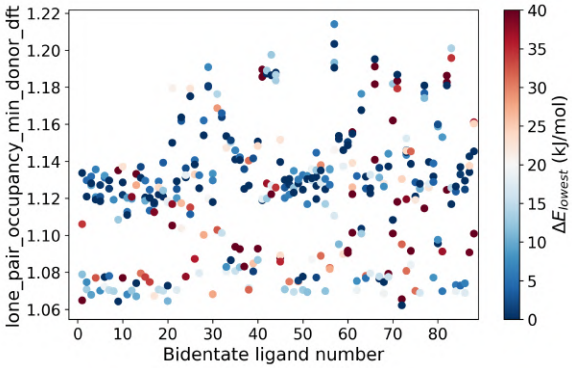

hardness_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

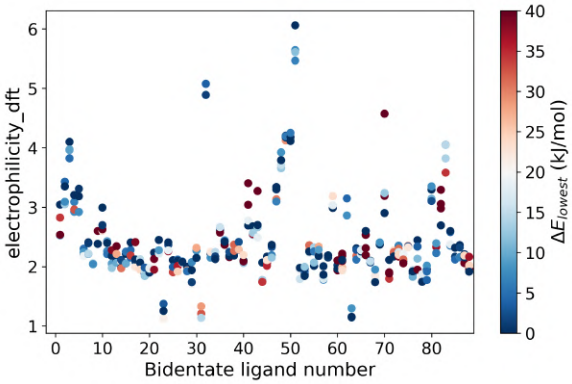

nbo_charge_Rh_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

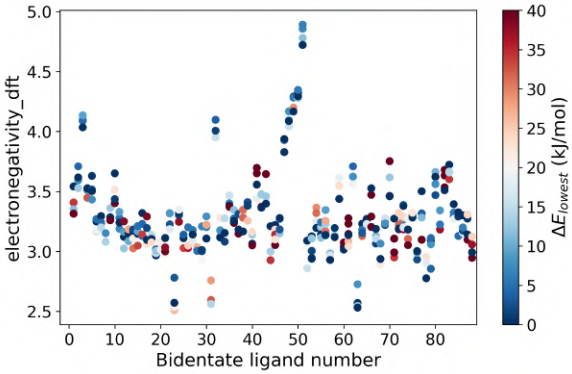

nbo_charge_max_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

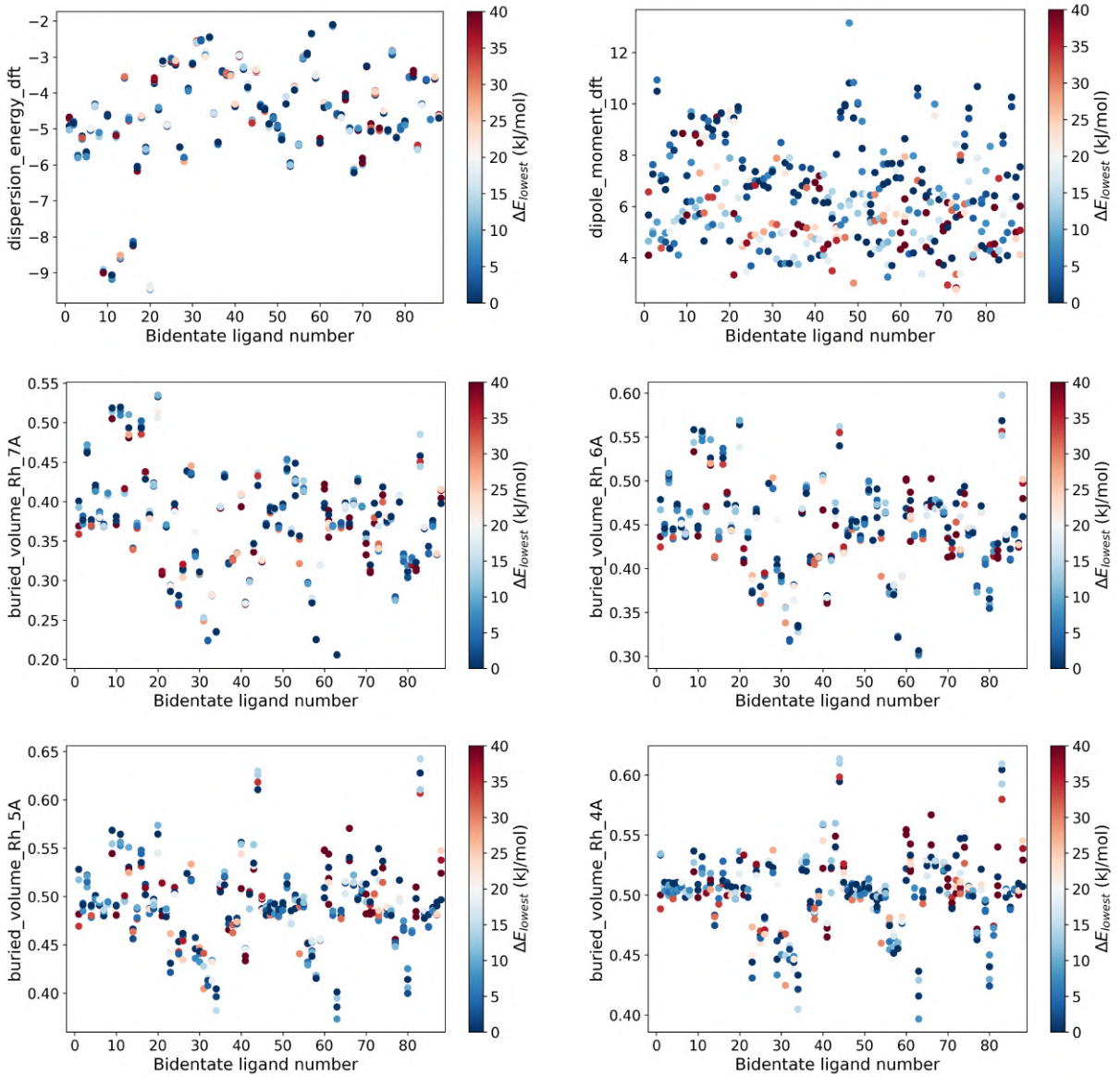

nbo_charge_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


lone_pair_occupancy_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

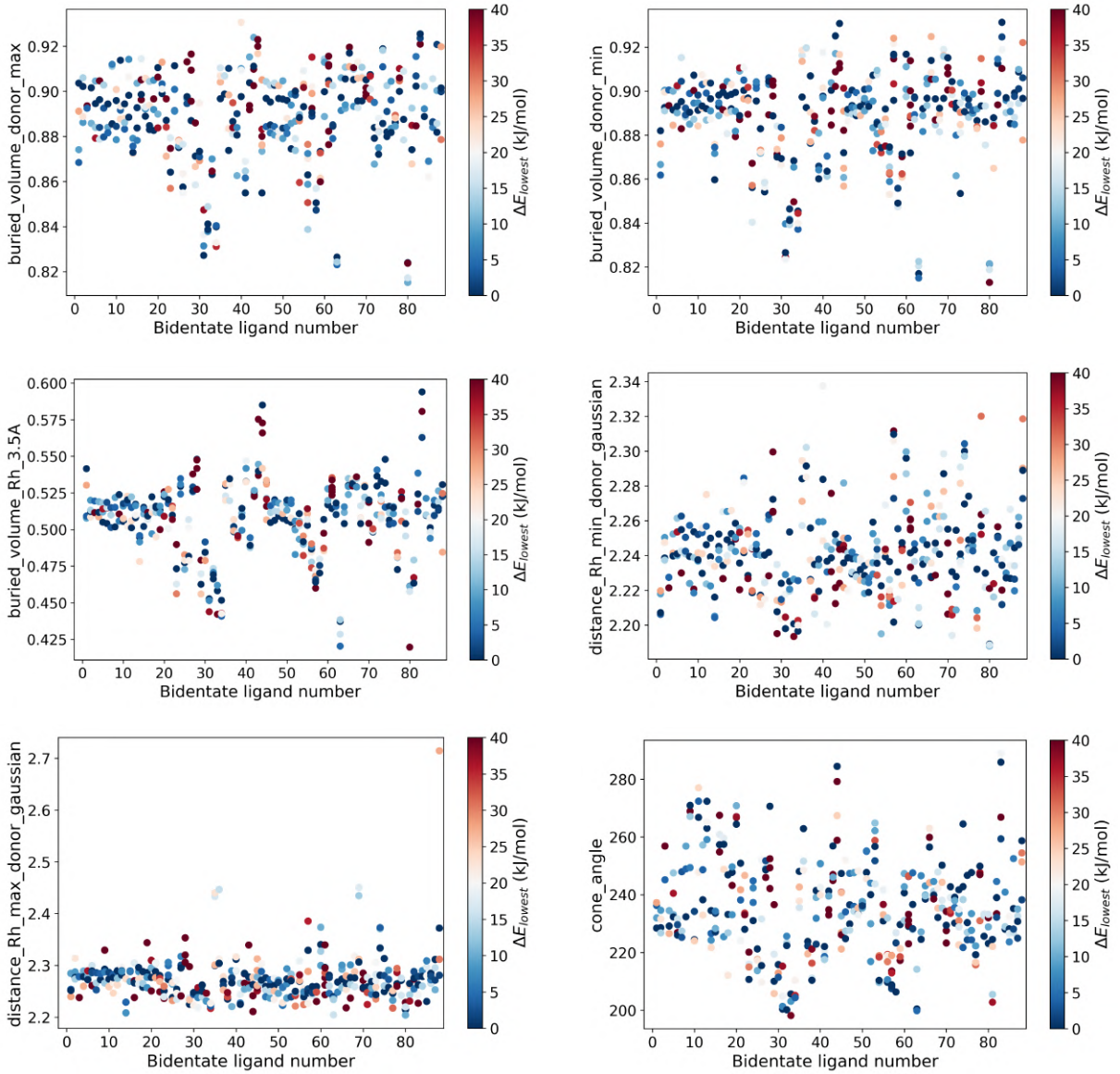

lone_pair_occupancy_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

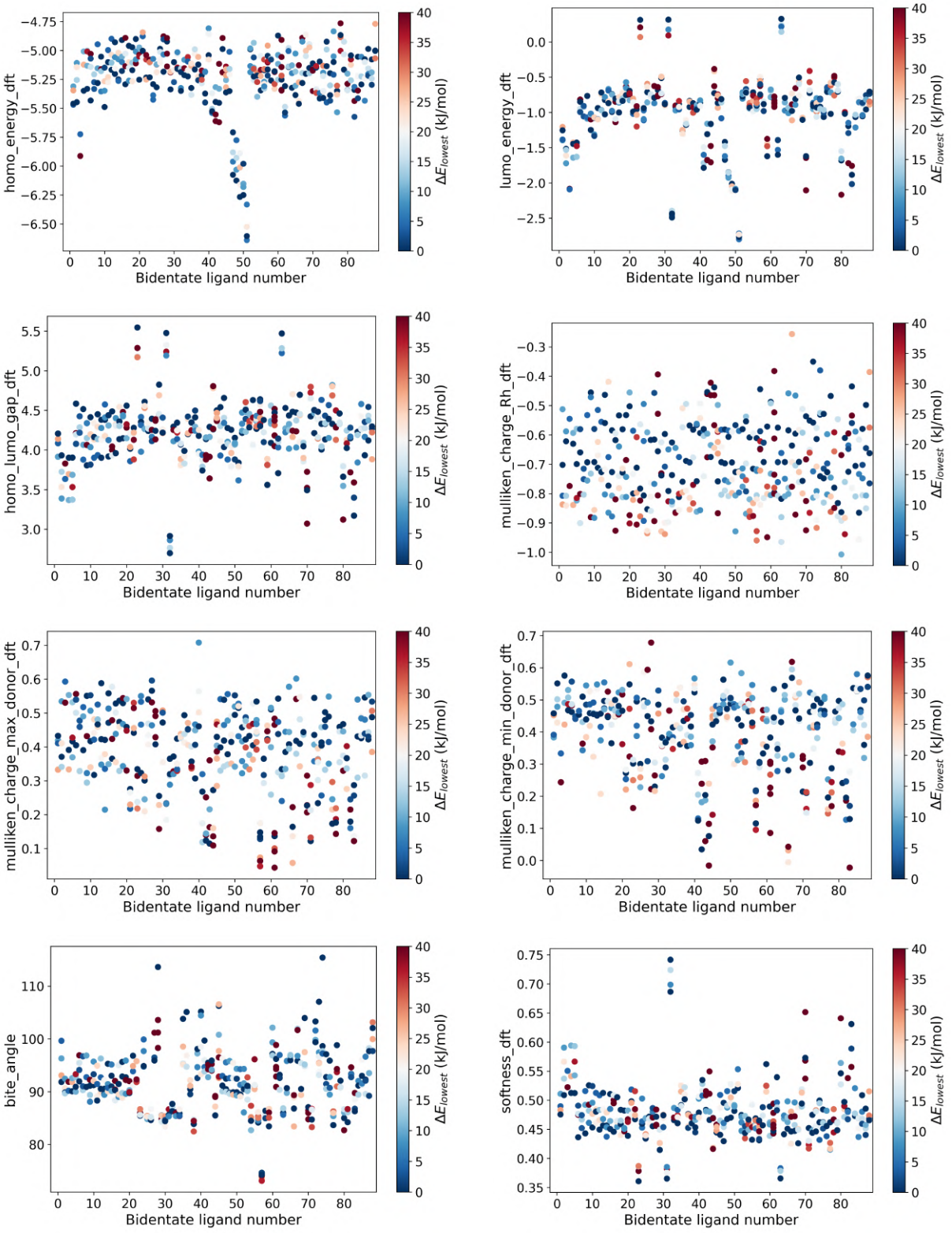

electrophilicity_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

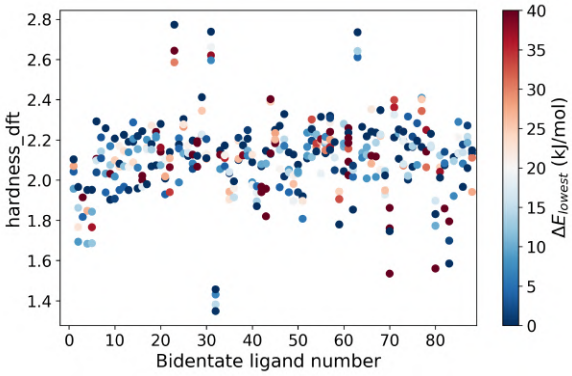

electronegativity_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

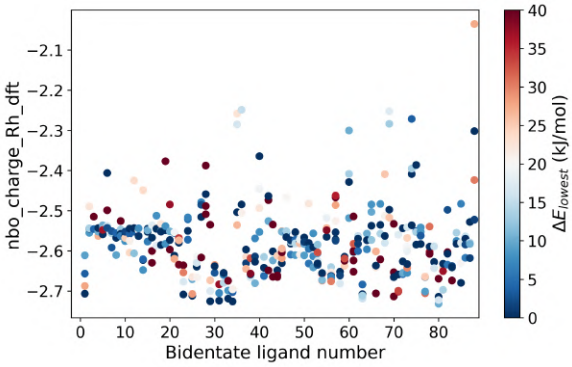

dispersion_energy_dft
dipole_moment_dft
buried_volume_Rh_7A
buried_volume_Rh_6A
buried_volume_Rh_5A
buried_volume_Rh_4A
Bidentate ligand number
ΔE_lowest (kJ/mol)

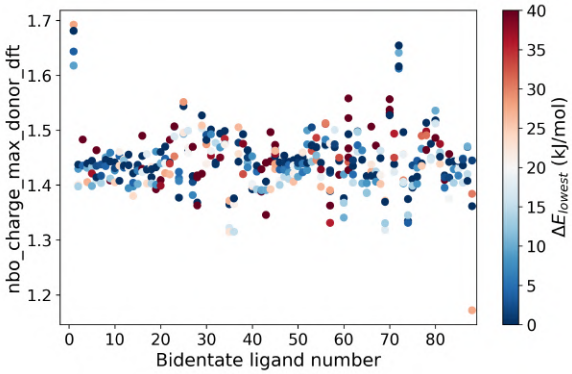
F.3 MnH(CO)2CH3CN plots

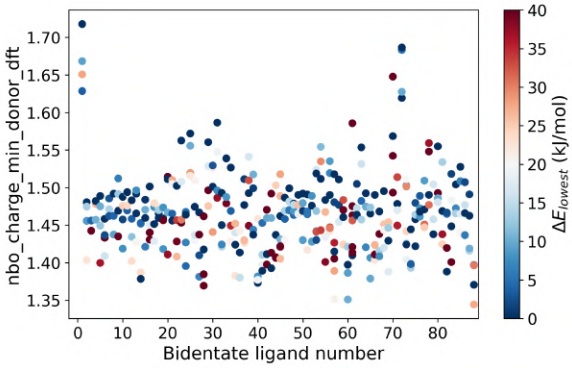


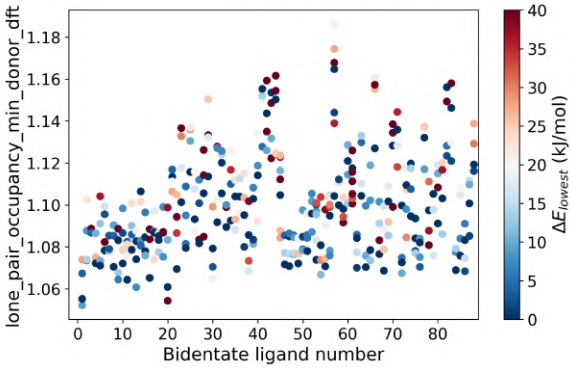

homo_energy_dft
lumo_energy_dft
homo_lumo_gap_dft
mulliken_charge_Rh_dft
mulliken_charge_max_donor_dft
mulliken_charge_min_donor_dft
bite_angle
softness_dft
Bidentate ligand number
ΔE_lowest (kJ/mol)

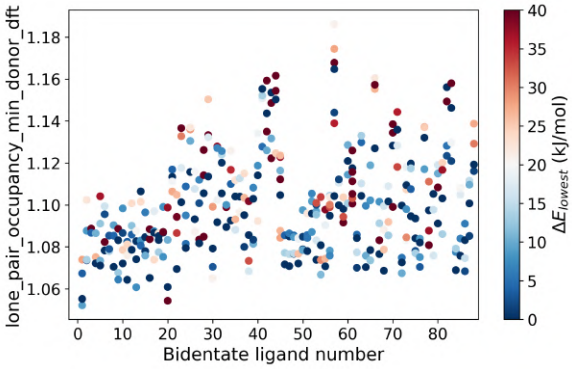

hardness_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

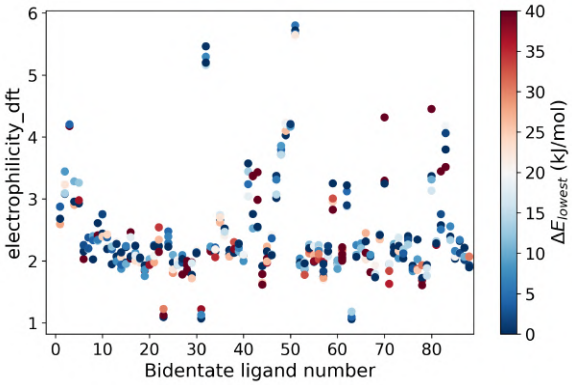

nbo_charge_Rh_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

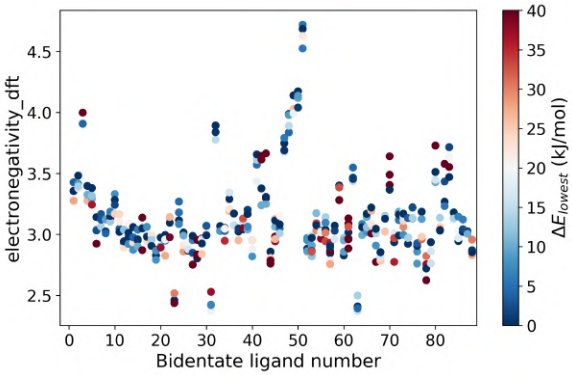

nbo_charge_max_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)

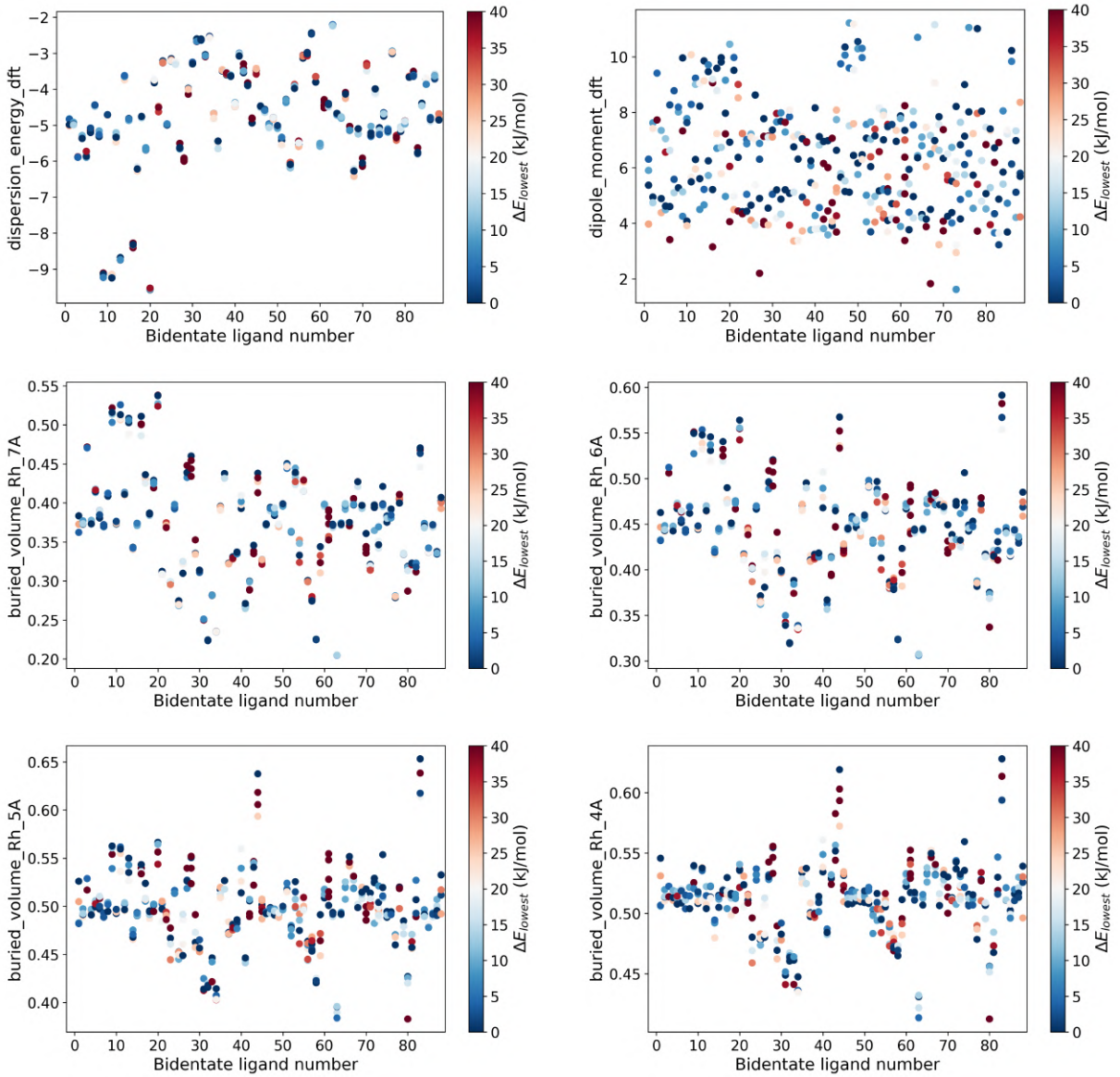

nbo_charge_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


lone_pair_occupancy_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


lone_pair_occupancy_min_donor_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


electrophilicity_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


electronegativity_dft
Bidentate ligand number
ΔE_{lowest} (kJ/mol)


dispersion_energy_dft
dipole_moment_dft
buried_volume_Rh_7A
buried_volume_Rh_6A
buried_volume_Rh_5A
buried_volume_Rh_4A
Bidentate ligand number
ΔE_lowest (kJ/mol)

G Multiple ligand configurations

The following figures display the number of ligand configurations identified within a 10 kJ/mol range from the most stable ligand configuration. A value of 1 indicates that only the lowest energy configuration is identified within the specified energy range.

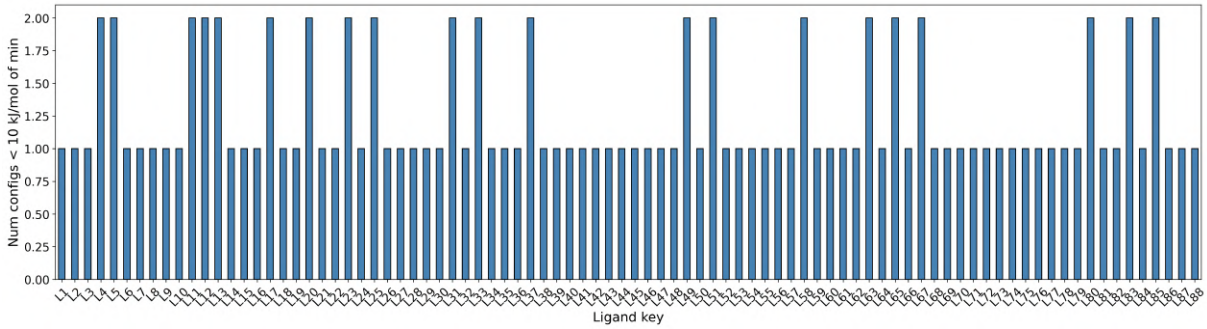


Figure 21: Ir(III) complexes

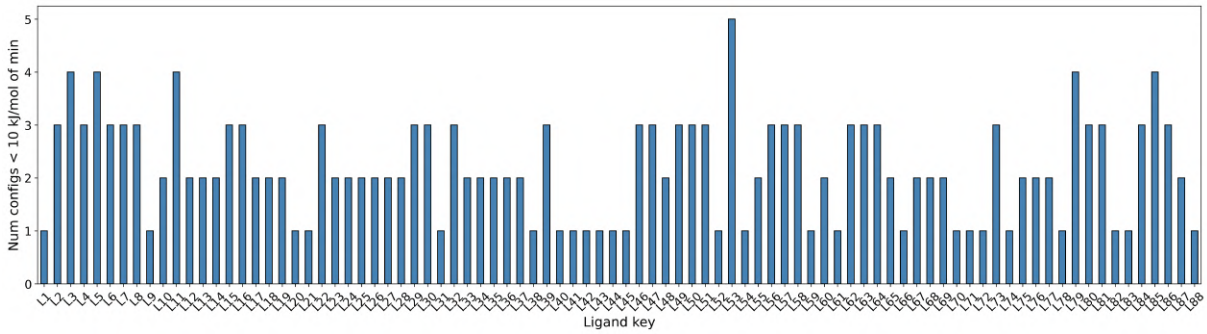


Figure 22: Ru(II) complexes

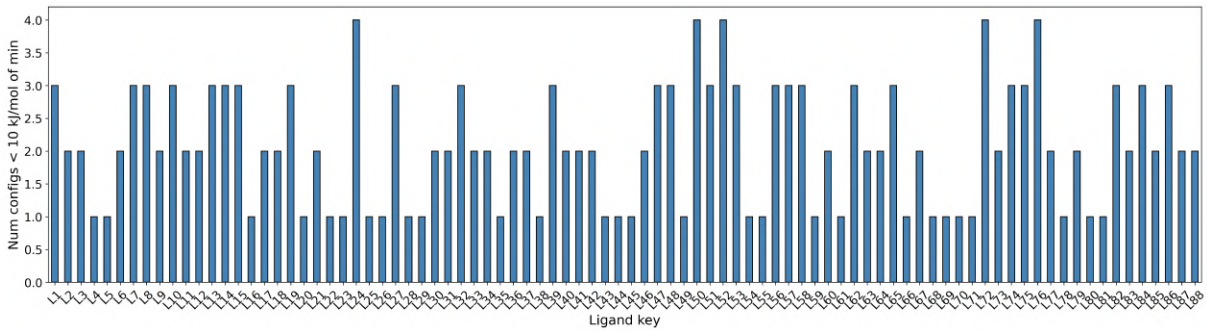


Figure 23: Mn(I) complexes