Energy Landscapes for Reaction Barriers in Metal-Organic Framework Catalysts and Non-Covalent Interactions in Polycyclic Aromatic Hydrocarbons: A Theoretical Insight

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Nikita Matsokin

Referentin: Prof. Dr. Karin Fink
Korreferent: Prof. Dr. Felix Studt
Tag der mündlichen Prüfung: 17.07.2026

Abstract

The energy landscape or potential energy surface underlies the description of any chemical system, ranging from molecules and crystals to polymers and metal clusters. This concept represents the energy of a chemical system in all possible configurations that the system can adopt. Its accurate characterization enables reliable predictions of diverse micro- and macroscopic properties. Among the most demanding cases for theoretical chemistry are catalytic energy profiles in metal–organic frameworks (MOFs) and non-covalent interactions in polycyclic aromatic hydrocarbon (PAH) dimers. Density functional theory (DFT) can be applied on extended systems but carries systematic errors, whereas coupled cluster (CC) methods deliver benchmark accuracy at prohibitive computational cost. The influence of the choice of method on reaction barriers, kinetic predictions, and relative stability of weakly bound complexes is often underestimated. I showed it through a theoretical study of two chemically distinct systems: ethylene oligomerization in a Ni-NU-1000 MOF and non-covalent dimers of PAH and their derivatives. In case of MOFs, a hierarchical cluster correction scheme based on CC substantially refined the energy profiles obtained from periodic DFT. Uncorrected periodic DFT reaction energies were found to deviate from CC-corrected energies by up to 20 kJ/mol for the adsorption and desorption steps and 15 kJ/mol for a hydrogen transfer barrier. Moreover, a microkinetic analysis revealed that the predicted product distribution depends markedly on both the level of barrier description and the catalytic environment. For PAH systems, barely investigated acene X-dimer was identified as the most stable conformer, and the origins of this stability, together with the electronic and optical properties of this and other conformers, were investigated.

These findings expose the limitations of standard methods for describing reaction energies and dimer stability. The hierarchical cluster correction scheme provides a transferable route to near-CC accuracy for periodic catalytic models at a fraction of the cost, while the dimer study provides a quantitative basis for the computational design of functional organic materials.

Kurzfassung

Die Energielandschaft oder Potentialhyperfläche bildet die Grundlage für die Beschreibung jedes chemischen Systems, von Molekülen und Kristallen bis hin zu Polymeren und Metallclustern. Die Potentialhyperfläche stellt die Energie eines chemischen Systems in allen möglichen Konfigurationen dar, die das System annehmen kann. Ihre genaue Charakterisierung ermöglicht zuverlässige Vorhersagen verschiedener mikro- und makroskopischer Eigenschaften. Zu den anspruchsvollsten Fällen für die theoretische Chemie zählen katalytische Energieprofile in metallorganischen Gerüsten (MOFs) und nichtkovalente Wechselwirkungen in Dimeren polyzyklischer aromatischer Kohlenwasserstoffe (PAH). Die Dichtefunktionaltheorie (DFT) lässt sich auch auf ausgedehnte Systeme anwerden, weist jedoch systematische Fehler auf, während Coupled-Cluster-Methoden (CC) eine Benchmark-Genauigkeit bei sehr hohen Rechenkosten liefern. Der Einfluss der Methodenwahl auf Reaktionsbarrieren, kinetische Vorhersagen und die relative Stabilität schwach gebundener Komplexe wird jedoch häufig unterschätzt. In dieser Arbeit wurde dies anhand einer theoretischen Untersuchung zweier chemisch unterschiedlicher Systeme gezeigt: der Ethylen-Oligomerisierung in einem Ni-NU-1000-MOF und nichtkovalenter Dimere von PAH und deren Derivaten. Im Fall von MOFs führte ein auf CC basierendes hierarchisches Clusterkorrekturschema zu einer erheblichen Verbesserung der aus periodischer DFT erhaltenen Energieprofile. Es zeigte sich, dass die unkorrigierten periodischen DFT-Reaktionsenergien um bis zu 20 kJ/mol für die Adsorptions- und Desorptionsschritte und um 15 kJ/mol für eine Wasserstoffübertragungsbarriere von den CC-korrigierten Energien abwichen. Darüber hinaus ergab eine mikrokinetische Analyse, dass die vorhergesagte Produktverteilung in hohem Maße sowohl vom Grad der Barrierebeschreibung als auch von der katalytischen Umgebung abhängt. Bei PAH-Systemen wurde das bislang wenig erforschte Acen-X-Dimer als stabilsters Konformer identifiziert, und die Ursachen dieser Stabilität wurden zusammen mit den elektronischen und optischen Eigenschaften dieses und anderer Konformere untersucht.

Dies Untersuchung dieser beiden Verbindungsklassen zeigen die Grenzen von DFT zur Beschreibung von Reaktionsenergien und Dimerstabilität auf. Das hierarchische Clusterkorrekturverfahren bietet einen übertragbaren Weg zu einer Genauigkeit nahe der CC-Genauigkeit für periodische katalytische Modelle zu einem Bruchteil des Aufwands, während die Dimer-Studie eine quantitative Grundlage für das computergestützte Design funktioneller organischer Materialien liefert.

Eidesstattliche Erklärung

Bei der eingereichten Dissertation zu dem Thema

Energy Landscapes for Reaction Barriers in Metal-Organic Framework Catalysts and Non-Covalent Interactions in Polycyclic Aromatic Hydrocarbons: A Theoretical Insight

handelt es sich um eine eigenständig erbrachte Leistung.

Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Karlsruhe, 05.06.2026 Nikita Matsokin

Contents

Abstract i

Kurzfassung iii

1 Introduction 1

2 Theoretical Background 5

2.1 Wave Function Theory 5

2.1.1 Coupled Cluster Theory 5

2.1.2 Complete Basis Set Extrapolation 8

2.1.3 Symmetry-Adapted Perturbation Theory 9

2.2 Density Functional Theory 10

2.2.1 The Kohn–Sham Formalism 11

2.2.2 Jacob's Ladder of Exchange–Correlation Functionals 12

2.2.3 Dispersion Corrections 13

2.2.4 Time-Dependent DFT 16

3 Ethylene Oligomerization in Ni-NU-1000 MOF Catalyst 19

3.1 Introduction 20

3.1.1 General Overview 20

3.1.2 Computational Details 25

3.2 The Hierarchical Cluster Correction Scheme for Ethylene Dimerization 29

3.2.1 The Reaction Mechanism of Ethylene Dimerization 30

3.2.2 The Importance of Correction Scheme in Catalysis 32

3.2.3 The Foundation of the Hierarchical Cluster Correction Scheme 33

3.2.4 DLPNO-CCSD(T) Correction of Ethylene Dimerization 37

3.2.5 DFT Functional Benchmark for Ethylene Dimerization 40

3.2.6 Discussion 44

3.3 Ethylene Oligomerization in Ni-NU-1000 45

3.3.1 Overview of the Catalytic Cycle and Reaction Network 46

3.3.2 Comparison of Activation Energy Barriers in the Simplified Mechanism . . . 48
3.3.3 Microkinetic Analysis of Product Distribution . . . 49
3.3.4 Discussion . . . 54
3.4 Chapter Conclusions . . . 55

4 Non-Covalent Interactions of PAHs and their Derivatives 57
4.1 Introduction . . . 57
4.1.1 General Overview . . . 57
4.1.2 Computational Details . . . 61
4.2 Non-Covalent Interactions in Acene Dimers . . . 64
4.2.1 SAPT Interaction Energy Surface Screening of Selected Pentacene Dimers . . . 64
4.2.2 The SAPT Analysis of the glike-dimer and the X-dimer Trajectory 68
4.2.3 DFT and SAPT Interaction Energies for Acene Molecules . . . 70
4.2.4 Theoretical Insights from PNO-LCCSD(T*)-F12 . . . 74
4.2.5 Comparison of Different Methods to PNO-LCCSD(T*)-F12 . . . 76
4.2.6 Electronic Structure and Optical Properties . . . 80
4.2.7 Discussion . . . 89
4.3 Chemical Substitution Effects on Anthracene Dimers . . . 90
4.3.1 Monomer Properties of Substituted Anthracenes . . . 93
4.3.2 Dimer Construction . . . 95
4.3.3 Dimer Interaction Energies . . . 96
4.3.4 Optical Properties of Substituted Dimers . . . 101
4.3.5 Discussion . . . 103
4.4 Extended Linear Acenes . . . 103
4.4.1 Interaction Energy Trends . . . 105
4.4.2 Electronic Structure Evolution in Extended Acenes . . . 107
4.4.3 Discussion . . . 110
4.5 *cata*-Condensed Polybenzoid Isomers . . . 111
4.5.1 Construction of the COMPAD Project Database . . . 112
4.5.2 Monomer and Dimer Electronic Properties . . . 113
4.5.3 Interaction Energy Correlation with BSA and Isotropic Polarizability . . . 117
4.5.4 Discussion . . . 120
4.6 Chapter Conclusions . . . 120

5 Summary and Conclusion **123**

A Appendix **127**
A.1 Appendix for Chapter 3, Section 3.2 127
A.2 Appendix for Chapter 3, Section 3.3 128
A.3 Appendix for Chapter 4, Section 4.2 131
A.4 Appendix for Chapter 4, Section 4.3 132
A.5 Appendix for Chapter 4, Section 4.4 133
A.6 Appendix for Chapter 4, Section 4.5 134

Acronyms and Abbreviations **141**

Bibliography **145**

Permissions to Print **173**

List of Publications **175**

Acknowledgments **179**

Declaration **183**

1 Introduction

The energy landscape (Figure 1) or potential energy surface (PES) is one of the central concepts in the natural sciences and beyond, representing a mathematical mapping of energy as a function of the system through all possible configurations[1, 2]. For example, these configurations can be represented as atomic coordinates[3]. Of particular interest are stationary points at which the gradient of the energy function vanishes[4]. Analysis of such landscapes reveals fundamental properties of the system, as well as characteristic patterns governing its behavior and evolution[2, 5].

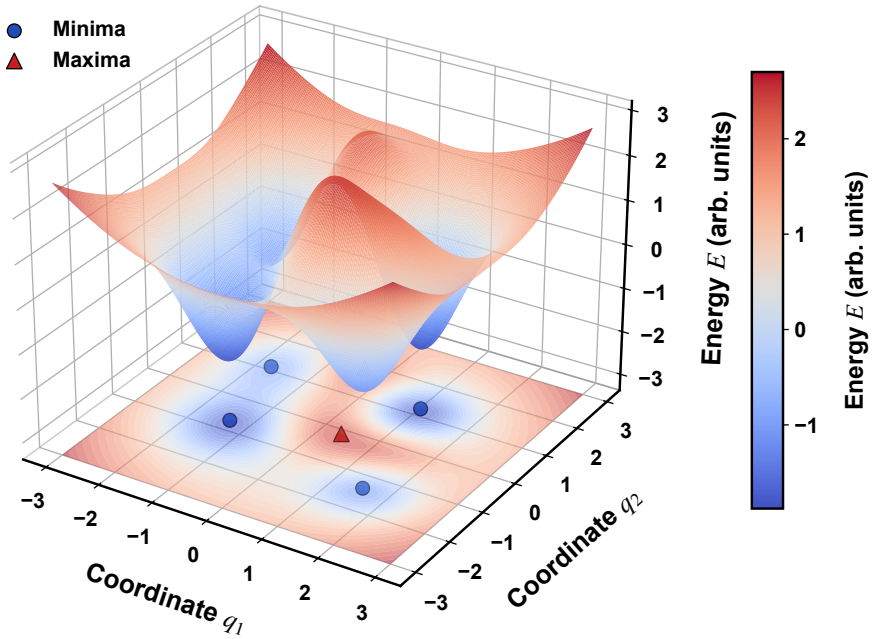


Figure 1: The example of the energy landscape with two coordinates q_1 and q_2. This picture represents minima and maximum for an abstract chemical system.

In chemistry, the most significant features of the energy landscape are the minima and the pathways that connect them via saddle points. These topological features of the landscape correspond to well-defined states of chemical systems, such as molecules[6,

7], crystals[1, 8], atomic clusters[4] and biomolecular complexes[9, 10]. Chemical reactions can also be described in terms of energy landscapes or PES projected onto a constrained set of degrees of freedom[11–13]. The residence of a system at a given point may be characterized in terms of its stability.

Transitions between minima are possible, but require activation: the system must acquire a sufficient energy to reach the saddle point that separates two minima[6]. Within the framework of transition state theory (TST), this saddle point defines the activation energy barrier that the system must overcome through the transition state (TS) to reach an adjacent minimum[14].

Two particularly important fields in which the well-described energy landscape plays a crucial role are catalysis[15, 16] and non-covalent complexes[17–21]. In the theoretical description of catalysis, the accurate characterization of minima and TSs is of direct importance, as these features are intimately connected to macroscopic observables measured for a reaction[22, 23]. An incorrectly determined activation energy barrier, for example, will lead to a misinterpretation of the catalytic mechanism and the associated reaction pathways[24]. For instance, changes in the TS energy may have significant influences on the product distribution.

In non-covalent complexes, the energy differences between minima can be relatively small, as are the barriers separating them, which complicates the theoretical rationalization of relative stability and conformational interconversion[17, 21, 25, 26]. The present thesis is dedicated to the theoretical investigation of both phenomena through specific case studies: ethylene oligomerization in a Ni-NU-1000 metal-organic framework (MOF) catalyst and non-covalent interactions in acene dimers and their derivatives.

In the case of MOFs, many experimental and theoretical studies on Ni-NU-1000 exist for application in ethylene oligomerization[27–29]. Consequently, both the reaction pathway and the catalyst itself are well characterized. However, as demonstrated for zeolites[30, 31], conventional studies employ a truncated cluster model and limited computational methods to simulate the catalyst, which restricts the predictive power of the modeling results. In the case of Ni-NU-1000 in particular, this can affect the reaction kinetics in the transition from energy barriers to rate constants[24]. For non-covalent acene dimers, various conformations have been studied in the literature[25, 32–34]. However, some of these investigations remained unsystematic for some conformers. Additionally, data on the influence of both the monomer structure and the dimer conformation on electronic and optical properties were largely absent.

The thesis is organized as follows. In Chapter 2, the theoretical background of the quantum chemical methods employed throughout this work is presented. In Chapter 3, the computational modeling of ethylene oligomerization is described for the Ni-NU-1000 MOF catalyst. In Chapter 4, non-covalent dimers for acene molecules and their derivatives are investigated through their stability, electronic and optical properties. Within these studies, I will present four key results:

1. A correction scheme was developed for MOF catalysts to accurately describe reaction intermediates and transition states in ethylene oligomerization, achieving near coupled-cluster accuracy while accounting for long-range interactions.
2. The influence of energy corrections on product distribution was investigated via microkinetic modeling.
3. The relative stability of acene conformers and their influence on electronic and optical properties were studied.
4. The influence of size, substitution, and shape on key dimer properties was characterized, including interaction energy, and optical and electronic properties.

2 Theoretical Background

Overview of Chapter

In this chapter a brief introduction to the underlying theoretical principles of used methods is provided. The chapter is divided into two main parts. In Section 2.1 selected wave function theory methods are presented, including coupled cluster (CC) and its local correlation modifications, complete basis set (CBS) extrapolation techniques, and symmetry-adapted perturbation theory. In Section 2.2, the fundamentals of density functional theory (DFT), its different levels of approximation, dispersion corrections, and finally an extension to excited states through the time-dependent DFT (TD-DFT) approach are described.

2.1 Wave Function Theory

2.1.1 Coupled Cluster Theory

The coupled cluster approach[35, 36] belongs to the family of post-Hartree-Fock methods. In Hartree-Fock (HF) theory, the wave function is represented by a single Slater determinant, that is a model where each electron interacts only with the mean field generated by all other electrons. The Slater determinant takes the following form:

$$\Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{x}_1) & \varphi_2(\mathbf{x}_1) & \cdots & \varphi_N(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) & \varphi_2(\mathbf{x}_2) & \cdots & \varphi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{x}_N) & \varphi_2(\mathbf{x}_N) & \cdots & \varphi_N(\mathbf{x}_N) \end{vmatrix}, \tag{1}$$

where $\Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ is the Hartree-Fock wave function, $\mathbf{x}_i = (\mathbf{r_i}, s_i)$ is the spatial and spin coordinate of the electron i, $\varphi_j(\mathbf{x}_i)$ is a one-electron spin orbital and N is the number of electrons.

Based on the variational principle, HF energy (E_{HF}) in an infinite basis set (complete basis set limit) will be higher than the exact energy (E_{exact}):

$$E_{HF} = \frac{\langle \Psi_{HF} \mid \hat{H} \mid \Psi_{HF} \rangle}{\langle \Psi_{HF} \mid \Psi_{HF} \rangle} \geq E_{exact}. \tag{2}$$

It means that the HF approach will have an underestimation in energy, which is called correlation energy:

$$E_{corr} = E_{exact} - E_{HF}. \tag{3}$$

There are different approaches to estimate the correlation energy, such as the Møller-Plesset perturbation theory (typically second-order perturbation theory or MP2) or the full configuration interaction (CI) method[37]. If the MP2 method has not enough accuracy and the full CI method is too demanding, there is also something in between, and it is called the coupled cluster approach.

In the CC approach, the wave function has the following form:

$$\Psi_{\mathrm{CC}} = e^{\hat{T}} \Psi_{\mathrm{HF}}, \tag{4}$$

where $e^{\hat{T}}$ is the exponential of the cluster operator $\hat{T}$:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_n, \tag{5}$$

where n is the total number of electrons in the system and each $\hat{T}_i$ generates all possible excitations to the Ψ_{HF} reference. For example, the $\hat{T}_2$ operator describes double excitations:

$$\hat{T}_2 \Psi_{\mathrm{HF}} = \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{vir}} t_{ij}^{ab} \, \Psi_{ij}^{ab}, \tag{6}$$

where t_{ij}^{ab} is the double-excitation amplitude, where occupied orbitals i, j are promoted to virtual orbitals a, b The expansion of $\hat{T}$ in the exponential format is defined as:

$$e^{\hat{T}} = \hat{I} + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2!}\hat{T}_1^2) + (\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{3!}\hat{T}_1^3) + \cdots, \tag{7}$$

Expanding the exponential operator yields contributions from single, double, triple, and higher excitations, allowing the correlation energy to be systematically improved

by retaining successively higher-order terms. This gives rise to a hierarchy of coupled cluster approximations:

$\hat{T} = \hat{T}_1$	CC singles (CCS)
$\hat{T} = \hat{T}_2$	CC doubles (CCD)
$\hat{T} = \hat{T}_1 + \hat{T}_2$	CC singles and doubles (CCSD)
$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$	CC singles, doubles, and triples (CCSDT)
$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4$	CC singles, doubles, triples, and quadruples (CCSDTQ)
...	

The typical choice between cost and accuracy is the CCSD(T) method with a scaling of $\mathcal{O}(M^7)$[38]. This method can be obtained from CCSD with adding triple excitations via perturbation theory. This perturbation expansion yields triple amplitude corrections to the energy in 4th and 5th order, which are then recalculated using converged CCSD amplitudes[39]:

$$E_{\mathrm{CCSD(T)}} = E_{\mathrm{CCSD}} + E^{(4)} + E^{(5)}, \tag{8}$$

$$E^{(4)} = \frac{1}{36} \sum_{ijkabc} t_{ijk}^{abc*} \, D_{ijk}^{abc} \, t_{ijk}^{abc}, \tag{9}$$

where D_{ijk}^{abc} is the energy denominator for the triple excitation.

$$E^{(5)} = \frac{1}{4} \sum_{ijkabc} \langle jk \| bc \rangle \, t_i^{a*} \, t_{ijk}^{abc}. \tag{10}$$

The CCSD(T) method becomes the most common realization of the CC theory and is widely used in benchmarks as the reference[40–44]. It is applied in many fields, such as catalysis[45], atmosphere chemistry[46], non-covalent interactions[42, 43] and many others[47–49].

Unfortunately, the steep scaling of CCSD(T) precludes its routine application to systems with more than 20-30 atoms without prohibitive computational effort[50]. However, there are approaches to circumvent this limitation by employing various local correlation schemes based on the ideas of occupied orbital localization and domain construction[51–56]. In these approaches, the virtual space is spanned by projected atomic orbitals (PAOs), and all excitations from localized occupied orbitals

to PAOs are restricted to the same domain. A further development of the PAO-based approach was the introduction of pair natural orbitals (PNOs), defined as linear combinations of PAOs that represent the correlating virtual space orbitals associated with a specific electron pair[57]. Building on this idea, several widely used methods have been developed, notably pair natural orbital local CCSD(T) (PNO-LCCSD(T)[58, 59]) and domain-based local pair natural orbital CCSD(T) (DLPNO-CCSD(T)[60, 61]). The differences between these two methods lie primarily in the domain and pair approximations[62]. First, they employ different algorithms for the construction of PAO orbital domains. In PNO-LCCSD(T), the same PAO orbital domains are used throughout the LMP2, LCCSD, and triples calculations, with the domain limits determined by a threshold applied to the localized molecular orbitals (LMOs), whereas in DLPNO-CCSD(T) the PAO orbital domains are selected on the basis of a differential overlap criterion with a specific threshold. Second, the criteria for constructing PNO domains differ: the PNO-LCCSD(T) approach employs both occupation number and energy completeness criteria, while DLPNO-CCSD(T) relies solely on an occupation number threshold. Finally, an important distinction concerns how the two methods treat different pair categories, such as distant and weak pairs. For example, PNO-LCCSD(T) employs a higher-order multipole expansion ($p = 3$) for approximating distant pair energies, while DLPNO-CCSD(T) uses only a lower-order expansion ($p = 1$), which reduces computational cost but introduces larger errors at long-range pairs[62]. Consequently, the PNO-LCCSD(T) method provides a more accurate treatment of medium- and long-range correlation effects than DLPNO-CCSD(T), which is especially important for non-covalent interactions[62, 63].

Finally, it is essential to note that the thresholds for domain and pair selection must be sufficiently low to avoid significant loss of correlation energy relative to canonical CCSD(T)[64]. For example, the "TightPNO" truncation criteria in ORCA software are designed to retain 99.9% of the canonical CCSD(T) correlation energy in some cases[65].

2.1.2 Complete Basis Set Extrapolation

A common challenge in WFT methods is the basis set incompleteness error (BSIE). Although it can in principle be reduced by increasing the basis set size, this approach quickly becomes impractical due to the increase in computational cost associated with larger basis sets. This issue is particularly critical for the CCSD(T) method, which scales as $\mathcal{O}(M^7)$. It is therefore standard practice to extrapolate the results

obtained with a sequence of finite basis sets toward the complete basis set limit in order to mitigate the BSIE. This procedure is known as CBS extrapolation and is routinely applied to both energy components in WFT: the Hartree-Fock energy and the correlation energy. In ORCA 5.0.4 quantum chemical package[66], different extrapolation schemes are available for each of these two components. To estimate the Hartree-Fock energy at the CBS limit, the two-point exponential extrapolation scheme of Zhong *et al.*[67] was used, using basis-set-specific exponent coefficients[68]. The extrapolation formula equals:

$$E_{\mathrm{HF}}(\mathrm{CBS}) = \frac{E_{\mathrm{HF}}(X) \cdot e^{-\alpha\sqrt{Y}} - E_{\mathrm{HF}}(Y) \cdot e^{-\alpha\sqrt{X}}}{e^{-\alpha\sqrt{Y}} - e^{-\alpha\sqrt{X}}}, \tag{11}$$

where $E_{\mathrm{HF}}(X)$ and $E_{\mathrm{HF}}(Y)$ are the Hartree-Fock energies computed with basis sets X and Y (where X and Y also denote the cardinal numbers of the respective basis sets), α is the exponent coefficient, and A is the pre-exponential coefficient.

Additionally, the correlation energy can be extrapolated using a two-point scheme[69] with basis set-specific coefficients[68], based on the expression[70]:

$$E_{\mathrm{corr}}(\mathrm{CBS}) = \frac{X^{\beta} \cdot E_{\mathrm{corr}}(X) - Y^{\beta} \cdot E_{\mathrm{corr}}(Y)}{X^{\beta} - Y^{\beta}}, \tag{12}$$

where $E_{\mathrm{corr}}(X)$ and $E_{\mathrm{corr}}(Y)$ are the correlation energies computed with basis sets X and Y, β is the exponent coefficient, and B is the pre-exponential coefficient. For example, when extrapolating with the def2-TZVPP and def2-QZVPP basis sets[71], the coefficients take the values $\alpha = 7.88$ and $\beta = 2.97$[68].

2.1.3 Symmetry-Adapted Perturbation Theory

In the case of a non-covalent dimer system AB, for each monomer (A and B) the Schrödinger equation can be written ($X = A, B$)[72]:

$$H_X \Psi_X = E_X \Psi_X. \tag{13}$$

The total Hamiltonian of the dimer system can be formulated as:

$$H = H_A + H_B + V. \tag{14}$$

That makes a problem when looking at direct wave function product of monomers ($\Psi = \Psi_A \Psi_B$), one can find the lack of antisymmetry with respect to electron exchange between monomers, that results in the neglect of Pauli repulsion effects[73]. Therefore, there is a need to restore the correct symmetry, for example, through the symmetrized Rayleigh-Schrödinger (SRS) perturbation theory[74]. The reformulation of SAPT in SRS form is applicable in the interpretation of contributions in non-covalent interaction. For example, using second-order perturbation in V, the interaction energy can be reformulated as a sum of electrostatic (elst), induction (ind), dispersion (disp), and exchange-repulsion (exr) contributions[73]:

$$E_{int} = E_{elst}^{(1)} + E_{exr}^{(1)} + E_{ind}^{(2)} + E_{exr-ind}^{(2)} + E_{disp}^{(2)} + E_{exr-disp}^{(2)}. \tag{15}$$

In practice, SAPT is built on Hartree-Fock approximation or density functional theory. For example, sSAPT0 that is used in this thesis based on Hartree-Fock level without intramolecular correlation effect and has the following formula[75–77]:

$$E_{int}^{\text{sSAPT0}} = \underbrace{E_{elst}^{(10)}}_{E_{elst}} + \underbrace{E_{exr}^{(10)}}_{E_{exr}} + \underbrace{E_{ind,r}^{(20)} + p_{\text{EX}}(\alpha) E_{exr-ind,r}^{(20)} + \delta E_{HF}^{(2)}}_{E_{ind}} + \underbrace{E_{disp}^{(20)} + p_{\text{EX}}(\alpha) E_{exr-disp}^{(20)}}_{E_{disp}}, \tag{16}$$

where E_{exr} is exchange-repulsion contribution, $p_{\text{EX}}(\alpha)$ is the scale factor with empirical $\alpha = 3.0$ parameter and $\delta E_{HF}^{(2)}$ is the Hartree-Fock correction. In $E^{(20)}$ the first digit denotes intermolecular order and the second denotes intramolecular order of perturbation. This scale factor is calculated using the following formula:

$$p_{\text{EX}}(\alpha) = \left(\frac{E_{exr}^{(10)}}{E_{exr}^{(10)}(S^2)} \right)^{\alpha}, \tag{17}$$

where S is the orbital overlap. The Hartree-Fock correction can be found as:

$$\delta E_{HF}^{(2)} = E_{int}^{\text{HF}} - \left(E_{elst}^{(10)} + E_{exr}^{(10)} + E_{ind,r}^{(20)} + E_{exr-ind,r}^{(20)} \right). \tag{18}$$

2.2 Density Functional Theory

Density functional theory is one of the most widely used computational approaches in theoretical chemistry and condensed-matter physics for describing the electronic structure of many-body quantum systems[78]. Its formal foundation is rested on the two Hohenberg–Kohn theorems[79], which establish that the ground-state energy of

an interacting many-electron system is a unique functional of the electron density $\rho(\mathbf{r})$, and that the true ground-state density minimizes this functional. Although these theorems showed the existence of such a universal functional without explicit form for it. The Kohn-Sham (KS) formalism[80] transformed DFT into a practical computational tool implemented in modern quantum-chemistry packages[81].

2.2.1 The Kohn–Sham Formalism

The central idea of the Kohn–Sham formalism[80] is the introduction of an auxiliary system of non-interacting electrons (N) whose collective behavior reproduces exactly the same ground-state electron density ($\rho(\mathbf{r})$) as that of the fully interacting many-electron system:

$$\rho(\mathbf{r}) \;=\; \sum_{i=1}^{N} \big|\psi_i(\mathbf{r})\big|^2. \tag{19}$$

The electron density is constructed directly from the N occupied one electron KS-orbitals ($\psi_i(\mathbf{r})$). In Kohn-Sham formalism, kinetic energy is computed exactly for the non-interacting reference system using KS-orbitals, while many-body contributions (except of $J[\rho]$) implemented into the exchange-correlation functional (XC). The total KS energy functional is formulated as:

$$E[\rho] \;=\; T_s[\rho] \;+\; E_{Ne}[\rho] \;+\; J[\rho] \;+\; E_{XC}[\rho], \tag{20}$$

where $T_s[\rho]$ is the kinetic energy of the non-interacting reference system, $E_{Ne}[\rho]$ is the electron–nuclear attraction energy, $J[\rho]$ is Hartree electron-electron repulsion, and $E_{XC}[\rho]$ is the XC energy.

If the variational principle is applied (Eq. (2)) to the KS energy(Eq. (20)), a set of effective single-particle equations appears (KS equations):

$$\hat{h}_{\mathrm{KS}}\,\psi_i(\mathbf{r}) \;=\; \left[-\frac{1}{2}\nabla^2 + v_{Ne}(\mathbf{r}) + v_J(\mathbf{r}) + v_{XC}(\mathbf{r})\right]\psi_i(\mathbf{r}) \;=\; \varepsilon_i\,\psi_i(\mathbf{r}), \tag{21}$$

where $v_{Ne}(\mathbf{r})$ is the external nuclear potential, $v_J(\mathbf{r})$ is Hartree electron-electron repulsion, and $v_{XC}(\mathbf{r}) = \delta E_{XC}/\delta\rho(\mathbf{r})$ is the XC potential. The challenge lies in the unknown exact XC functional, making available only different approximations of $E_{XC}[\rho]$.

All effective potentials in Eq. (21) depend on the density or consequently on KS orbitals. Therefore, KS equations must be solved in a self-consistent field cycle to update iteratively orbitals and density by diagonalization until convergence.

2.2.2 Jacob's Ladder of Exchange–Correlation Functionals

As mentioned above, the exact form of $E_{XC}[\rho]$ remains unknown. Consequently, various approximations have been developed for the evaluation of the exchange-correlation energy, which were subsequently classified with means of Jacob's Ladder by Perdew and Schmidt[82]. This classification corresponds to incorporating progressively more complex and computationally demanding ingredients into the functional with each ascending class, with the expectation of converging toward the "heaven" of chemical accuracy (1 kcal/mol). However, this does not imply that a higher-class functional will necessarily yield greater accuracy in practice compared to a lower-class functional, as has been observed in recent years with the development of new functionals[83]. The four selected classes of Jacob's Ladder are summarized in Table 1.

Table 1: Jacob's Ladder of DFT exchange-correlation approximations[82]. $\tau(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla\psi_i|^2$ denotes the KS kinetic energy density.

Class	**Functional Components**	**Examples**
LDA / LSDA	$\rho(\mathbf{r})$	VWN[84], PW92[85]
GGA	$\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$	PBE[86]
meta-GGA	$\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$, $\tau(\mathbf{r})$	TPSS[87]
Hybrid	GGA/*meta*-GGA + exact HF exchange	B3LYP[88], TPSSh[89]

Local Density Approximation (LDA/LSDA)

The basic idea of the simplest approximation is to assume the homogeneity of the density everywhere. So, it means that the XC energy can be determined as a uniform electron gas (UEG) with the same local density[84, 85]:

$$E_{XC}^{\mathrm{LDA}}[\rho] \;=\; \int \rho(\mathbf{r})\, \varepsilon_{XC}^{\mathrm{UEG}}\big[\rho(\mathbf{r})\big]\, d\mathbf{r}, \tag{22}$$

where $\varepsilon_{XC}^{\mathrm{UEG}}$ is the XC energy per particle of the UEG with the density $\rho(\mathbf{r})$. LDA/LSDA can provide a quite accurate exchange-correlation function, especially for solid-state systems[90]. However, it systematically has inaccurate values of the XC energy for molecules[91].

Generalized Gradient Approximation (GGA)

GGA functionals modify LDA by including the density gradient ($\nabla\rho(\mathbf{r})$), therefore capturing the inhomogeneity of real electron densities[86]:

$$E_{XC}^{\mathrm{GGA}}[\rho] \;=\; \int \rho(\mathbf{r})\, \varepsilon_{XC}\big[\rho(\mathbf{r}),\, \nabla\rho(\mathbf{r})\big]\, d\mathbf{r} \tag{23}$$

The parameter-free PBE functional[86] is among the most widely used GGAs in both chemistry and materials science[92].

meta-GGA

meta-GGA functionals extend and improve GGA functional by additionally incorporating the KS kinetic energy density ($\tau(\mathbf{r})$):

$$\tau(\mathbf{r}) \;=\; \frac{1}{2}\sum_i \big|\nabla\psi_i(\mathbf{r})\big|^2, \tag{24}$$

and in some cases the Laplacian $\nabla^2\rho(\mathbf{r})$. This addition to GGA allows *meta*-GGA functionals[93] describe different bonding types (covalent, metallic and weak bonds). The TPSS functional[87] was one of the first *meta*-GGA.

Hybrid Functionals

Hybrid functionals are based on the idea to mix a fraction of non-local exact exchange from the Hartree-Fock method and semi-local GGA or *meta*-GGA functionals[88]:

$$E_{XC}^{\mathrm{hybrid}} \;=\; a\, E_X^{\mathrm{exact}} + (1-a)\, E_X^{\mathrm{GGA}} + E_C^{\mathrm{GGA}}, \tag{25}$$

where a is the mixing coefficient.This inclusion of Hartree-Fock exact exchange can partially compensate the self-interaction error (SIE) included in pure semi-local DFT functionals, that can improve the performance for thermochemistry, reaction barriers and others[94–96].

2.2.3 Dispersion Corrections

Another problem of DFT is the lack of long-range dispersion interactions description[97], while it is very important for modeling the most of the systems in chemistry, physics and biology[98]. For this reason different methods, mostly incorporated with empirical components are used to describe this type of the systems. Grimme suggested to classified dispersion corrected method it in the following three groups: density-based,

C_6-based and effective one-electron potential methods[97]. Currently, the most popular methods are C_6-based or DFT-D corrections, when the full energy of the system (E_{full}) can be obtained by:

$$E_{\text{full}} = E_{\text{DFT}} + E_{\text{disp}}, \tag{26}$$

where E_{DFT} and E_{disp} are the DFT energy and the dispersion correction energy. This approach is called semi-classical because of atom pairwise additive treatment of the dispersion energy that contains semi-classical isotropic nth-order dispersion coefficient (C_n).

The general view of the E_{disp} is the following one:

$$E_{\text{disp}} = -\sum_{AB} \sum_{n=6,8,10,\ldots} \frac{s_n C_n^{AB}}{R_{AB}^n} f_{\text{damp}}(R_{AB}), \tag{27}$$

where $f_{\text{damp}}(R_{AB})$ indicates the damping function (standard D2 and D3 use zero-damping) and s_n are scaling constants fitted to specific functionals, AB is atom pair and R_{AB} is the interatomic distance.

D2 (DFT-D2)

The D2 dispersion energy is a purely pairwise sum over all pairs of atoms A, B[99]:

$$E_{\text{disp}}^{\text{D2}} = -\sum_{A<B} \frac{s_6 C_6^{AB}}{R_{AB}^6} \cdot f_{\text{damp}}(R_{AB}), \tag{28}$$

where $C_6^{AB} = \sqrt{C_6^A \cdot C_6^B}$ (geometric mean of tabulated coefficients), and the zero-damping function is:

$$f_{\text{damp}}(R_{AB}) = \frac{1}{1 + e^{-d(R_{AB}/R_r - 1)}}, \tag{29}$$

where d is an empirical parameter and R_r is a sum of atomic van der Waals radii. The importance of damping function to prevent divergence at short distance and avoid double counting with the DFT correlation. In D2 C_6 are fixed atomic coefficients, making the same values for two atoms in a different environment.

D3 (DFT-D3)

The D3 correction models the dispersion energy as an atom-pairwise sum using geometry-dependent C_6 and C_8 coefficients[100]:

$$E_{\text{disp}}^{\text{D3}} = -\sum_{A,B}\sum_{n=6,8}\frac{s_n C_n^{AB}}{R_{AB}^n} f_{\text{damp}}(R_{AB}), \tag{30}$$

In D3 (and also in D4) it is possible to add a three-body Axilrod–Teller–Muto (ATM) correction $E^{(3)}$[101, 102]. While the C_8/R^8 term captures dipole–quadrupole dispersion, the three-body $E^{(3)}$ term accounts for non-additive triple interactions via:

$$E^{(3)} = \sum_{A<B<C}\frac{s_9 C_9^{ABC}(3\cos\theta_a\cos\theta_b\cos\theta_c + 1)}{(r_{AB}r_{BC}r_{CA})^3}. \tag{31}$$

The zero-damping function is also different to D2 and now tends to zero as $R \to 0$:

$$f_{d,n}(R_{AB}) = \frac{1}{1 + 6(R_{AB}/(s_{r,n}R_0^{AB}))^{-\alpha_n}}, \tag{32}$$

where $s_{r,n}$ is the order-dependent scaling factor, R_0^{AB} is the cutoff radii and α_n is the "steepness" parameter.

What is important especially for D3 that C_6^{AB} and C_8^{AB} coefficients are dependent on geometry and coordination number, making the difference, for example, in the case of aliphatic and aromatic carbon atoms.

D3(BJ) (DFT-D3 with Becke-Johnson Damping Function)

D3(BJ) dispersion correction[103] modifies the damping function to the Becke-Johnson[104, 105] (BJ) form for improved short-range accuracy (makes the dispersion correction finite and non-zero at $R \to 0$):

$$f_{d,n}(R_{AB}) = \frac{R_{AB}^n}{R_{AB}^n + (a_1 R_0^{AB} + a_2)^n} \tag{33}$$

The BJ damping has two additional parameters (a_1, a_2) which are empirically specified for each DFT functional, increasing flexibility and transferability. So, the D3 form of expression remained the same with the BJ damping substitution.

D4 (DFT-D4)

In D4 dispersion correction, the atomic charge-dependent dispersion coefficients are included[106]. For example, the C_6^{AB} coefficient is calculated through atomic polarizabilities (α_A and α_B) that depends on atomic charges (q_A and q_B):

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty \alpha_A(i\omega, q_A)\, \alpha_B(i\omega, q_B)\, d\omega, \tag{34}$$

where $i\omega$ is an imaginary frequency. So, now dispersion coefficients depend on both geometry and atomic charges, which improves the description of non-covalent interactions, especially in ionized or metal-organic systems.

D4 incorporates atomic charge dependent dispersion coefficients and includes a default three-body Axilrod-Teller-Muto term:

$$E_{\text{disp}}^{\text{D4}} = E_{\text{pairwise}} + E^{(3)}. \tag{35}$$

Here, the pairwise coefficients C_n depend on both geometry and atomic partial charges, improving transferability and accuracy, especially in ionized or metal-organic systems.

2.2.4 Time-Dependent DFT

If there is need to model excited states in molecules and solids, it is not possible to do with standard Kohn-Sham DFT that can treat a system only in the ground-state. Therefore, the extension of Kohn-Sham DFT was developed and is called time-dependent DFT or TD-DFT[107, 108]. As was the foundation of DFT in the Hohenberg-Kohn theoremes, originally TD-DFT also had the basis known as the Runge-Gross theoremes[109]. Time-dependent DFT has the same Kohn-Sham formalism and final set of effective single-particle equations has time-dependence for each term take the form:

$$\hat{h}_{\text{KS}}\, \psi_i(\mathbf{r}, t) \;=\; \left[-\frac{1}{2}\nabla^2 + v_{Ne}(\mathbf{r}, t) + v_J(\mathbf{r}, t) + v_{XC}(\mathbf{r}, t) \right] \psi_i(\mathbf{r}, t) \;=\; i\frac{\partial}{\partial t}\psi_i(\mathbf{r}, t). \tag{36}$$

In the end time-dependent electron density is reconstructed from the N occupied KS orbitals:

$$\rho(\mathbf{r},t) = \sum_{i=1}^{N} |\psi_i(\mathbf{r},t)|^2. \tag{37}$$

Then it is possible to use linear response formalism of TD-DFT that reduces Kohn-Sham equations to the Casida equations[107, 110]:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \omega \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}, \tag{38}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are the excitation and de-excitation amplitude vectors, ω is the corresponding excitation energy, and the matrices ($\mathbf{A}$,$\mathbf{B}$) are:

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) \ + (ia|jb) + (ia|f_{\mathrm{XC}}|jb), \tag{39}$$

$$B_{ia,jb} = (ia|bj) + (ia|f_{\mathrm{XC}}|bj). \tag{40}$$

Here i, j index occupied and a, b virtual Kohn-Sham orbitals, $(ia|jb)$ are two-electron Coulomb integrals, and $f_{\mathrm{XC}}(\mathbf{r})$ is the exchange-correlation kernel:

$$f_{\mathrm{XC}}(\mathbf{r}) = \frac{\delta v_{\mathrm{XC}}(\mathbf{r})}{\delta\rho(\mathbf{r})}. \tag{41}$$

3 Ethylene Oligomerization in Ni-NU-1000 MOF Catalyst

Overview of Chapter

In this chapter, a hierarchical cluster correction scheme for energies of reactions, based on the DLPNO-CCSD(T) method, was developed and applied to the kinetic investigation of the MOF-based Ni-grafted NU-1000 catalyst (Ni-NU-1000) for ethylene oligomerization. In Section 3.1.1 an introduction to the topic of ethylene oligomerization are provided, covering known mechanisms and catalysts, the role of MOF-supported catalysts within this landscape, and the current state of computational modeling of such catalysts with respect to the thermodynamics and kinetics of the reaction mechanism. The objectives and challenges to be addressed throughout this chapter are also stated. In Section 3.1.2 computational details of all calculations within this chapter are described. In Section 3.2 the DLPNO-CCSD(T)-based hierarchical cluster correction scheme for the ethylene dimerization process is presented, followed by a systematic benchmark of various DFT functionals and MP2 against the DLPNO-CCSD(T) reference. These results were obtained in collaboration with co-first author Thanh-Nam Huynh, who developed the original reaction pathway for ethylene dimerization and conducted initial tests of various active sites and cluster models for reaction energies[111] (Publication 3, p. 175). In Section 3.3 the implementation of the hierarchical cluster correction scheme for the complete ethylene oligomerization mechanism is described. Results are presented for corrected and uncorrected DFT approaches. Subsequently, within the framework of a simplified mechanism, the three methods and two pore types (trigonal and hexagonal) are compared on the basis of the computed energy barriers. The microkinetic analysis was then performed using these energy barriers. The evolution of concentration profiles of all products and the final product distributions as results are discussed. Furthermore, the microkinetic results are discussed in the context of available experimental selectivity data for Ni-NU-1000. The complete and simplified reaction mechanisms for olefin oligomerization and the

kinetic workflow were developed and published in co-authorship with Alexander Avdoshin and Thanh-Nam Huynh[112] (Publication 2, p. 175). Unless stated otherwise, the construction of the hierarchical correction scheme, the DFT benchmarking, the application of the hierarchical cluster correction scheme to the complete oligomerization mechanism, and the execution of the microkinetic simulations presented in this chapter were performed by the author of this thesis.

3.1 Introduction

3.1.1 General Overview

Ethylene Oligomerization

The oligomerization of ethylene into linear α-olefins (LAOs) occupies a central position in the petrochemical industry[113, 114]. LAOs serve as co-monomers in the production of linear low-density polyethylene, as well as feedstocks for detergents, plasticizers, and synthetic lubricants[115–118]. The global demand for these products continues to grow, placing significant pressure on developing selective and efficient catalytic processes[119, 120].

Historically, the industrial production of LAOs has relied on homogeneous transition-metal catalysts[113]. Prominent examples include the Shell Higher Olefin Process, employing Ni-based phosphine-oxide catalysts, and the Cr-based Phillips trimerization and Sasol/Linde processes[114, 121, 122]. In these systems, the key mechanistic pathway governing oligomer formation is the Cossee-Arlman mechanism, where ethylene coordinates to the metal center, undergoes migratory insertion into the metal-carbon bond, and the chain is terminated by β-hydride elimination, releasing the oligomeric product and regenerating the active metal hydride[27]. The competition between further insertion (chain growth) and β-hydride elimination (chain termination) determines the distribution of oligomeric products.

Metal-Organic Framework Catalysts

Although homogeneous catalysis is widely used in various applications, there are still practical challenges present such as catalyst separation and its recycling. The desire for confinement-driven selectivity has motivated the transition toward heterogeneous single-atom catalysts (SACs) and, in particular, toward MOF-supported catalytic systems[123–125]. The deposition of catalytic metal centers within a well-defined porous environment offers not only the predictable catalytic activity of heterogeneous

systems but also the site isolation and structural tunability, characteristic of molecular catalysts.

MOFs are a class of crystalline porous materials formed by the self-assembly of inorganic metal nodes (secondary building units) and organic linkers[126, 127]. Their well-defined porous structures provide a uniquely characterized environment for catalysis. Each metal site exists in an identical coordination environment, conferring a SAC character that is difficult to achieve on porous and mixed composition supports, such as zeolites[128, 129]. Moreover, the modular nature of MOF chemistry, in which linker length, functionality, and metal identity can be systematically varied, allowing rational design of the catalytic environment and pore geometry[130].

Among zirconia-based MOFs, NU-1000 and UiO-66 have emerged as the most widely investigated supports for ethylene oligomerization catalysts[27, 131–133]. Their thermal and chemical stability, combined with large pore sizes and the availability of under-coordinated Zr-based nodes, make them well-suited for the installation of single active metal sites via atomic layer deposition or solvent-assisted ligand incorporation[125]. Specifically, in NU-1000 material there is the varied pore architecture: comprising of hexagonal meso-channels (32 Å diameter) and triangular micropores (c-pores, 12 Å diameter)[134]. This architecture creates conditions under which local site chemistry and pore-level transport simultaneously influence the catalytic behavior[135]. Structural characterization studies have confirmed that grafted metal atoms preferentially occupy the c-pore environments, where confinement can influence both the stability of transition states and the mobility of growing oligomeric chains[136].

A DFT-based computational screening by Ye *et al.* identified Ni(II) as the most promising metal center among a series of transition-metal sites in NU-1000, which is based on calculated free energy profiles for ethylene dimerization[29]. The experimental catalytic performance of metal-doped NU-1000 systems has been investigated across several groups. Goetjen *et al.* reported that Cr(III)-NU-1000 exhibits measurable ethylene oligomerization activity under mild conditions, yielding products consistent with Schulz-Flory statistics toward C_8-C_{18} oligomers[137]. In subsequent works, the same group characterized the product distributions and deactivation behavior of these systems under varying conditions[138, 139]. Li *et al.* reported that Ni-grafted NU-1000 (Ni-NU-1000) producing mainly C_4-C_8 oligomers under flow conditions at 318 K and 2 bar ethylene pressure, consistent with predominantly dimerization activity[131]. Importantly, the system undergoes deactivation over time through accumulation of the oligomeric material within the pore channels, reducing accessible

pore volume and lowering ethylene conversion. Collectively, these experimental and computational studies have established Ni-NU-1000 as a model system in which the Cossee-Arlman mechanism operates[27], and where the kinetics of chain propagation versus termination govern the observed product distribution.

Despite the availability of experimental data, a quantitative computational description of ethylene oligomerization in Ni-NU-1000 requires resolving two challenges: (i) the appropriate choice of structural model for the catalytic active site, and (ii) the accuracy of the electronic structure method used to compute reaction energetics.

Structural Models

The complex, extended architecture of SACs in NU-1000, comprising of inorganic Zr-based nodes, pyrene-based organic linkers, and the grafted metal center, poses substantial challenges for quantum-chemical modeling. Two principal approaches have been employed: cluster models, in which a finite fragment of the MOF is carved out and saturated with capping groups, and periodic models, which impose translational symmetry over the full unit cell. Bernales and co-workers were the first to compute the mechanism of ethylene dimerization in Ni-AIM-NU-1000 using DFT calculations on a formate-cluster model, in which the pyrene linkers were replaced by formate groups[27]. They concluded that the Cossee-Arlman mechanism was energetically favored. Ye *et al.* later employed analogous models to screen a series of transition metals and demonstrated that the inclusion of partial benzoate linkers did not substantially alter the free energy profiles[28]. Similar formate-cluster model has been applied to butene dimerization in UiO-66-supported systems[140, 141].

The implicit assumption underlying these truncated is that the steric environment imposed by the pore confinement is negligible for small substrates. The validity of this approximation depends on the interplay between compound size and pore shape, and becomes increasingly problematic for longer oligomeric chains where confinement effects on transition-state geometries may be non-negligible[142, 143]. Motivated by this fact, hierarchical cluster correction schemes that combine the computational efficiency of periodic model with sufficient and applicable method and cluster model with appropriate size and quite accurate method have been developed for extended catalytic systems[45]. This approach was applied especially to zeolites and some inorganic surfaces[31, 144–148].

Electronic Structure Methods

A fundamental challenge concerns the accuracy of electronic structure methods for reaction barrier calculations. DFT method has been the primary computational

tool in all prior studies of ethylene oligomerization in MOF-supported catalysts[27–29], because of their favorable balance between cost and accuracy[78]. However, the accuracy of DFT is strongly functional-dependent, and this dependence is especially pronounced for the computation of reaction barriers, which are critical for predicting reasonable kinetics[24, 30, 149].

Previous theoretical studies[27–29] of ethylene oligomerization in NU-1000-supported catalysts employed either GGA functional (PBE) or *meta*-GGA functional (M06-L). These functionals are known to underestimate reaction barriers due to SIE, which, for example, artificially stabilize delocalized transition states[92]. Hybrid functionals, which partially correct SIE by incorporating exact exchange, generally improve barrier heights but carry a substantially higher computational cost, particularly under periodic boundary conditions[148, 150–152].

The consequences of functional dependence for barrier heights in catalytic systems relevant to the present work have been demonstrated concretely. Dohrmann *et al.* benchmarked different DFT functionals against experimental activation energies for ethylene hydrogenation and found that DFT methods systematically failed to reproduce the experimental trend, with errors spanning tens of kJ/mol across the functional ladder[153]. While DFT benchmarks for catalytic reactions on zeolite systems are available[30, 31, 148], analogous systematic assessments for MOF-based catalysts remain rare[153]. The magnitude of functional-induced errors is particularly consequential for kinetic predictions, since barrier heights are exponentially incorporated into rate constants. An error of even 5-10 kJ/mol in a rate-limiting activation barrier propagates to an order-of-magnitude error in the predicted rate at considered temperature and pressure, directly affecting the predicted product distributions in kinetic models[24].

Methods based on WFT provide a systematic order of approximations which are free from SIE and can, in principle, be converged towards the exact electronic energy[37, 154–156]. Among WFT approaches, coupled cluster theory at the CCSD(T) level[155] is widely recognized as the "gold standard" for molecular reaction energetics, providing excellent predictions of bond dissociation energies, reaction enthalpies, and activation barriers for systems dominated by a single electronic configuration[37, 49, 157–160]. The primary limitation of canonical CCSD(T) is its formal $\mathcal{O}(N^7)$ scaling with system size, which restricts its application to small chemical systems[161].

The DLPNO-CCSD(T) method overcomes this bottleneck by exploiting the locality of electron correlation through a localized molecular orbital basis and pair natural orbital truncation[162, 163]. DLPNO-CCSD(T) achieves relatively lower scaling with

system size, than canonical CCSD(T), while retaining the accuracy of canonical CCSD(T) for ground-state energies[164, 165]. Studies have demonstrated that DLPNO-CCSD(T) with tight PNO thresholds (so-called "TightPNO" settings) agree with canonical CCSD(T) within 1 kJ/mol for thermochemical properties[64, 166]. It is important to employ sufficiently tight PNO thresholds. Less stringent settings can introduce significant errors in the computed energetics, in particular to non-covalent contributions[63, 167].

Kinetic Modeling of Catalysts

While electronic structure calculations provide the foundation for understanding catalytic mechanisms, a direct comparison with experimental product distributions requires kinetic modeling that translates energy barriers into predicted rates and selectivities[168]. However, the literature on kinetic modeling of ethylene oligomerization in MOF-based catalysts is minimum compared to the electronic structure literature[112].

Microkinetic analysis based on TST provides a tractable route from DFT-derived free energy barriers to macroscopic rate constants[169, 170]. In the case of ethylene oligomerization, its simplest form contains a network of elementary reaction steps (e.g., coordination, insertion, termination) as master-equations to yield predicted product distributions as a function of temperature and pressure. There are many examples where master-equations were implemented, but not to MOF catalysts[171–175].

In contrary, for stochastic approaches, kinetic Monte Carlo (kMC) methods provide an alternative to the deterministic microkinetic framework, allowing explicit simulation of individual reaction events on the catalyst surface[176, 177]. kMC can naturally incorporate coverage-dependent adsorption and stochastic fluctuations in product distributions[178]. However, kMC becomes computationally demanding for systems with slow, rare events or where high statistical convergence is required. Deterministic master-equation methods offer a computationally more efficient alternative by directly solving the time evolution of the probability distribution over system states[172, 174].

Prior computational studies of ethylene oligomerization in Ni-NU-1000 and related systems have generally been omitted from microkinetic analysis. The independent studies of Bernales[27] and Ye[28, 29] provide Gibbs free energy profiles for individual reaction steps but do not translate these into predicted product distributions under realistic operating conditions. Le *et al.* explored kinetic aspects for related systems[177], but a combination of multiscale approaches, including DFT barriers and diffusion study for MOF-based oligomerization has been reported only once in co-authorship with the author of this thesis[112].

This chapter addresses the interconnected challenges of electronic structure accuracy, structural model selection, and kinetic prediction for ethylene oligomerization in Ni-NU-1000. The work presented here includes two complementary parts. Firstly, a hierarchical cluster correction scheme was developed. Especially, a systematical benchmark of the performance for different DFT functionals was performed against DLPNO-CCSD(T) for the Cossee-Arlman mechanism of ethylene dimerization in Ni-NU-1000[111]. Then the development of a microkinetic analysis was done and evaluated. This microkinetic model couples a deterministic master-equation based on the corrected DFT Gibbs free energy barriers with a microkinetic model of oligomerization, enabling quantitative prediction of the product distribution as a function of the correction method and type of pores.

Specifically, the objectives of this chapter are as follows:

(i) To develop the hierarchical cluster correction scheme for applying DLPNO-CCSD(T) energies to periodic DFT calculations of ethylene dimerization in Ni-NU-1000, and to benchmark different DFT functionals and MP2 against the DLPNO-CCSD(T) reference.

(ii) To identify based on the benchmark, the most accurate methods for the description of reaction's energetics in this system, and to validate the dependence of results on the structural cluster model.

(iii) To map with DFT-based correction on the complete oligomerization mechanism, including the following reaction stages: dimerization, trimerization, tetramerization.

(iv) To construct a microkinetic model based on the corrected Gibbs free energy barriers and to analyze how DFT functional, type of pores reshape the predicted product distribution.

3.1.2 Computational Details

Periodic DFT Calculations

To make a periodic structure of Ni-grafted NU-1000 material, a unit cell of NU-1000[179] with dimensions of 39.90 Å×39.90 Å×16.50 Å, with $\alpha = \beta = 90°$ and $\gamma = 120°$, was adopted (see Figure 2a). The Ni atom was deposited as described in the work of Bernales[27]. The optimizations and free energy calculations of the periodic model were performed using the PBE functional[86] with Grimme's D3(BJ) dispersion correction[100, 103] using the VASP code[180–183]. The plane-wave cutoff

was 520 eV, and the calculations were done at the Γ-point. All periodic calculations were performed by Thanh-Nam Huynh from Institute of Catalysis Research and Technology, Karlsruhe Institute of Technology. Throughout this work, the Ni center was investigated in the +2 oxidation state [Ni(II)], which was confirmed to be the electronic ground state. Accordingly, all subsequent references to Ni-NU-1000 denote the Ni(II)-grafted NU-1000 catalyst.

Cluster Models Construction

For the correction scheme and its validation, two cluster models (formate- and benzoate-cluster models), were considered, especially because of their widespread use[27, 29] (see Figure 2b,c). These two models involved truncation of organic linkers, while the Zr-based node with Ni was not changed. This representation of active sites is a typical approach for investigating catalytic processes in MOFs[184]. All intermediates and transition states of catalytic pathways in the formate- and benzoate-cluster models were extracted from the optimized periodic structures. Then, all dangling bonds in the clusters were terminated with hydrogen atoms. Finally, all hydrogen atoms were optimized using the PBE-D3(BJ) functional and the def2-TZVP basis set[71] in ORCA 5.0.4[66]. The final structures were used for hierarchical cluster correction scheme single-point energy calculations.

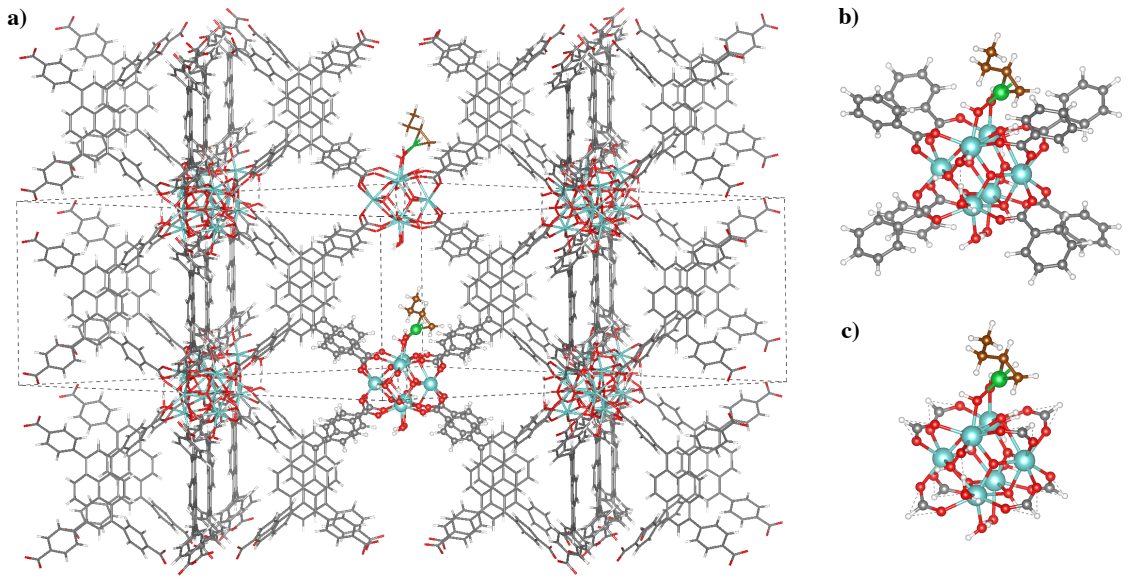


Figure 2: Three different models of Ni-NU-1000: (a) periodic model [pbc], (b) benzoate-cluster model, (c) formate-cluster model. In the visualization, chemical elements are represented by specific colors: Zr (cyan), Ni (green), O (red), framework C (gray); molecular C (brown), and H (white). Adopted with permission from [111]. Copyright 2025 American Chemical Society.

The Multireference Character Investigation

Firstly, the fractional occupation number weighted electron density analysis (FOD analysis) was performed for formate-cluster model at the B3LYP level with T = 9000 K and the def2-TZVP basis set[185, 186] (see Figure 3). Finally, the T2-diagnostic using DLPNO-CCSD(T) was performed, and the obtained T2 amplitudes were 0.2 or below, on average near 0.15. Structures with T2 amplitudes below 0.2 have been shown to be adequately treated with a single-configuration wavefunction[187, 188]. This supports a single-configuration description for all structures in the work of Bernales[27].

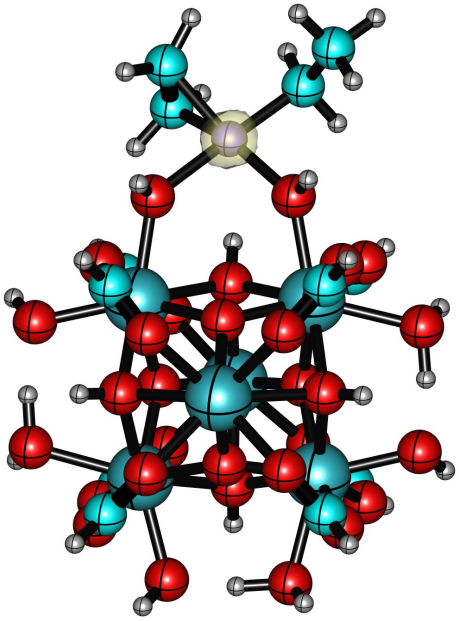


Figure 3: FOD plots for Intermediate **2** at an iso-contour value of $\sigma = 0.005\, a_0{}^{-3}$. In the visualization, chemical elements are represented by specific colors: Zr (dark cyan), Ni (violet), O (red), framework C (cyan), and H (gray). Reprinted with permission from [111]. Copyright 2025 American Chemical Society.

The Hierarchical Cluster Correction Calculations

All calculations for the hierarchical cluster correction scheme were performed in ORCA 5.0.4, including RI-DFT, RI-MP2 (names as DFT and MP2 later), DLPNO-MP2, and DLPNO-CCSD(T) single-point energy calculations. Basis sets ranging from def2-TZVP to def2-QZVPP were used in different parts of the hierarchical cluster correction scheme[71]. Additionally, RI-MP2, DLPNO-MP2, and DLPNO-CCSD(T) calculations were performed with CBS extrapolation. For the final DLPNO-CCSD(T) electronic energies, tight pair natural orbital thresholds were used ("TightPNO" settings). To benchmark DFT functionals, following functionals were chosen: PBE[86],

TPSS[87], M06-L[189], TPSSh[89], PBE0[190], B3LYP[191], ωB97M[192], M06[193], BHandHLYP[88], and M06-2X[194]. For the benchmark, the D3(BJ) dispersion correction was used[100, 103] (and D3zero for M06-family functionals) and the def2-QZVPP basis set was utilized.

To benchmark DFT and MP2 methods two statistical descriptors were used: mean absolute deviation (MAD), and mean signed deviation (MSD), which captures the systematic bias of a method relative to the reference. Both descriptors are defined as follows:

$$\mathrm{MAD} = \frac{1}{N}\sum_{i=1}^{N}\left|E_i^{Method} - E_i^{Reference}\right|, \tag{42}$$

$$\mathrm{MSD} = \frac{1}{N}\sum_{i=1}^{N}\left(E_i^{Method} - E_i^{Reference}\right), \tag{43}$$

where N is the number of complexes considered and i is each reaction energy.

Microkinetic Modeling

The standard harmonic oscillator approximation was used for periodic boundary conditions (pbc) structures at $T = 300$ K to obtain the Gibbs free energy for each stationary point i:

$$G_i(T) = E_i^{\mathrm{DFT}} - k_{\mathrm{B}}T\ln Q_{pbc,i}^{PBE-D3(BJ)}(T), \tag{44}$$

where E_i^{DFT} is the DFT-corrected electronic energy, k_{B} is Boltzmann constant, and $Q_{pbc,i}^{PBE-D3(BJ)}(T)$ is the partition function obtained from the partial Hessian matrix at the PBE-D3(BJ) level. All periodic calculations were performed by Thanh-Nam Huynh from the Karlsruhe Institute of Technology. The dependence of the Gibbs free energy on the temperature T and pressure p is calculated as:

$$G_g(T,p) = G_g^{\circ}(T) + k_{\mathrm{B}}T\ln\left(\frac{p}{p^{\circ}}\right), \tag{45}$$

where $G_g^{\circ}(T)$ is the standard state Gibbs free energy and $p^{\circ} = 1$ bar is the pressure at standard state. In this case, $T = 300$ K and $p = 1$ bar were chosen. The activation Gibbs free energy barrier is defined as:

$$\Delta G_r^{\ddagger}(T,p) = G_{\mathrm{TS},r}(T,p) - G_{\mathrm{reactant},r}(T,p). \tag{46}$$

The reaction Gibbs free energy difference is defined as:

$$\Delta G_r(T,p) = G_{\text{product},r}(T,p) - G_{\text{reactant},r}(T,p). \tag{47}$$

Based on the obtained activation free energy barriers (or Gibbs free energy barriers), the elementary rate constant for reaction i ($k_i(T)$) can be calculated using Eyring transition-state theory[195]:

$$k_i(T) = \frac{k_\text{B}T}{h} \exp\left(-\frac{\Delta G_i^\ddagger(T)}{k_\text{B}T}\right), \tag{48}$$

where h is Planck's constant, and $\Delta G_i^\ddagger(T)$ is the Gibbs free energy barrier for reaction i at $T = 300$ K and $p = 1$ bar.

The kinetic modeling framework[112] developed by Alexander Avdoshin, Thanh-Nam Huynh and the author of this thesis was used to perform the microkinetic analysis. In this framework, chemical transformations were treated as a unified set of ordinary differential equations based on master-equation formalism:

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = \sum_j \nu_{ij}\, r_j \quad (i = 1, \ldots, N_\text{sp}), \tag{49}$$

where C_i is the concentration of species i, ν_{ij} is the stoichiometric coefficient in elementary step j, and r_j is the mass-action rate.

For the microkinetic analysis only concentration of ethylene is presented and limited to a fixed amount at the initial step of the simulation. The initial concentration of ethylene in the reactor was approximately 13 molecule per unit cell. The concentration of catalytic centers was 10^{-2} per unit cell. It should be noted that this microkinetic analysis does not account for diffusion or capacity effects (no mass-transfer), which considerably limits the applicability of the predictions when compared to experimental data and real operating conditions.

3.2 The Hierarchical Cluster Correction Scheme for Ethylene Dimerization

This section addresses the first stage of ethylene oligomerization, namely ethylene dimerization in Ni-NU-1000 catalyst. Firstly, the limitation of model and method

choice is discussed. Using the reaction pathway given in Figure 4 as the reference, the construction of the hierarchical cluster correction scheme based on DLPNO-CCSD(T), along with various corrections and basis sets, is demonstrated. Additionally, the influence of the basis set on the correlation energy in DLPNO-CCSD(T) and CBS extrapolation values with MP2 for different structures are presented. Subsequently, various DFT functionals were benchmarked. Their deviations are presented with respect to DLPNO-CCSD(T). The deviations for different size of clusters of two most accurate DFT functionals are also discussed. Additionally, turnover frequency (TOF) values for DLPNO-CCSD(T) as well as for the best DFT functionals were calculated and compared against the experimental data. This section is methodological, as it establishes the foundation upon which all energy barriers will be subsequently calculated when considered the complete mechanism of ethylene oligomerization.

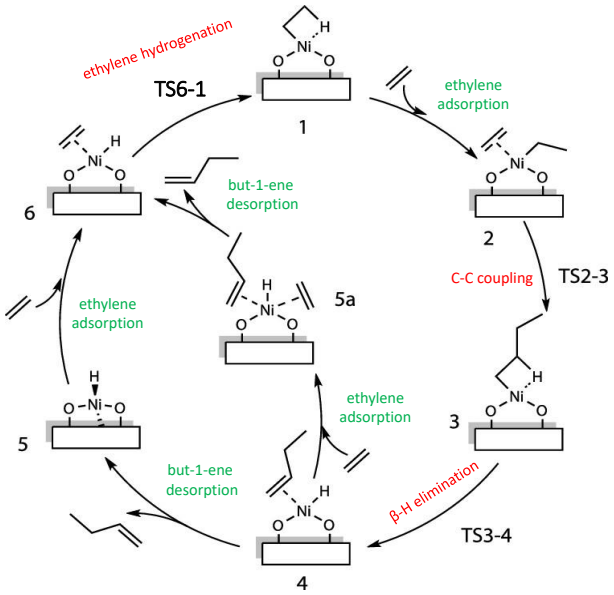


Figure 4: The Cossee-Arlman mechanism of ethylene dimerization in Ni-NU-1000. The scheme illustrates the truncated Ni catalytic site as the location at which ethylene conversion directly takes place. Adopted with permission from [111]. Copyright 2025 American Chemical Society.

3.2.1 The Reaction Mechanism of Ethylene Dimerization

The ethylene dimerization reaction mechanism illustrated in Figure 4 begins with ethylene adsorption ([Ni-ethyl] **1** → [et-Ni-ethyl] **2**), followed by a C-C coupling reaction (**TS2-3**) that yields a β-H-butyl intermediate [Ni-butyl] (**3**). Subsequently, a

β-H elimination reaction (**TS3-4**) transforms the butyl species into adsorbed but-1-ene [H-Ni-but], which can be released via two distinct pathways: direct desorption (**5**) or substitutive desorption (**5a**). In the case of direct desorption (**4** $\rightarrow$ **5**), the final state consists of a Ni center bearing a hydrogen atom [Ni-H] (**5**), onto which ethylene can subsequently be adsorbed (**5** $\rightarrow$ **6**). Alternatively, in the substitutive pathway, ethylene first is adsorbed onto the [but-Ni-H] intermediate (**5a**, [et-Ni-but-H]), after which but-1-ene is displaced and removed (**5a** $\rightarrow$ **6**). Both pathways finish at the intermediate [et-Ni-H] (**6**), in which both hydrogen and ethylene are coordinated to the Ni center. Finally, ethylene hydrogenation (**TS6-1**) regenerates the [Ni-ethyl] species (**1**), therefore closing the catalytic cycle. All intermediate and transition state geometries for active site can be found in Figure 51 (Appendix for Chapter 3, Section 3.2).

Based on this reaction pathway, five principal activation energy barriers of ethylene dimerization in Ni-NU-1000 can be identified:

1. C-C coupling:

$$\text{[et-Ni-ethyl]} \rightarrow \text{[Ni-butyl]}, \quad \Delta\Delta E_1^{\ddagger} = \Delta E(\textbf{TS2-3}) - \Delta E(\textbf{2}). \tag{50}$$

2. β-H elimination:

$$\text{[Ni-butyl]} \rightarrow \text{[H-Ni-but]}, \quad \Delta\Delta E_2^{\ddagger} = \Delta E(\textbf{TS3-4}) - \Delta E(\textbf{3}). \tag{51}$$

3. Direct but-1-ene desorption:

$$\text{[H-Ni-but]} \rightarrow \text{[Ni-H]} + \text{C}_4\text{H}_{8\,\textbf{(gp)}}, \quad \Delta\Delta E_3^{\ddagger} = \Delta E(\textbf{5}) + \Delta E(\text{C}_4\text{H}_8) - \Delta E(\textbf{4}). \tag{52}$$

4. Substitutive but-1-ene desorption:

$$\begin{aligned} &\text{[H-Ni-but]} + \text{C}_2\text{H}_{4\,\textbf{(gp)}} \rightarrow \text{[et-Ni-but-H]}, \\ &\Delta\Delta E_4^{\ddagger} = \Delta E(\textbf{5a}) - \Delta E(\textbf{4}) - \Delta E(\text{C}_2\text{H}_4). \end{aligned} \tag{53}$$

5. Ethylene hydrogenation:

$$\text{[H-Ni-Et]} \rightarrow \text{[Ni-ethyl]}, \quad \Delta\Delta E_5^{\ddagger} = \Delta E(\textbf{TS6-1}) - \Delta E(\textbf{6}), \tag{54}$$

where the subscripts (**gp**) denote the gas-phase molecules.

3.2.2 The Importance of Correction Scheme in Catalysis

Before describing the hierarchical cluster correction, it is important to emphasize its necessity for the modeling of MOF catalysts. Two principal questions must be addressed in order to correctly describe the energetics of the catalytic process: which model to use for representing the catalytic center, and which theoretical method to employ. This is very important because, for example, an error of 10 kJ/mol in the rate-limiting activation barrier of a given reaction can introduce an error of 10^1-10^2 in the theoretical rate constant relative to the experimental value[24]. Figure 2a-c shows three different models: the periodic unit cell of the Ni-NU-1000 crystal and two clusters cut from it, with hydrogen termination at the boundaries (see Section 3.1.2). First, all structures relevant to the ethylene dimerization mechanism were computed using the PBE-D3(BJ) method for two types of catalytic model: periodic and formate-cluster models. The data and structures of the periodic model were obtained by Thanh-Nam Huynh, while the structures of the formate-cluster model were created and computed by the author of this thesis. The results of the comparison between the two models are presented in Table 2. It can be seen that the deviations of the formate-cluster model from the periodic model can reach up to 17 kJ/mol, as in the case of Intermediate **5a**. On average, the mean absolute deviation of the formate-cluster model from the periodic model amounts to 7 kJ/mol. Thus, the absence of long-range interactions in the formate-cluster model introduces errors in the computed energies, which may subsequently affect the predicted kinetics of the process. It should also be noted that PBE, and certain DFT methods in general, exhibit substantial deviations from reference values, particularly in the description of catalytic mechanisms in related systems[30]. Therefore, a more accurate method is required to correctly describe the electronic structure of the catalytic center. The DLPNO-CCSD(T) method, for instance, enables calculations on systems considerably larger than those tractable with canonical CC, while incurring only a modest loss in accuracy (see Section 2.1.1). The crucial limitation is that any implementation of CC is currently restricted to finite cluster models and cannot be applied to extended periodic systems[196]. Consequently, for porous catalysts such as MOFs, where the chemical environment plays an important role, both long-range interactions and an accurate description of the catalytic center must be accounted for simultaneously. There is a QM:QM embedding scheme approach, which allows both requirements to be fulfilled.

Table 2: Relative electronic energies (in kJ/mol) of all structures along the ethylene dimerization pathway in Ni-NU-1000, computed with the periodic and formate-cluster models using the PBE-D3(BJ) method. Finally, the differences between the periodic model and formate-cluster model electronic energy values are also presented. Both models are uncorrected.

Structure	**periodic model**	**formate-cluster model**	**Δ periodic/cluster**
1	0	0	0
2	-107	-101	-6
TS2-3	-25	-18	-7
3	-134	-120	-14
TS3-4	-94	-83	-11
4	-101	-94	-7
5	-5	-5	0
5a	-138	-121	-17
6	-98	-98	0
TS6-1	-90	-87	-3
1 + C_4H_8	-127	-126	-1

3.2.3 The Foundation of the Hierarchical Cluster Correction Scheme

The hierarchical cluster correction scheme is based on the idea of treating the electronic structure in two separate parts, in this case as a QM:QM embedding scheme[31, 197, 198]. It is important to note that only electronic energies were considered. Within the QM:QM treatment, the formula for the correction is:

$$E^{corrected} = E_{cluster}^{Method\ 1} + [E_{pbc}^{Method\ 2} - E_{cluster}^{Method\ 2}], \tag{55}$$

where Method 1 has relatively higher level of theory for treating the electronic structure of the active site than Method 2. Importantly, to replace the cluster electronic energy, the same method must be used as for the periodic electronic energy calculations. In this case, it is PBE-D3(BJ).

The challenge lies in selecting a model of appropriate size: it must be large enough not to miss important electronic structure features for describing intermediates and yet small enough to remain computationally tractable. As described previously, the formate-cluster model used in the literature is sufficient. It comprises the full Zr-based node with the Ni atom. Therefore this type of system was adopted for final

electronic energy corrections. Furthermore, the benzoate-cluster model was also tested with two DFT functionals, as will be described later. However, the formate-cluster model was still too large to employ large basis sets and CBS extrapolation with the DLPNO-CCSD(T) method.

As the reference method, DLPNO-CCSD(T) was selected. Accordingly, the formula of the hierarchical cluster correction scheme for Ni-NU-1000 evolves to:

$$E^{corrected} = E^{DLPNO-CCSD(T)}_{cluster} + [E^{PBE-D3(BJ)}_{pbc} - E^{PBE-D3(BJ)}_{cluster}], \tag{56}$$

where $E^{\text{DLPNO-CCSD(T)}}_{\text{cluster}}$ is the electronic energy of the cluster obtained at the DLPNO-CCSD(T) level of theory. It is important to note that, in its simplest form, $E^{\text{DLPNO-CCSD(T)}}_{\text{cluster}}$ comprises two contributions: the Hartree–Fock energy and the correlation energy. The Hartree–Fock energy was computed separately using the CBS extrapolation based on the def2-TZVPP/def2-QZVPP basis set pair (see Section 2.1.2), in order to minimize the potential basis set superposition error.

However, it has been demonstrated that, due to limitations in computational resources, even for the formate-cluster model one is constrained to a finite basis set. A common strategy to reduce computational cost without significant loss of accuracy is the use of hybrid (mixed) basis sets, in which a basis set with more basis functions is assigned to the atoms or groups of atoms directly participating in the process of interest (catalytic site), whereas the remaining spectator atoms are described with a smaller basis set[152]. It should be noted, however, that the basis set quality across different regions of the system must be balanced: assigning an excessively large basis set to the central atom while treating the surrounding ligands with a considerably smaller one may lead to an imbalanced description and introduce systematic errors[199]. Specifically, the mixed basis sets def2-TZVPP/def2-TZVP and def2-QZVPP/def2-TZVP were constructed such that the larger basis set is placed on the atoms belonging to the small-cluster model (Figure 5).

Firstly, different basis sets were tested, both in full and in mixed forms, by calculating the correlation energies relative to Intermediate **1**. It should be noted that this assessment was performed before the complete set of final structures was determined; however, it is sufficient for identifying general trends. The following basis sets were employed: def2-TZVP, def2-TZVPP/def2-TZVP, def2-TZVPP, and def2-QZVPP/def2-TZVP. The results are presented in Table 3. One can see that the deviations obtained with the def2-TZVP and def2-TZVPP/def2-TZVP basis sets can be substantial,

reaching up to 5 kJ/mol for Intermediate **5**. This highlights the importance of using the largest feasible basis set in order to minimize BSIEs. Interestingly, the difference between the mixed def2-TZVPP/def2-TZVP and the full def2-TZVPP basis sets is relatively small, with a maximum deviation of 2 kJ/mol, suggesting that a similar trend is likely to persist at higher basis sets. Furthermore, for nearly all structures the correlation energy contribution shifts in the same direction, It becomes more negative with increasing basis set size with the exception of Intermediate **2** and Intermediate **5**. Therefore, the def2-QZVPP/def2-TZVP basis set was chosen as the optimal compromise for the final correlation energy calculations.

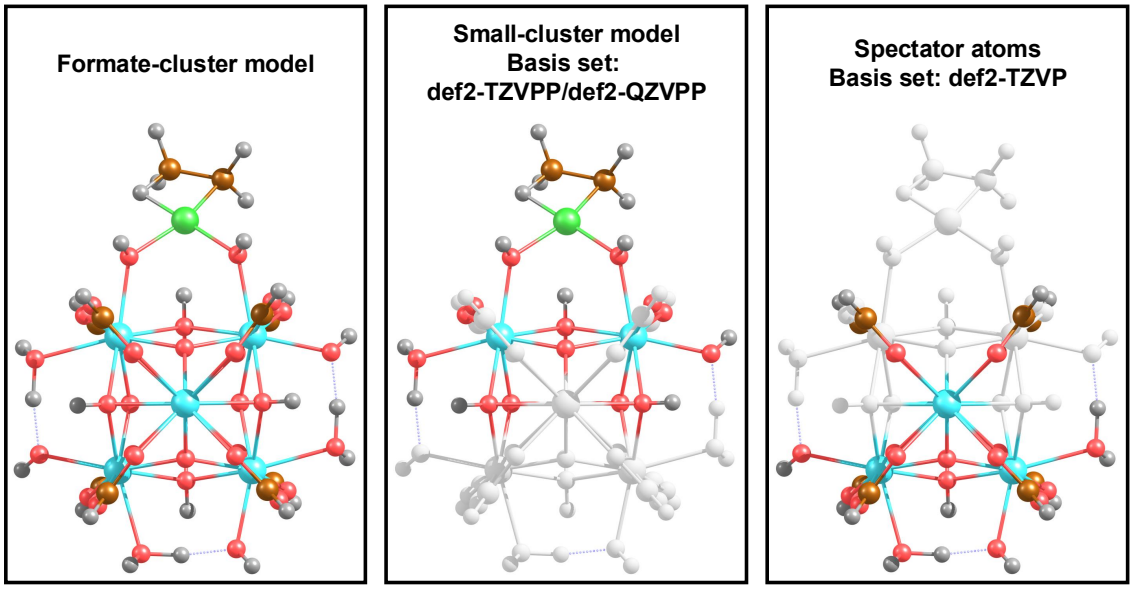


Figure 5: The view of the formate-cluster model and the atom selection scheme for the mixed basis set, for example by the def2-QZVPP/def2-TZVP or def2-TZVPP/def2-TZVP combinations, in which the small-cluster model of Ni catalytic center is covered by def2-QZVPP or def2-TZVPP basis functions, while the atoms belonging to the remaining part (spectator atoms) of the Zr-based node are described with the def2-TZVP basis set. In the visualization, chemical elements are represented by specific colors: Zr (cyan), Ni (green), O (red), C (brown), and H (grey).

The following step was to estimate the energy corrections arising from the CBS extrapolation that was already used in the same correction scheme[30]. First, the CBS extrapolation correction was calculated as the difference between the MP2 energy obtained with CBS extrapolation technique (see Section 2.1.2) using the def2-TZVPP/def2-QZVPP basis set pair and the MP2 energy computed with the

Table 3: Relative electronic energies (in kJ/mol) of all structures along the ethylene dimerization pathway in Ni-NU-1000, computed with the formate-cluster model using the MP2 method with different basis sets. TZVP an TZVPP denotes def2-TZVP and def2-TZVPP, TZVPP/TZVP and QZVPP/TZVP denotes the mixed def2-TZVPP/def2-TZVP and def2-QZVPP/def2-TZVP basis sets.

Structure	**TZVP**	**TZVPP/TZVP**	**TZVPP**	**QZVPP/TZVP**
1	0	0	0	0
2	-125	-121	-120	-123
TS2-3	-75	-76	-76	-77
3	-39	-42	-40	-45
TS3-4	-155	-157	-158	-163
4	-95	-97	-97	-101
5	100	95	97	99
6	-65	-66	-65	-66
TS6-1	-75	-78	-78	-81

def2-QZVPP basis set alone. The relative energies with respect to Intermediate **1** are shown in Table 4. It is important to note that the relative energies decrease monotonically from def2-TZVPP to def2-QZVPP and further to the CBS limit. However, the magnitude of this decrease varies across different structures. Therefore, it is essential to account for this CBS effect, as it may have a non-negligible influence on the final results. One can see that the difference between the CBS and def2-TZVPP values (ΔCBS/TZVPP) can reach up to 9 kJ/mol, whereas the difference between the CBS and def2-QZVPP values (ΔCBS/QZVPP) amounts to approximately 4 kJ/mol. This MP2 CBS correction was subsequently incorporated into the total energy expression.

Additionally, to assess the effect of the DLPNO approximation, the difference between canonical MP2 and DLPNO-MP2[200] energies with def2-QZVPP basis set was computed and implemented in the scheme. This approach has been shown to provide a reliable estimate of the discrepancy between DLPNO-CCSD(T) and canonical CCSD(T)[201].

Finally, the $E_{\text{cluster}}^{\text{DLPNO-CCSD(T)}}$ equation has the following contributions (*Hybrid* designated hybrid basis set def2-QZVPP/def2-TZVP):

$$E_{cluster}^{DLPNO-CCSD(T)} = E_{CBS}^{HF} + E_{Hybrid}^{CCSD} + E_{Hybrid}^{Triples} + \Delta E_{DLPNO}^{MP2} + \Delta E_{CBS}^{MP2}, \quad (57)$$

where

$$\Delta E_{DLPNO}^{MP2} = E_{def2-QZVPP}^{MP2} - E_{def2-QZVPP}^{DLPNO-MP2}, \quad (58)$$

$$\Delta E_{CBS}^{MP2} = E_{CBS}^{MP2} - E_{def2-QZVPP}^{MP2} \quad (59)$$

Table 4: Relative electronic energies (in kJ/mol) of all structures along the ethylene dimerization pathway in Ni-NU-1000, computed with the formate-cluster model using the MP2 method with different basis sets. TZVPP denotes def2-TZVPP, QZVPP denotes def2-QZVPP, and CBS denotes the complete basis set limit obtained via the two-point CBS extrapolation scheme using the def2-TZVPP/def2-QZVPP basis set pair (see Section 2.1.2). Finally, the differences between the CBS values and those of the two individual basis sets are also presented.

Structure	**TZVPP**	**QZVPP**	**CBS**	**Δ CBS/TZVPP**	**Δ CBS/QZVPP**
1	0	0	0	0	0
2	-145	-148	-151	-6	-3
TS2-3	-91	-94	-95	-4	-2
3	-56	-57	-58	-2	-1
TS3-4	-141	-146	-150	-9	-4
4	-130	-133	-136	-6	-3
5	129	126	123	-6	-2
5a	-165	-169	-172	-7	-3
6	-82	-84	-85	-3	-1
TS6-1	-103	-106	-109	-6	-2
1 + C_4H_8	-35	-35	-36	-1	-1

3.2.4 DLPNO-CCSD(T) Correction of Ethylene Dimerization

All structures along the mentioned ethylene dimerization reaction pathway were calculated using the hierarchical cluster correction scheme based on the DLPNO-CCSD(T) method (denoted as CC) for both electronic energies and Gibbs free energies. Thermal correction to Gibbs free energy, were calculated on the periodic model of Ni-NU-1000 with PBE-D3(BJ) by Thanh-Nam Huynh. A comparison with periodic PBE-D3(BJ) calculations (denoted as PBE) is presented in Figures 6,7. Figure 6 shows

the reaction pathway for direct desorption (**5**), while Figure 7 shows the reaction pathway for substitutive desorption (**5a**).

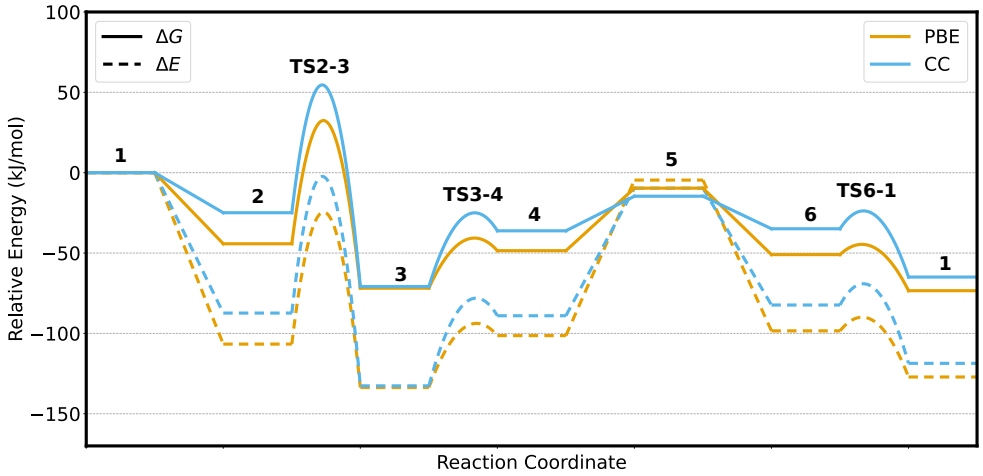


Figure 6: Comparisons of relative values (in kJ/mol) in electronic energy (dashed lines) and Gibbs free energy profiles at 298.15 K (solid lines) for the ethylene dimerization in Ni-NU-1000 obtained from periodic PBE and the DLPNO-CCSD(T)-based (CC) correction scheme calculations. Reprinted with permission from [111]. Copyright 2025 American Chemical Society.

First, the adsorption and desorption energies are overestimated by more than 20 kJ/mol at the periodic PBE level compared to the CC correction scheme. For example, the energy of ethylene adsorption in intermediate **1** (**1** → **2**) is -107 kJ/mol with PBE, while with CC this value is -86 kJ/mol. A similar situation is observed for the direct but-1-ene desorption (**4** → **5**): at the PBE level, the reaction energy is 97 kJ/mol, while with CC it is reduced to 78 kJ/mol.

Some energy barriers are nearly equal for both approaches. For instance, the PBE activation energy barriers for C-C coupling (**TS2-3**) and ethylene hydrogenation (**TS6-1**) are 82 kJ/mol and 8 kJ/mol, respectively. With the CC correction, these values increase slightly to 85 kJ/mol and 13 kJ/mol, respectively.

More importantly, as has been demonstrated previously for hydrogen transfer reactions, DFT tends to significantly underestimate the activation energy for such processes[202]. The same trend is observed for β-H elimination: periodic PBE yields an activation energy (**TS3-4**) of 40 kJ/mol, whereas CC gives an activation barrier of 55 kJ/mol. This discrepancy can significantly affect the kinetic modeling of such processes.

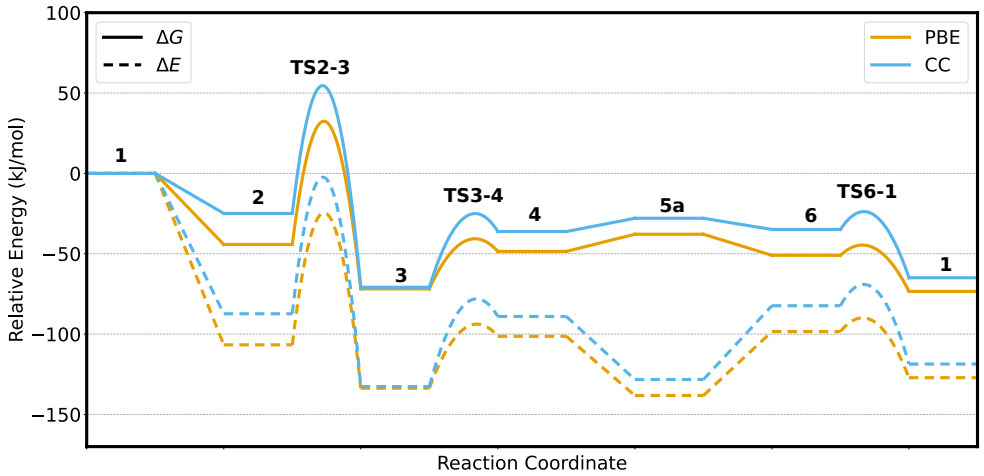


Figure 7: Comparisons of relative values (in kJ/mol) in electronic energy (dashed lines) and Gibbs free energy profiles at 298.15 K (solid lines) for the ethylene dimerization in Ni-NU-1000 obtained from periodic PBE and the DLPNO-CCSD(T)-based (CC) correction scheme calculations. Reprinted with permission from [111]. Copyright 2025 American Chemical Society.

When thermal correction is taken into account (Gibbs free energies evaluated at 298.15 K), the relative energies of all structures with respect to intermediate **1** are shifted upward, rendering, for instance, the adsorption of ethylene less thermodynamically favorable compared to the electronic energy profile (Figure 6). This effect is particularly pronounced for both desorption pathways - direct and substitutive. In the case of direct desorption, the energy difference between intermediates **4** and **5** on the Gibbs free energy surface is considerably smaller than that on the electronic energy surface. For the substitutive pathway, the inclusion of thermal effects qualitatively alters the character of the process: whereas the electronic energy decreases from intermediate **4** to intermediate **5**, the Gibbs free energy exhibits a slight increase along the same step (Figure 7).

Based on the corrected Gibbs free energy profile, the ethylene insertion barrier remains the rate-determining step, with an activation energy of 80 kJ/mol. The β-H elimination activation energy is moderate (46 kJ/mol), while the hydrogenation barrier is readily overcome. Furthermore, it is evident that substitutive desorption is considerably more favorable than direct desorption from a kinetic standpoint. With the corrected energy profile established as a reference, it is now possible to benchmark various DFT methods and identify the most appropriate functionals on the basis of statistical deviations.

3.2.5 DFT Functional Benchmark for Ethylene Dimerization

Firstly, to assess the performance of the different methods, the energy difference relative to intermediate **1** was considered. Figure 8 presents the performance of selected density functionals and of the MP2 method, assessed through the mean absolute deviation and the mean signed deviation computed for adsorbed intermediates (top), transition states (center), and the complete set of stationary points (bottom) along the catalytic pathway, respectively. The PBE functional consistently overstabilizes both intermediates (11 kJ/mol) and transition states (20 kJ/mol), resulting in an overall MAD of 14 kJ/mol. The error is particularly pronounced for transition states, that makes this functional artificial, because other methods do not have so large deviation between intermediates and transition states. In contrast, the PBE-based hybrid functional (PBE0) performs noticeably better than PBE, with an overall MAD of 5 kJ/mol. Examining the overall trend, most of the considered functionals (*meta*-GGA and hybrids) yield smaller MAD values than PBE, ranging from 5 to 10 kJ/mol. The three best-performing functionals, with overall MAD in the range of 5-6 kJ/mol, are TPSSh, PBE0, and M06-L. Interestingly, a high fraction of Hartree-Fock exchange correlates with larger deviations, as observed for BHandHLYP (50%) and M06-2X (54%). The MP2 method also falls outside the range of acceptable accuracy for these relative energies (overall MAD equals 26 kJ/mol). On average, the MSD values for GGA and *meta*-GGA methods are negative, indicating that these methods tend to underestimate the relative energies, whereas hybrid functionals show positive MSD values, meaning that they overestimate considered relative energies.

Then, as was shown before, five principal activation Gibbs free energy barriers were considered, and the deviation for each of them is additionally shown in Figure 9 for the selected methods (BHandHLYP, M06-2X, and MP2 were excluded due to their excessively high MAD values). As discussed previously, C-C coupling is one of the highest barrier among five, making it the rate-determining step. Therefore, the accurate prediction of this barrier is one of the most important. For C-C coupling (**TS2-3**), except for the M06-L and TPSS functionals, the other methods underestimate this activation energy. While most of the functionals show deviations within 10 kJ/mol, ωB97M, PBE0, and M06 exhibit deviations in the range of 10-20 kJ/mol. The β-H elimination reaction (**TS3-4**) was challenging for most methods, except for TPSSh and B3LYP, which show a maximum deviation within 5 kJ/mol. Overall, PBE and *meta*-GGA functionals yield lower values for this energy barrier, whereas hybrid functionals overestimate it. The hydrogenation barrier (**TS6-1**) is rather small, making any

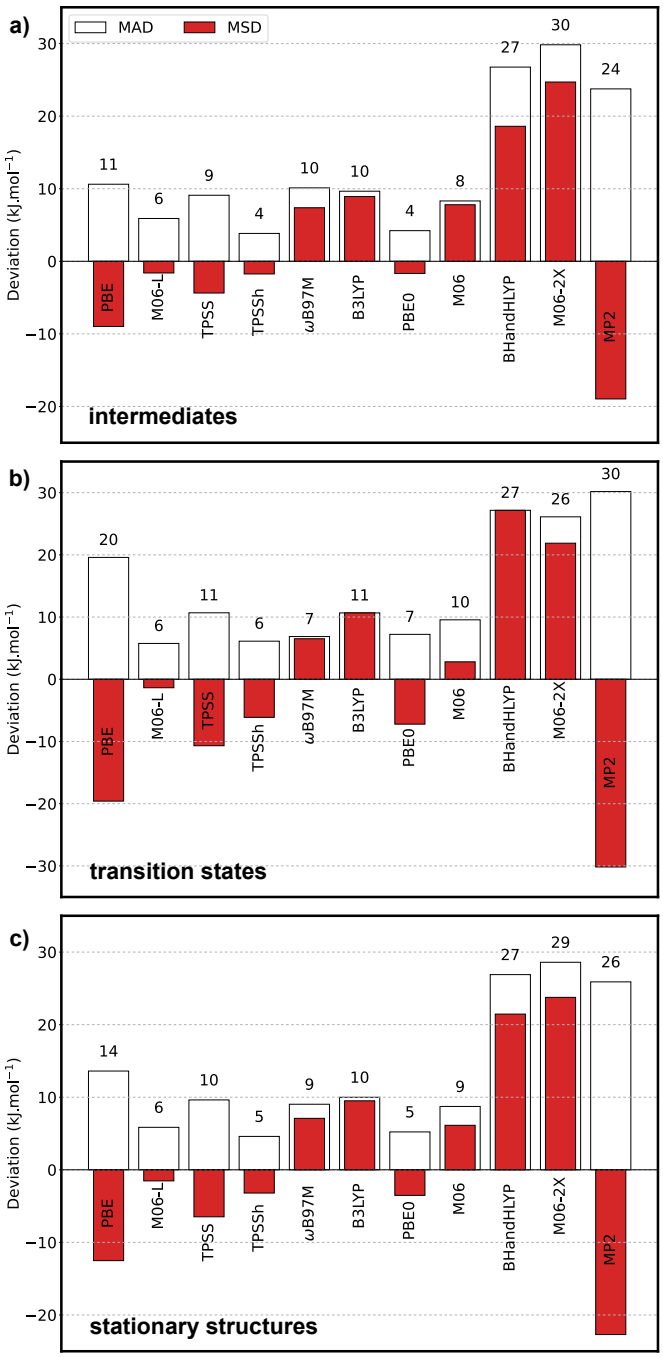


Figure 8: Performance of the MP2 method and a range of density functionals for ethylene dimerization in the Ni–NU-1000 catalyst, estimated via the mean absolute deviation (MAD) and mean signed deviation (MSD) relative to the CC reference (in kJ/mol); a) and b) show the deviations in the energies of intermediates and transition states, respectively, referenced to intermediate **1**; c) presents the overall deviations aggregated across all intermediates and transition states. The value above each bar indicates the corresponding MAD (in kJ/mol). Reprinted with permission from [111]. Copyright 2025 American Chemical Society.

deviation negligible, so all methods are on average in good agreement with the CC reference.

Desorption activation Gibbs free energy barriers are also important for describing the overall kinetic behavior of the catalysts. In the case of direct desorption, the best-performing functionals are M06-L, PBE, and PBE0, with deviations within approximately 7 kJ/mol. TPSS and TPSSh also do not show large deviations from the CC reference. In contrast, for substitutive desorption the best-performing functionals are TPSSh and B3LYP, which show deviations within 5 kJ/mol. PBE, M06-L, and TPSS underestimate this energy barrier, while ωB97M, PBE0, and M06 functionals overestimate the substitutive desorption barrier.

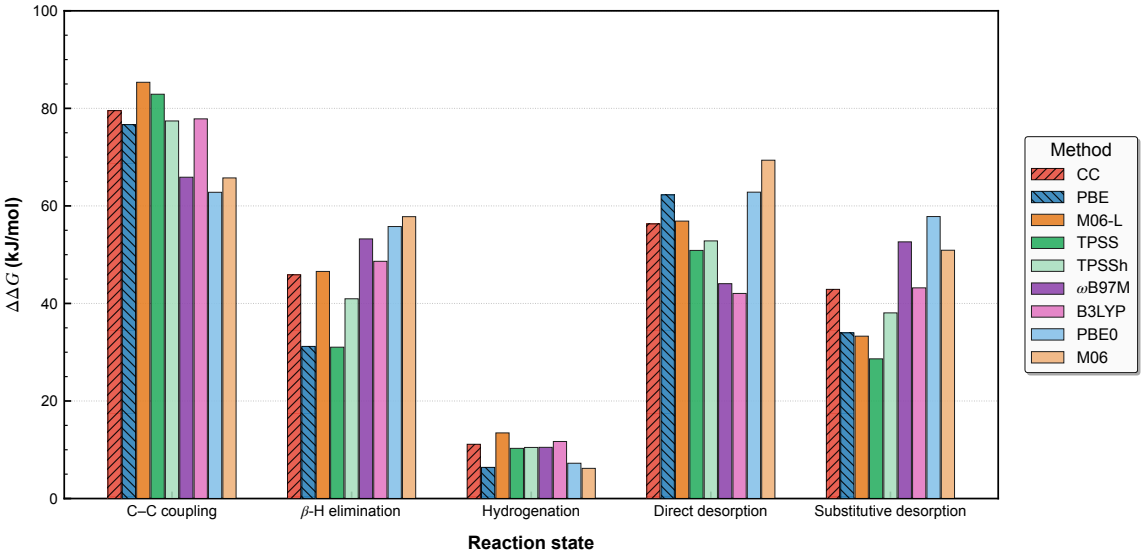


Figure 9: Activation free energies (in kJ/mol) for five principal energy barriers computed using selected DFT functionals within the hierarchical cluster correction scheme. Adopted with permission from [111]. Copyright 2025 American Chemical Society.

Finally, the overall MAD and MSD for the five principal steps and for the TS-related steps only are shown in Figure 10. The three best-performing functionals are TPSSh (with an MAD of 3 kJ/mol), M06-L (4 kJ/mol), and B3LYP (4 kJ/mol). This means that these three functionals, especially TPSSh and M06-L, accurately predict the activation energy barriers relative to the CC reference. This is supported by previous studies where M06-L was used for cluster-based modeling of Zr-based catalysts[27, 28, 131].

Based on the benchmark presented above, two methods, TPSSh and M06-L, showed the best performance both for the relative energies (Figure 8) and for the five principal

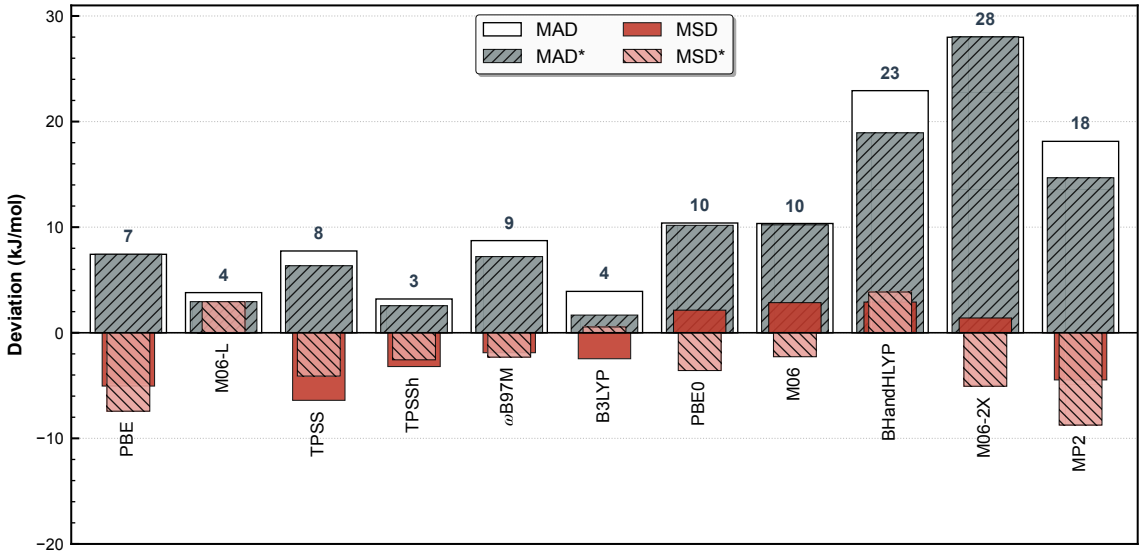


Figure 10: Performance of MP2 and DFT functionals assessed via MAD and MSD relative to the CC reference (in kJ/mol). MAD* and MSD* account exclusively for barriers involving transition states, whereas MAD and MSD reflect the deviations computed over the five principal energy barriers. The value above each bar indicates the corresponding MAD (in kJ/mol). Adopted with permission from [111]. Copyright 2025 American Chemical Society.

steps in ethylene dimerization (Figure 10). Therefore, these two methods were selected for the hierarchical cluster correction scheme applied to the complete mechanism of ethylene oligomerization.

However, before proceeding to the complete mechanism, the hypothesis of a sufficient size of the modeled system was additionally verified. To this end, for both the TPSSh and M06-L functionals, calculations within the hierarchical cluster correction scheme were performed using two different models: the benzoate- and formate-cluster models. The benzoate-cluster model was cut out from the periodic structure in the same way, with the edges terminated by hydrogen atoms that were subsequently optimized (see Section 3.1.2). The relative energies with respect to intermediate **1** were computed for all the remaining structures. The results are presented in Table 5. In all cases relative energy was slightly lower for benzoate-cluster model than for formate-cluster model. It can be seen that, in the case of M06-L, noticeable deviations were observed, particularly for intermediate **3** (a difference of 4 kJ/mol) and intermediate **5a** (6 kJ/mol). In contrast, for TPSSh the difference between the two model sizes does not exceed 2 kJ/mol (intermediate **5a**), while the other structures either coincide or differ by only 1 kJ/mol. However, in the case of M06-L, these differences in the relative

energies will partially cancel each other out, which on average decreases the effect of employing a different cluster model.

Table 5: Relative electronic energies (in kJ/mol) of all structures along the ethylene dimerization pathway in Ni–NU-1000, computed with the formate-cluster and benzoate-cluster models in hierarchical cluster correction scheme using the M06-L and TPSSh functionals.

Structure	**M06-L**		**TPSSh**	
	formate	benzoate	formate	benzoate
1	0	0	0	0
2	-104	-106	-99	-100
TS2-3	-13	-16	-16	-17
3	-127	-131	-128	-128
TS3-4	-71	-73	-78	-78
4	-89	-92	-90	-91
5	-3	-4	-8	-8
5a	-132	-138	-128	-130
6	-85	-87	-86	-86
TS6-1	-69	-71	-73	-73
1 + C_4H_8	-117	-117	-119	-119

Therefore, it can be concluded that the formate-cluster model provides, on average, a qualitatively correct description of the active site, since increasing the size of the model leaves the relative energies essentially unchanged. Consequently, the methods and the model selected above can now be used for the calculations of the complete ethylene oligomerization mechanism.

3.2.6 Discussion

In summary, the hierarchical cluster correction scheme for the Ni-NU-1000 catalyst was developed based on the DLPNO-CCSD(T) method using the formate-cluster model. Subsequently, a range of DFT functionals and MP2 were benchmarked. Based on the MAD and MSD values for the relative energies of all structures along the reaction pathway, as well as for the five principal reaction barriers, the best-performing functionals were identified. TPSSh and M06-L emerged as the top candidates, exhibiting MADs of approximately 3–4 kJ/mol relative to the DLPNO-CCSD(T) reference.

In the absence of direct experimental data, it is difficult to unambiguously judge which method is the most accurate. Nevertheless, a comparison based on the turnover frequency, which reflects the rate-determining process is possible, using the following formula for the energetic span approximation[22]:

$$\mathrm{TOF} = \frac{k_\mathrm{B}T}{h} \exp\left(\frac{-\delta G}{k_\mathrm{B}T}\right), \tag{60}$$

where δG is the energetic span, which is the activation free energy of the rate-determining state (C-C coupling reaction), k_B is the Boltzmann constant, h is the Planck constant, and T is the temperature. TOF shows that TPSSh reproduces the experimental order of magnitude (612 h^{-1}), whereas M06-L underestimates it (25 h^{-1}). The TOF in 260 h^{-1} for the DLPNO-CCSD(T) method almost exactly matches the experiment value of 252 h^{-1}[131]

Finally, two different models of the Ni-NU-1000 catalytic site were tested: the formate-cluster and the benzoate-cluster models. The results demonstrated that the two models yield nearly identical relative energy values. The largest deviation was observed for M06-L at 6 kJ/mol, while for TPSSh the maximum discrepancy was only 2 kJ/mol.

On the basis of the findings described above, appropriate DFT methods have been identified that enable the calculation of the complete mechanism of ethylene oligomerization with an accuracy pretending to the DLPNO-CCSD(T) method. These results will serve as a foundation for obtaining the reaction barriers required to gain insight into the kinetics of the oligomerization process in the Ni-NU-1000 catalyst.

3.3 Ethylene Oligomerization in Ni-NU-1000

In this section, the complete mechanism of ethylene oligomerization in Ni-NU-1000 up to the formation of oct-1-ene is presented for periodic PBE-D3(BJ) calculations (shortly, PBE), as well as for hierachical cluster corrections calculated with TPSSh and M06-L functionals. Two types of pores were considered — hexagonal and trigonal. The results for both pore types were analyzed. Based on the energy barriers of all oligomerization reactions, a simplified model was employed that includes only four stages. These four reactions were analyzed with respect to variations in the computational method as well as the types of pores. Finally, a microkinetic analysis was carried out for the simplified mechanism over a one-hour reaction time. The

changes in the concentrations of reactants and products, as well as their distribution for three methods and two pores, were examined and analyzed.

3.3.1 Overview of the Catalytic Cycle and Reaction Network

All Gibbs free energy barriers for ethylene oligomerization mechanism were computed using the same hierarchical cluster correction scheme proposed in the previous section. The periodic PBE calculations were conducted by Thanh-Nam Huynh and based on these structures again the correction with TPSSh and M06-L were performed by the author of this thesis. The same formate-cluster model extracted from periodic model was used (see Figure 11A). The computed Gibbs free energy profile (shown schematically as the complete mechamism in Figure 11B and the energy profile for TPSSh correction in Figure 12) describes ethylene oligomerization proceeding through the Cossee-Arlman mechanism in the Ni-NU-1000 catalyst.

The catalytic cycle proceeds through the following elementary steps: (I) ethylene adsorption on the Ni site, forming a π-complex; (II) coupling of the coordinated ethylene to the Ni site, yielding the first Ni–alkyl species (β-agostic complex); (III) adsorption of a second ethylene molecule to the Ni–alkyl intermediate, producing the [et-Ni-ethyl] π-complex, taken as the reference initial structure for consistency with the kinetic model discussed later; (IV) C-C coupling of ethylene and ethyl species, generating a longer Ni–alkyl species and completing one propagation step of the Cossee-Arlman mechanism; (V) repetition of steps (III)–(IV) via further ethylene adsorption and C-C coupling, leading to continued chain growth and the formation of higher α-olefins; (VI) alternatively, chain termination via β-H elimination from the Ni–alkyl intermediate, generating a [H-Ni-but] species with a coordinated α-olefin (e.g., but-1-ene); (VII) direct desorption of the coordinated olefin from the Ni site, releasing the product and leaving a highly unstable [Ni-H] species that is subsequently stabilized later by adsorption of an ethylene molecule; (VIII) alternatively, adsorption of ethylene to the post-elimination [but-Ni-H], promoting displacement of the coordinated α-olefin and forming the [et-Ni-but-H], thereby enabling re-entry into the propagation manifold; (IX) desorption of the coordinated α-olefin, regenerating the initial structure and completing the catalytic cycle. The full equations of each reaction can be found in Appendix for Chapter 3, Section 3.3.

The complete mechanism in Figure 11B leads to numerical instabilities when combined with the Eyring equations in a kinetic model because of barrierless and near-barrierless elementary steps yielding in nonphysically large rate constants[112]. To make it

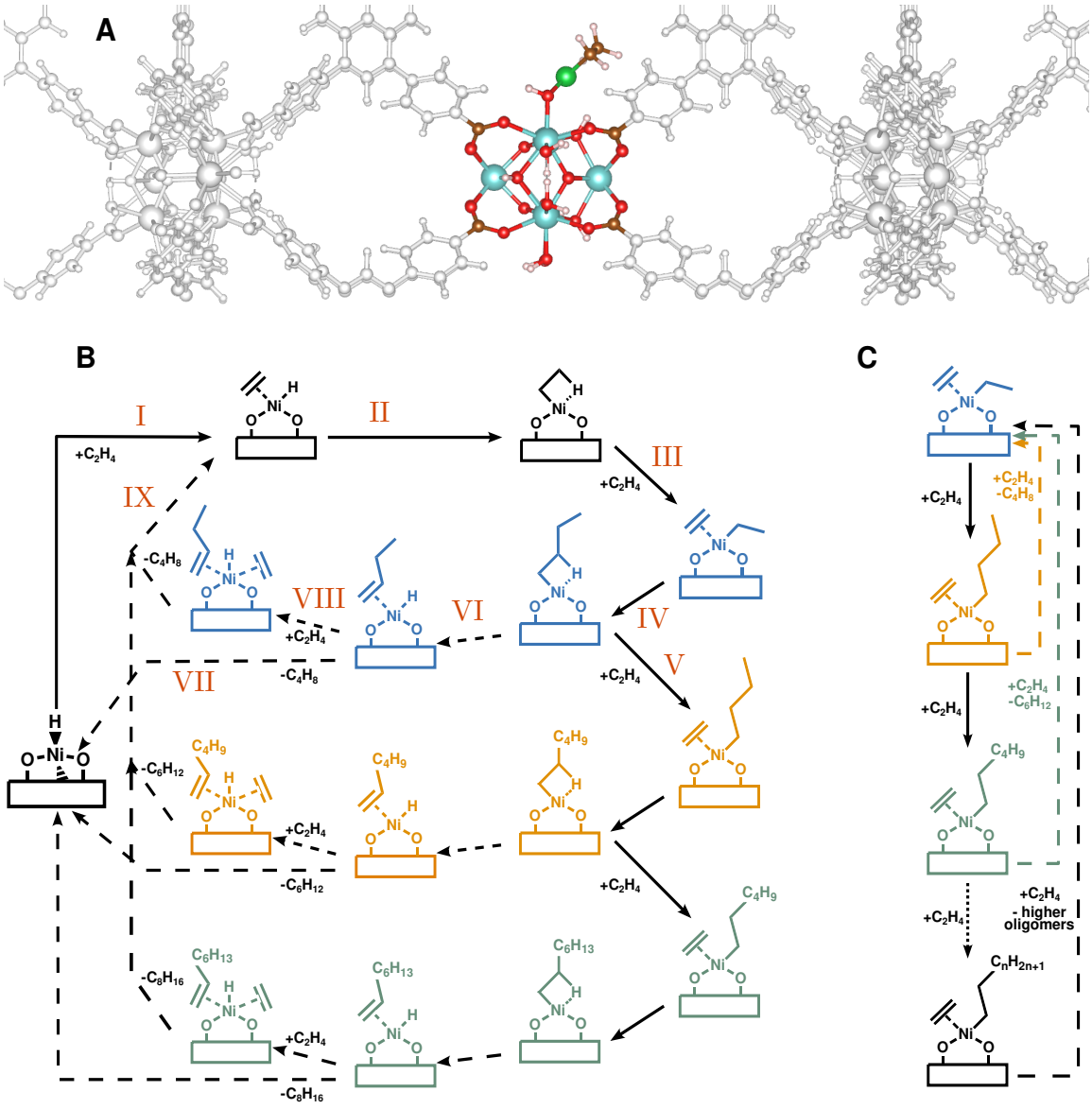


Figure 11: **A** Visual representation of the formate-cluster model within periodic structure of Ni-NU-1000. Atom colors in the cluster model are as follows: Zr (cyan), Ni (green), O (red), C (brown), and H (white). **B** Complete and **C** simplified catalytic reaction network for ethylene oligomerization in Ni–NU-1000. The complete network extends up to C_8H_{16} (C_8 or oct-1-ene) product formation. The simplified network includes olefin chains of up to oct-1-ene. In both **B** and **C**, solid arrows indicate chain growth reactions and dashed arrows indicate chain termination reactions. Reprinted with permission from [112]. Copyright 2026 Nature Publishing Group UK.

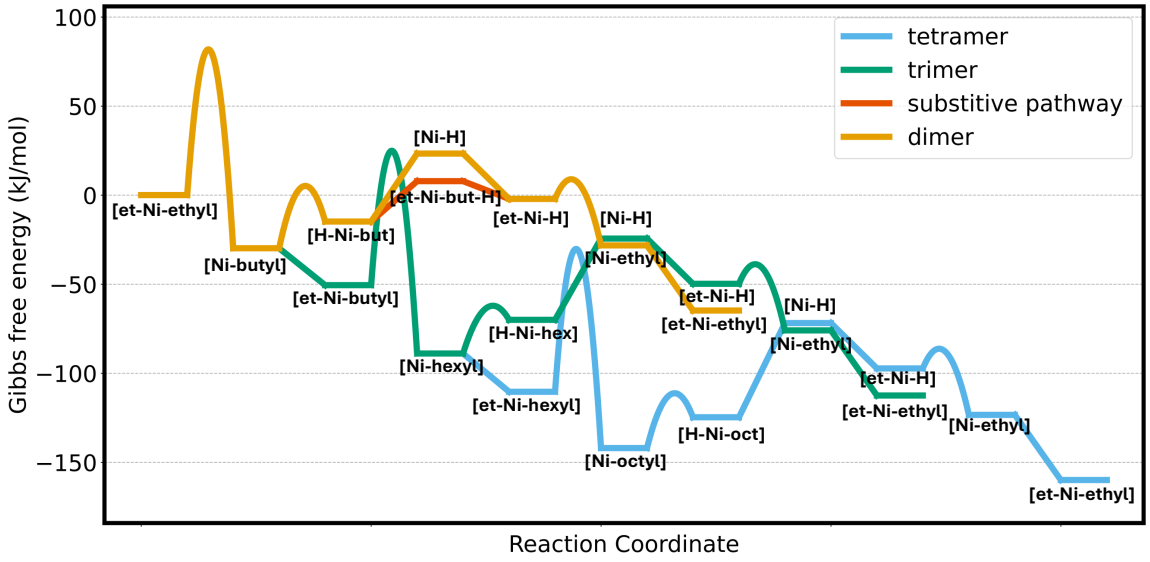


Figure 12: The Gibs free energy profile of the ethylene oligomerization pathway up to the tetramerization step (C_8) for TPSSh-corrected method in trigonal pore, computed at 300 K and 1 bar. For clarity, branching pathways are shown exclusively for the dimerization step. Reprinted with permission from [112]. Copyright 2026 Nature Publishing Group UK.

stable, Alexander Avdoshin from Institute of Nanotechnology in Karlsruhe Institute of Technology suggested a simplified reaction mechanism which contains two main steps as shown in Figure 11C:

(i) Chain growth via ethylene insertion and its reverse reaction (reverse migratory insertion or de-insertion), with rate constants r_{ins} and $r_{\text{de-ins}}$, respectively

(ii) Chain termination and its reverse reaction (de-termination), with rate constants r_{ter} and $r_{\text{de-ter}}$, respectively.

Additionally, r_{ter} was determined by selecting the highest Gibbs free energy transition state (i.e., the lower effective Gibbs free energy barrier) among the two competing pathways connecting the chain-growth and termination manifolds. Based on this approach, the microkinetic analysis was performed using Alexander Avdoshin's and Thanh-Nam Huynh's workflow[112].

3.3.2 Comparison of Activation Energy Barriers in the Simplified Mechanism

Before performing the microkinetic analysis, the four reaction steps of the simplified mechanism were analyzed for three methods for the trigonal pore. Figure 13a shows the barriers of the simplified mechanism obtained with three methods (PBE, TPSSh-

corrected, and M06-L-corrected) for the three product reactions: dimerization (C_4), trimerization (C_6), and tetramerization (C_8). The insertion barriers obtained with PBE and TPSSh are very close, with TPSSh yielding slightly higher values, whereas M06-L predicts the highest insertion barriers for all three product sizes. For reverse insertion, PBE behaves similarly to M06-L and remains close to the TPSSh values. Termination shows the largest deviation among the methods: M06-L lies clearly outside the range of the other two, predicting much higher barriers, while PBE and TPSSh are in close agreement. Interestingly, TPSSh values exceed those of PBE only starting from C_6. The de-termination barriers follow the same size trend for TPSSh and M06-L, with M06-L values being higher than TPSSh by about 5 kJ/mol. PBE yields values close to TPSSh for all product sizes except for C_4.

The comparison between the two different pores (trigonal and hexagonal) was carried out using TPSSh-corrected activation free energy barriers. It is shown in Figure 13b. For the insertion barrier, the trigonal pore is slightly higher for C_4, but for C_6 and C_8 the hexagonal pore exhibits larger barriers, making it less favorable for chain growth. A similar trend is observed for reverse insertion, although the difference between the two pores is smaller in this case. The termination barriers for C_4 and C_6 are considerably higher in the hexagonal pore, while for C_8 the two pores reach nearly the same value, with the hexagonal pore being only marginally lower. For the de-termination barriers, the values in the trigonal pore are close to each other and consistently lower than in the hexagonal pore, where the C_6 and C_8 barriers are higher than that of C_4.

Overall, these results with TPPSh indicate that longer chains preferentially grow in the trigonal pores due to the lower insertion barriers, even though the corresponding reverse-insertion barriers are on average substantially higher. In addition, the termination barriers in the trigonal pore are lower than in the hexagonal pore for C_4 and C_6, and comparable for C_8. For C_4 and C_6, the de-termination barriers are higher in the trigonal pore, whereas in the hexagonal pore only the C_4 termination barrier lies below the corresponding de-termination barrier. Similar observations have been reported in the literature for NU-1000[131, 134], supporting our finding that long olefin chains preferentially reside in the trigonal pores, as this pore type effectively traps the olefins within its confined environment.

3.3.3 Microkinetic Analysis of Product Distribution

As already mentioned above, the microkinetic analysis was carried out using the framework developed and described in previous study[112]. For simplicity, the con-

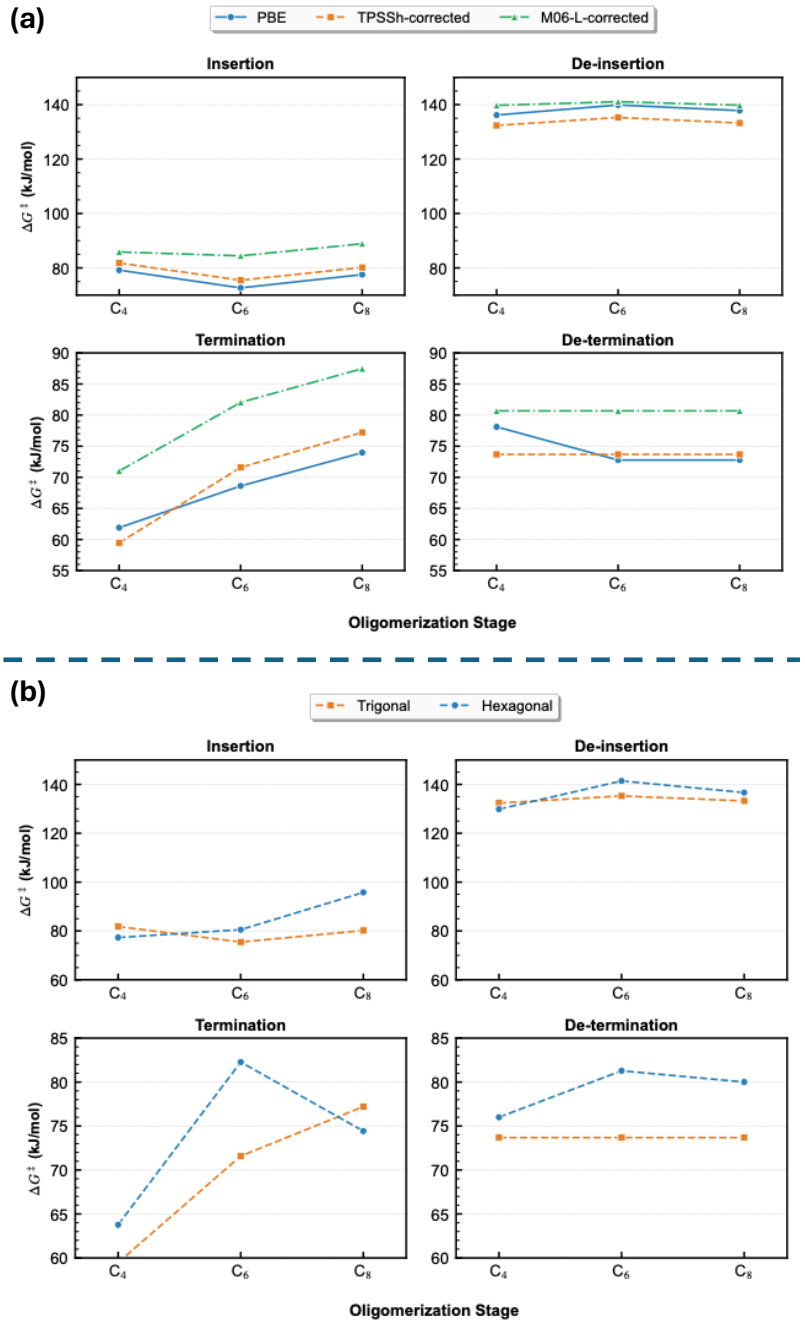


Figure 13: Activation free energies (in kJ/mol) at 300 K and 1 bar for the simplified ethylene oligomerization mechanism in Ni–NU-1000, comparing pathways with growing carbon chains corresponding to dimerization (C_4), trimerization (C_6), and tetramerization (C_8), all corrected values were obtained within the hierarchical cluster correction scheme.
(a) for trigonal pore with different methods: periodic PBE values are shown as blue points and serve as the baseline reference. Other energy barriers were obtained with the TPSSh (orange points) and M06-L (green points) functionals.
(b) for trigonal (orange points) and hexagonal (blue points) pores with TPSSh functional.

ditions $T = 300$ K and $p = 1$ bar were chosen. Under these conditions, the kinetic evolution of the reactants and products was simulated over a reaction time of 1 hour (3600 s). For further simplification, no additional reactants are introduced during the microkinetic simulation. The initial concentrations of all species except ethylene were set to zero, while the initial ethylene concentration was approximately 13 mol per unit cell. After 1 hour of simulated reaction time, the product distributions are reported as percentages, representing the relative formation probabilities.

The first comparison was performed between the different methods for the trigonal pore. As discussed above, PBE and TPSSh are expected to behave similarly, since their barriers are comparable for most steps, with the exception of some termination and de-termination barriers. In contrast, olefin formation with M06-L is expected to be hindered, owing to its substantially higher termination barriers as well as its slightly higher insertion barriers.

Figure 14a-c shows the results of the kinetic modeling for the three methods. PBE (Figure 14a) predicts an almost complete conversion of ethylene: in the final product distribution only about 2% of the total distribution remains as ethylene. But-1-ene reaches its peak concentration at around 500 s and then gradually decreases toward a limiting value of approximately 1 mol per unit cell. Hex-1-ene reaches its plateau at about 1000 s and remains essentially unchanged until the end of the simulation. In the final mixture, these two products account for roughly one quarter of the total mass each (25% and 23%, respectively). The dominant product is oct-1-ene, which grows steadily throughout almost the entire reaction time and reaches its plateau only toward the end. It accounts for nearly half of the final product distribution (49%).

In line with the high activation barriers discussed earlier, the kinetic modeling with M06-L shows that ethylene is barely converted into longer α-olefins (Figure 14b). The combined fraction of C_4–C_8 products in the final mixture amounts to only about 11%. It is worth noting, however, that under these conditions a significantly longer reaction time would be required to reach the corresponding concentration plateaus.

Finally, the results obtained with TPSSh were examined (Figure 14c), which turned out to lie in between those of PBE and M06-L. For instance, the conversion of ethylene proceeds much faster and reaches its plateau by the end of the 1 hour simulation in contrast to M06-L. The final ethylene concentration is about 2 mol per unit cell, corresponding to 37% of the final product distribution. The but-1-ene concentration reaches its maximum at around 1000 s and then slowly decreases, accounting for only 15% in the final distribution. Hex-1-ene approaches its plateau slowly and represents

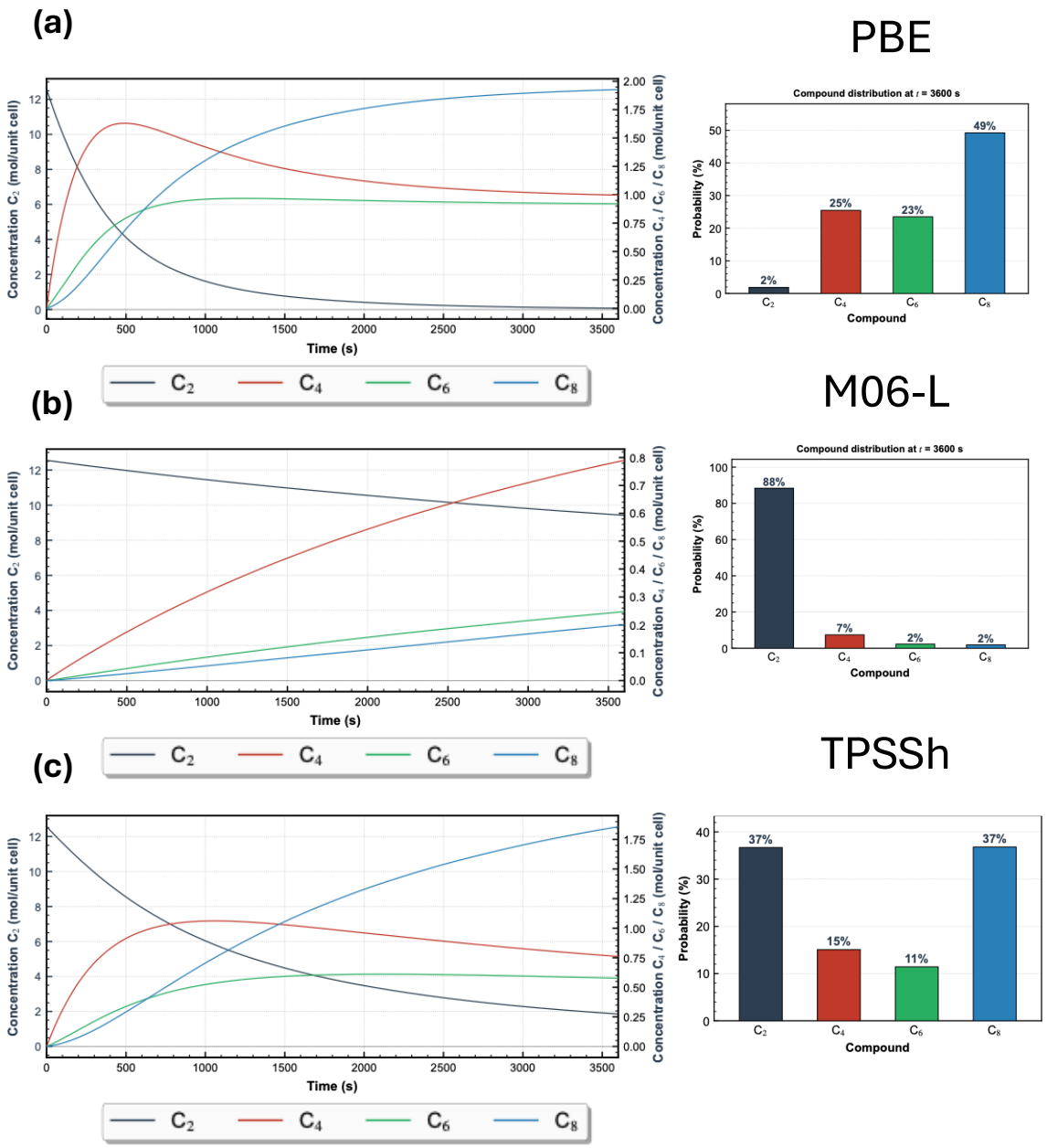


Figure 14: Time evolution of α-olefin concentrations: ethylene (C_2), but-1-ene (C_4), hex-1-ene (C_6), and oct-1-ene (C_8) (molecules per unit cell) at $T = 300$ K and $p = 1$ bar in a Ni-NU-1000 catalyst model over 1 hour (3600 s), and product distribution (mol%) after 1 hour of reaction. Reaction rates are derived from activation Gibbs free energy barriers computed using (a) periodic PBE, (b) M06-L-corrected, and (c) TPSSh-corrected methods. Only the trigonal pore is considered.

11% of the final mixture. By contrast, oct-1-ene grows faster than the other products starting from approximately 600 s and ultimately accounts for as much as 37% of the final mixture, which is in line with the behavior observed for the PBE functional.

In summary, even the relatively small differences in the activation barriers between PBE and TPSSh, compared to M06-L, lead to substantial differences in the simulated reaction kinetics. M06-L, in particular, deviates markedly from the trends observed with the other two methods. The choice of correction was guided by the available reference data, and that the two correction methods are in fact rather close to each other (TPSSh and M06-L show a mean absolute deviation of about 3–4 kJ/mol), with only minor discrepancies in the dimerization barriers. The differences accumulated for the subsequent oligomerization steps then translate into much larger differences in the final product distributions.

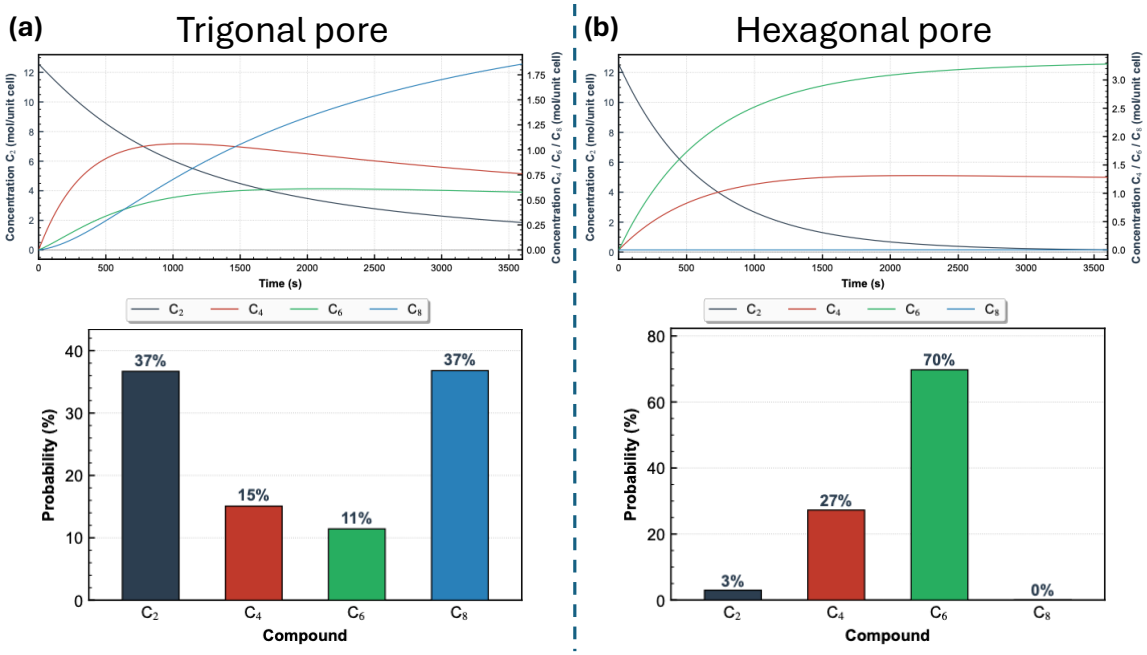


Figure 15: Time evolution of α-olefin concentrations: ethylene (C_2), but-1-ene (C_4), hex-1-ene (C_6), and oct-1-ene (C_8) (molecules per unit cell) at $T = 300$ K and $p = 1$ bar in a Ni-NU-1000 catalyst model over 1 hour (3600 s), and product distribution (mol%) after 1 hour of reaction. Reaction rates are derived from activation Gibbs free energy barriers computed at the TPSSh-corrected level of theory. Two types of pore are considered: (a) trigonal pore, (b) hexagonal pore.

The following step was to investigate, using the two different pore types, their influence on the time evolution of the product distribution. The same conditions as in the previous kinetic simulations were employed. The results are shown in Figure 15. The

behavior of the trigonal pore, obtained at the TPSSh level, was already discussed above (Figure 14c and Figure 15a are the same). It is characterized by an incomplete conversion of ethylene, mainly into oct-1-ene, whose concentration has not yet reached its plateau by the end of the simulation (indicating that further growth is possible beyond 1 hour of reaction time).

When compared with the hexagonal pore (Figure 15b), the differences in the barrier heights discussed earlier become clearly visible (Figure 13b): the insertion barriers are considerably higher starting from C_6, and the termination barriers for C_4 and C_6 are also higher than those in the trigonal pore (Figure 15a). Looking at the time evolution of the concentrations, almost all ethylene is consumed in the hexagonal pore, so that at the end of 1 hour it accounts for only about 3% of the total product distribution. But-1-ene reaches its plateau at around 1500 s and remains essentially constant afterwards, representing 27% of the final distribution. Interestingly, owing to the high barriers in the hexagonal pore, oct-1-ene is not formed at all (0% in the final distribution). Instead, hex-1-ene, which becomes the dominant product of the reactor already after approximately 500 s, accounts for 70% of the mixture after 1 hour of reaction. Experimental studies have reported the formation of products heavier than hex-1-ene, which clearly supports the conclusion that the hexagonal pores are not catalytically active for the formation of α-olefins larger than hex-1-ene[131, 203].

3.3.4 Discussion

The complete mechanism was computed using several methods: periodic PBE, TPSSh-corrected, and M06-L-corrected methods. Based on the simplified mechanism, all insertion and termination barriers (as well as their reverse reactions) were compared across the three methods and two distinct pore types. It was demonstrated that M06-L deviates substantially from the other methods, owing to considerably higher insertion and termination barriers, which is expected to result in lower yields of heavier olefins. Mostly, TPSSh performed comparably to periodic PBE, with the notable exception of the de-termination reaction. A comparison of the two pore types using TPSSh-corrected values revealed that nearly all barriers in the hexagonal pore were higher than those in the trigonal pore, particularly, for the termination reaction, suggesting that olefin formation would likewise be hindered in this pore geometry.

Subsequently, microkinetic analysis was performed on the basis of these barriers using automated workflow, simulating reactions at T = 300 K and p = 1 bar over the course of one hour, for all three methods and both pore types. The results showed that in the

case of M06-L, virtually no ethylene conversion takes place. With periodic PBE, nearly complete conversion of ethylene to olefins was observed, with C_8 being the dominant product (49%), whereas with TPSSh, only slightly more than half of the ethylene was converted, with a C_8 fraction of 37%. A comparison of the two pore types using TPSSh indicated that in the hexagonal pore, ethylene was almost entirely converted to C_4 (27%) and C_6 (70%) with no C_8 detected. This result is clearly inconsistent with literature data[131, 134, 203], whereas in the trigonal pore, C_8 accounted for 37% of the final product mixture. It is important to note that this microkinetic analysis was conducted solely for the formation mechanism of α-olefins up to C_8, and does not account for diffusion of components within the reactor.

3.4 Chapter Conclusions

This chapter was devoted to the development of the hierarchical cluster correction scheme for periodic DFT energies, aimed at achieving a more accurate description of all intermediates and transition states of the Ni-NU-1000 active site in ethylene dimerization reactions. The scheme was subsequently applied to the complete ethylene oligomerization mechanism up to oct-1-ene formation in order to perform the microkinetic analysis.

First, the hierachical cluster correction scheme based the DLPNO-CCSD(T) method was constructed as the reference and shown to yield energetics that differ significantly from uncorrected periodic PBE results (i.e., the transition state MAD equals 20 kJ/mol). A subsequent benchmarking of various DFT functionals identified the best-performing methods with the smallest deviations from the DLPNO-CCSD(T) reference — TPSSh and M06-L. A comparison of different cluster sizes was also carried out, revealing minimal sensitivity to cluster size, particularly for TPSSh.

Then these two methods were employed to compute the corrected complete ethylene oligomerization mechanism. Using a simplified mechanistic scheme, reaction barriers were calculated and compared across different computational methods and pore types in Ni-NU-1000. Microkinetic simulations of the reaction over one hour were subsequently performed on the basis of these barriers, and the resulting product distributions were compared. It was found that M06-L predicts essentially no ethylene conversion, whereas periodic PBE leads to complete conversion to α-olefins, and TPSSh to partial conversion; notably, both PBE and TPSSh show a predominance of the C_8 product. A separate comparison of the two pore types revealed that no

products beyond C_6 are formed in the hexagonal pore, whereas in the trigonal pore oct-1-ene is the dominant product in the distribution. It should be noted that the present microkinetic model is restricted to the formation of linear α-olefins up to C_8 and does not account for adsorption capacity and diffusion or competing side reactions, which may affect quantitative predictions.

To discuss the final results of corrected values, TPSSh yielded MADs for five energy barriers in ethylene dimerization of approximately 3 kJ/mol relative to the DLPNO-CCSD(T) reference, with a maximum cluster-size sensitivity of 2 kJ/mol. Microkinetic simulations with TPSSh predicted 50% ethylene conversion with 37% C_8 selectivity in the trigonal pore.

These results are of considerable significance, as they demonstrate the extent to which the choice of computational method and active site model can ultimately influence the observable outcomes in a kinetic simulation. For instance, the product distribution changes drastically depending on the method employed. Furthermore, the predicted difference in pore reactivity aligns with experimental and theoretical evidence for preferential chain growth within the trigonal pores of Ni-NU-1000[131, 134]. In future work, the availability of experimental data could enhance the predictive power of the simulations and guide hypothesis selection for the computational modeling of the complete catalytic cycle of Ni-NU-1000 and other catalytic materials.

4 Non-Covalent Interactions of PAHs and their Derivatives

Overview of Chapter

In this chapter, a theoretical study of PAHs and their derivatives non-covalently bound dimers is presented. In Section 4.1 an introduction to polycyclic aromatic hydrocarbons and their binding via non-covalent interactions is presented, along with the main objectives of the chapter and the open questions to be addressed. In Section 4.2 the results for acene dimers are presented, with a detailed analysis of their various properties. The stability for unique acene dimer type and acene dimer properties were obtained in collaboration and published together with the co-first author Mikhail Kalinin[77], where the author of this thesis performed quantum chemical investigation of acene non-covalent complexes (Publication 1, p. 175). In Section 4.3, the study of acene dimers is extended to the case of F- and BN-substituted anthracene monomers, with particular emphasis on dimer stability and the effects of substitution on optical properties. In Section 4.4, the trends in relative stability and other properties that were observed for acene dimers are extended to larger acenes with the number of linearly fused rings ranging from five up to ten (decacene). In Section 4.5, different geometries of *cata*-condensed polybenzoids are investigated to identify trends in the strength of $\pi - \pi$ stacking and its connection with electronic and structural properties of the monomer.

4.1 Introduction

4.1.1 General Overview

Non-covalent interactions (NCIs) play a significant role in determining the structural, thermodynamic, and dynamic properties of molecular[204, 205]. NCIs are forces, such as dispersion or electrostatics, which collectively define the stability and function of

supramolecular architectures[206]. These forces dictate, for example, the aggregation of polycyclic aromatic hydrocarbons in soot particles[207], organic semiconductor structures[208], DNA interactions[209, 210], and other phenomena[211, 212]. Their subtlety makes them pivotal in supramolecular chemistry and materials design[208, 213, 214]. Therefore understanding the nature and strength of these interactions is essential for rationalizing how molecular organization can be tuned[215, 216].

PAHs constitute prototypical systems for exploring $\pi - \pi$ interactions because of their extended π-system conjugation and mostly planar geometry[217, 218]. In PAHs, NCIs govern processes ranging from molecular recognition, self-assembly, and crystal packing to charge transport phenomena[219–226]. The interplay between NCI forces, such as dispersion attractions and electrostatics in PAHs drives their aggregation behavior in solution[227–231] and in the condensed phase[232–235]. Moreover, in materials science and catalysis, controlled $\pi - \pi$ interactions between PAH motifs dictate the electronic coupling and charge transferability within organic semiconductors and molecular surfaces[236–238].

$\pi - \pi$ stacking can be easily found in the crystal packing of PAHs and their derivatives[239, 240], where parallel-displaced or herringbone geometries minimize steric clashes while maximizing orbital overlap and dispersion energies[241, 242]. In organic field-effect transistors (OFETs), these interactions enhance charge carrier mobility by facilitating tight π-stacked arrangements, as seen in pentacene thin films[243]. Similarly, in organic light-emitting diodes (OLEDs), $\pi - \pi$ contacts between emissive PAH cores promote efficient exciton migration and energy transfer, boosting device efficiency[244]. Moreover, MOFs with aromatic linkers, such as those incorporating naphthalene[245], anthracene[246, 247], pyrene[248, 249], or pentacene[250] motifs, leverage interlayer $\pi - \pi$ stacking to tune porosity, guest adsorption, mechanical stability, and optical properties.

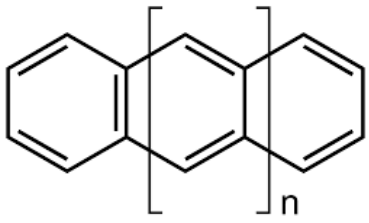


Figure 16: General view of an acene structural formula for n+2 fused rings.

Acenes (Figure 16) occupy a special place among PAHs and are of great interest as tunable materials for applications in organoelectronics and spintronics[251–254].

These molecules serve as ideal model systems for analyzing $\pi - \pi$ interactions from naphthalene to pentacene and beyond[251, 255]. This is due to their structural simplicity and rigidity, connection of fused benzene rings and tunable size, which systematically varies HOMO-LUMO gaps and stacking energetics[251, 256–258]. Most acene crystals adopt a herringbone structure[259–262], and only functionalized acenes exhibit other motifs[240, 263–271], such as glike-, cross-shaped, and parallel-shaped (Figure 17). Headen *et al.*[272] showed that in the liquid phase the parallel-displaced configuration is dominant for naphthalene. All these crystal packing types can be represented as dimer structure.

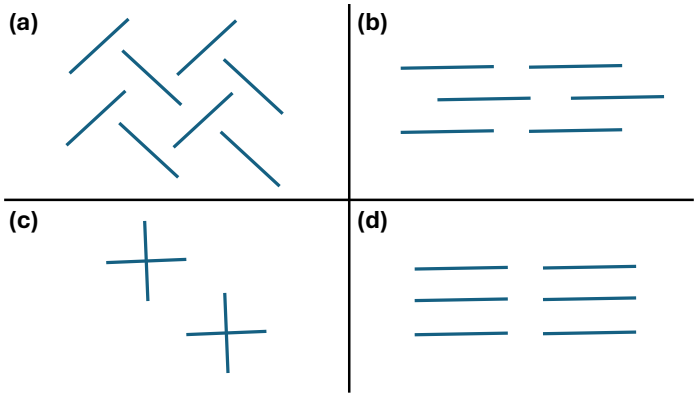


Figure 17: Abstract view of various crystal packing types for acenes and their derivatives: (a) herringbone; (b) glike; (c) cross-shaped; (d) parallel-shaped.

From a theoretical perspective, different dimers of acene molecules have been explored, including glike-, cross-[270], and parallel- or stacked-dimers[26, 32, 273–277] (Figure 18). One of the theoretical studies in this area was carried out by Cabaleiro-Lago *et al.*[32], who provided a systematic investigation of various dimers of aliphatic and aromatic polycyclic hydrocarbons. They also offered insights into the contributions of $\sigma - \sigma$, $\sigma - \pi$, and $\pi - \pi$ interactions to the dimer formation using SAPT[32], and demonstrated that the glike-dimer is the most stable with increasing size. Importantly, different acene dimers and their derivatives can exhibit unique electronic and optical properties[278, 279]. While dimers, such as parallel-, glike-, and slipped- configurations, are commonly studied in a systematic way[26, 32, 273–277], one particular type of dimer (hereafter referred to as the X-dimer) has received limited attention and has been limitedly reported in some studies[20, 25, 33, 34, 280–284]. Its stability trends with acene size remain unexplored. Most calculations carried out at the DFT level with dispersion corrections[25, 282–284], some employing semi-empirical[282, 284] methods. Naphtalene is the only one molecule for which a CCSD(T) reference for dimers is available[20, 281]. The energy differences between different dimers are small, which

renders methods such as DFT and semi-empirical approaches insufficiently accurate. Regarding optical properties, mostly monomeric acenes have been studied[285–287], while dimers have received only limited attention[284, 288].

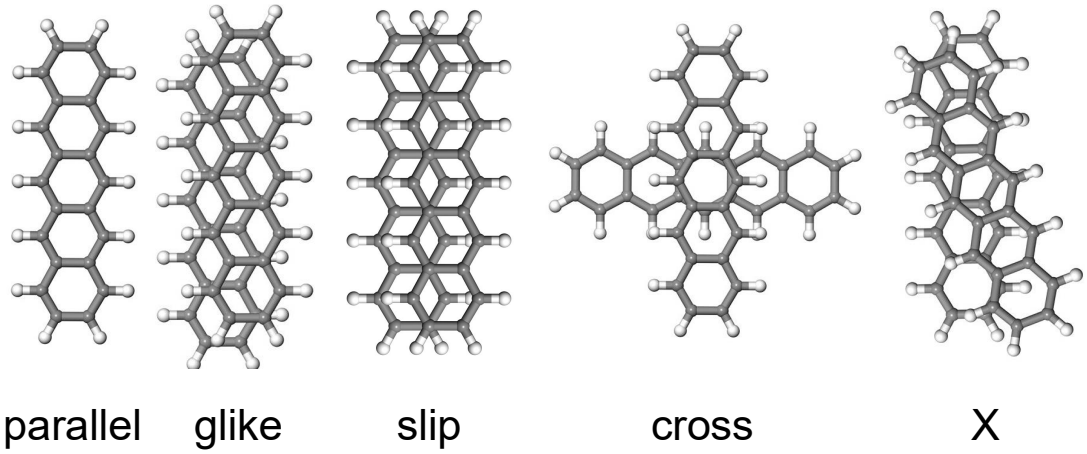


Figure 18: Top view on selected dimer structures for pentacene system.

However, with increasing acene size, the reactivity of this class of compounds also increases, and the open-shell configuration becomes energetically more favorable compared to the closed-shell one, leading to rapid chemical decomposition[254, 255, 289–291]. Several approaches have been employed to improve stability, including the attachment of side fused rings to block the radical centres of the molecule[292–294], the incorporation of different substituents[295–297], and the incorporation of heteroatoms into the acene core[298–301]. Consequently, all of these strategies inevitably perturb the electronic structure of the molecule, thus altering its properties and behavior in non-covalent complexes[302]. Indeed, several studies have demonstrated that substitution of acenes with BN or fluorination systematically modulates the stacking geometry, interaction energies, and frontier orbitals, thereby affecting the optical gaps, exciton character, and thermodynamic stability of non-covalent dimers and crystalline assemblies[303, 304]. Varying the orientation of fused rings within the acene framework gives rise to a broad range of structural isomers, which may exhibit distinctly different electronic properties and non-covalent interaction characteristics in their complexes[293, 305, 306].

This chapter aims to address the following questions:

(i) Which dimer type (the glike-dimer against the X-dimer) is more stable for acenes.

(ii) How does dimer interaction scale with the acene size and does it correlate with the contact surface area of the NCI. What the effect is on the electronic and optical properties.

(iii) How do chemical modifications (fluorination and BN-substitution) affect dimer stability and optical properties.

(iv) Up to what acene size do the established trends of electronic and thermodynamic properties hold (extended acenes).

(v) Do the observed trends satisfy to *cata*-condensed polybenzoids, and which monomeric properties can help to estimate the stability of the most stable non-covalent dimers.

4.1.2 Computational Details

Geometry Optimization and Interaction Energy Calculations

The DFT geometry optimization and single-point energy calculations have been carried out using ORCA 5.0.4 quantum-chemical package[66, 307] with the B3LYP exchange-correlation functional[191] with Grimme's D3 dispersion correction with Becke-Johnson damping[100, 103] and the def2-TZVP basis set[71]. The same method was used for electronic properties calculation.

The interaction energy ($E_{int} = E_{dimer} - 2E_{monomer}$) of a non-covalent dimer is defined as the energy difference between this dimer and its isolated constituent molecules, providing a quantitative measure of the strength of non-covalent binding. The interaction energy must be negative. A dimer with a smaller values of interaction energy is more stable than one with a larger values of interaction energy. The CP-correction formula[308] was used in all calculations of the interaction energy to reduce the effect of the BSSE[309]. The counterpoise (CP)-corrected interaction energy has the following formula:

$$E_{\text{int}} = E_{\text{dimer}} - E^{\text{unrelaxed}}_{\text{monomer 1,ghost}} - E^{\text{unrelaxed}}_{\text{monomer 2,ghost}}, \tag{61}$$

where E_{dimer} is the total energy of considered dimer and $E^{\text{unrelaxed}}_{\text{monomer 1,2,ghost}}$ is the total energy of considered monomer in its structure in the dimer as obtained with the basis of that monomer and the ghost basis of the other monomer.

To calculate the interaction energy and analyze the different interaction energy contributions of dimers the sSAPT0/jun-cc-pVDZ method[76] was used in Psi4[310, 311] through the perturbative SAPT approach (see Section 2.1.3). The SAPT analysis contains the *exchange-repulsion* (E_{exr}), the *induction* (E_{ind}), the *electrostatic* (E_{elst}), the *dispersion* (E_{disp}) contributions, and the total *interaction energy* (E_{int}). It was shown that for a system the size of pentacene, it is not feasible to increase the level of SAPT calculations due to the complexity of computational scaling. The interaction energy calculations for naphthalene, anthracene and tetracene dimers were done by Andrea Buchwald and Professor Reinhold F. Fink from the University of Tübingen.

Coupled Cluster Calculations

To validate the results of DFT and SAPT calculations, the calculations were done using PNO-LCCSD(T*)-F12[59, 62, 63, 312–315] level of theory in Molpro code[316, 317]. A composite basis set was used, composed of aug-cc-pVTZ for C atoms and cc-pVTZ for H atoms[318]. All calculations of the CP-corrected interaction energy of acene dimers with the PNO-LCCSD(T*)-F12 method have been performed by Professor Hans-Joachim Werner of the University of Stuttgart.

Performance of Electronic Structure Methods for Interaction Energy Benchmarking

To benchmark different methods, especially DFT functionals, interaction energy calculations of selected acene dimers have been done to compare with reference values (PNO-LCCSD(T*)-F12) for the interaction energy. Primarily, an objective was to benchmark the performance of different DFT methods with dispersion correction, using ORCA 5.0.4, ranging from GGA to *meta*-GGA and hybrid functionals. The following functionals were employed: PBE[86], M06-L[189], TPSS[87], TPSSh[89], CAM-B3LYP[319], B3LYP[191], PBE0[190], M06[193], and M06–2X[194]. Almost all functionals have been used with the D3 dispersion correction including Becke–Johnson damping function[104, 105], M06-family functionals have been used with D3zero dispersion correction[100]. Moreover, for two functionals TPSS and B3LYP, calculations were performed not only with D3(BJ) but also with D4 dispersion correction[106]. In addition, calculations were performed with the functional ωB97M-V[320] and semi-empirical GFN2-xTB[321], both using ORCA 6.1.0 code[322]. All DFT calculations for benchmarking were performed with the def2-TZVP basis set[71]. In addition to the DFT and semi-empirical methods, two neural network potentials (NNPs) were also used for interaction energy calculations: the small-conserving NNP based on Meta's equivariant Smooth Energy Network (eSEN) architecture, trained on the

Open Molecules 2025 dataset (OMol25's eSEN Cons.Sm.)[323–325], and the "small" version of the NNP "Universal Model for Atoms", also trained on OMol25 (UMA Small (OMol))[326]. Unlike DFT methods, for the NNP-based methods a correction for BSSE is not needed. For these methods and semi-empirical GFN2-xTB, the interaction energy is defined as:

$$E_{\text{int}} = E_{\text{dimer}} - E^{\text{unrelaxed}}_{\text{monomer 1}} - E^{\text{unrelaxed}}_{\text{monomer 2}}, \tag{62}$$

where E_{dimer} is the total energy of the considered dimer and $E^{\text{unrelaxed}}_{\text{monomer 1}}$ and $E^{\text{unrelaxed}}_{\text{monomer 2}}$ are the total energies of the respective monomers in the structure of the dimer.

To quantitatively assess the performance of the tested methods, two statistical descriptors were used: the mean absolute deviation, which measures the average magnitude of the errors regardless of sign, and the mean signed deviation, which captures the systematic bias of a method. A negative MSD indicates a systematic overestimation of the interaction energy ($E^{Method}_{int,i}$) relative to the reference ($E^{Referenece}_{int,i}$), while a positive MSD indicates overestimation. Both quantities are defined as follows, where N is the number of complexes considered and i is each acene system:

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N}\left|E^{Method}_{int,i} - E^{Reference}_{int,i}\right|, \tag{63}$$

$$\text{MSD} = \frac{1}{N}\sum_{i=1}^{N}\left(E^{Method}_{int,i} - E^{Reference}_{int,i}\right). \tag{64}$$

UV-Vis Spectra Calculations

TD-DFT calculations[108, 109] without Tamm-Dancoff approximation using the CAM-B3LYP functional[319] with the def2-TZVP basis set[71] were performed using ORCA 5.0.4[307] to simulate the UV-Vis spectra and first excitation energy (E_{exc}). To validate the reliability of this approach, calculations with the TD-CAM-B3LYP level of theory were additionally carried out using many-body Green's function theory (evGW/BSE)[327–332], RI-ADC(2)[333], and RI-CC2[334, 335] in Turbomole 7.6[336]. The experimental measurements of fluorescence and UV-Vis spectra have been executed by Mikhail Kalinin from the Martin-Luther-University Halle-Wittenberg to compare with theoretical results and give an interpretation of anthracene spectra in solution.

Evaluation of Database with *cata*-Condensed Polybenzoid Dimers

Finally, conformational searches were performed using the CREST 3.0.1 code[337] with default parameters and NCI mode for a series of *cata*-condensed polybenzoid dimers to identify candidate structures for the most stable dimer geometries with GFN2-xTB for **the COMPAD Project** database. As the initiation structure, a dimer in a parallel-stacked configuration, separated by a distance of 3.4 Å, was used. The six lowest-energy conformers obtained from each search were subsequently re-optimized at the B3LYP-D3(BJ)/def2-SVP level of theory, after which the most stable structure was selected and further re-optimized at the higher B3LYP-D3(BJ)/def2-TZVP level of theory. The final properties were obtained with B3LYP-D3(BJ)/def2-TZVPP and the first vertical excitation energy with TD-CAM-B3LYP/def2-TZVP.

4.2 Non-Covalent Interactions in Acene Dimers

This section constitutes the methodological foundation of this chapter, establishing all baseline results upon which subsequent sections build. The focus is on acene molecules and two distinct dimer configurations: the glike-dimer, which has been extensively studied in the literature, and the X-dimer, which has received comparatively less and no systematic attention. Using SAPT, the individual contributions to dimer stabilization are systematically decomposed and analyzed. The relative stability of the two dimer types has established for pentacene and then traced across the acene series, from naphthalene to pentacene, revealing clear trends in both the total interaction energy and the preferential stabilization of one motif over the other. Beyond the energetics of dimer formation, this section also provides a comparative analysis of the electronic and optical properties of both dimers relative to the corresponding monomer. Changes in the HOMO–LUMO gap (E_{H-L}), frontier orbital character and other electronic properties, as well as UV-Vis excitation energies upon dimerization are examined using DFT and TD-DFT methods.

4.2.1 SAPT Interaction Energy Surface Screening of Selected Pentacene Dimers

To verify the existence of the X-dimer configuration reported in the literature[25, 33, 34, 280, 282–284], the interaction energy surface as an analogue of PES was explored for pentacene system to determine whether this configuration corresponds to an interaction energy minimum and to evaluate its stability relative to the glike-

dimer, reported as the most stable dimer by Cabaleiro-Lago *et al.*[32]. The glike-dimer structure was constructed using the following procedure:

(i) The geometry-optimized pentacene monomer at the B3LYP-D3(BJ)/def2-TZVP level of theory was used as the starting structure.

(ii) The second pentacene monomer was placed in parallel at an intermolecular distance of 3.49 Å (Z-axis).

(iii) it was then displaced by 1.45 Å along the X- and Y-axes to repeat the glike-dimer structure from [32].

Then, a screening of the sSAPT0 interaction energy surface with respect to the rotation and translation of one monomer relative to the other was performed for the pentacene dimer (see Figure 19). A total of 100 points were sampled: 10 points corresponding to the translational displacement of one monomer relative to the other (from 0 Å to 1.45 Å along the X- and Y-axes), and 10 points corresponding to the rotational angle about the axis perpendicular to the molecular plane (from 0° to 90°). The boundary points of the surface correspond to distinct dimers: the parallel or stack-dimer (0°; 0 Å), the glike-dimer (0°; 1.45 Å), and the cross-dimer (90°; 0 Å) (structures and their corresponding rotation angles and displacements are shown in Figure 20). As was shown in [32], the stack-dimer is the most unstable one that confirms with current interaction energy surface (maximum). Importantly, an X-dimer minimum was identified. This screening indicates that the most stable orientations, corresponding to the lowest E_{int}, are the glike-dimer and the X-dimer (see Figure 19). While the glike-dimer lies at the boundary of the interaction energy surface, the X-dimer is located in the region of the second E_{int} minimum, with the second minimum at a rotation of 20° and a displacement of 1.13 Å (Figure 20). It is a highly constrained geometry optimization procedure and only grid points without exact minima, but the glike-dimer has slightly lower E_{int} than the X-dimer (-110.70 versus -109.15 kJ/mol).

It is important to note that the induction energy (E_{ind}) follows the same trend as the total interaction energy. However, its magnitude does not significantly contribute to the total interaction energy, as the variation in E_{ind} across the surface is approximately 10 kJ/mol. The electrostatic (E_{elst}) and dispersion (E_{disp}) contributions exhibit opposite trends to that of the exchange-repulsion (E_{exr}). E_{exch} reaches its repulsion maximum at small rotational angles due to increased orbital overlap between the monomers, while E_{elst} and E_{disp} reach their minimum values at the same geometry. The effect of translational displacement on these contributions (E_{exr}, E_{elst}, and E_{disp}) is considerably smaller.

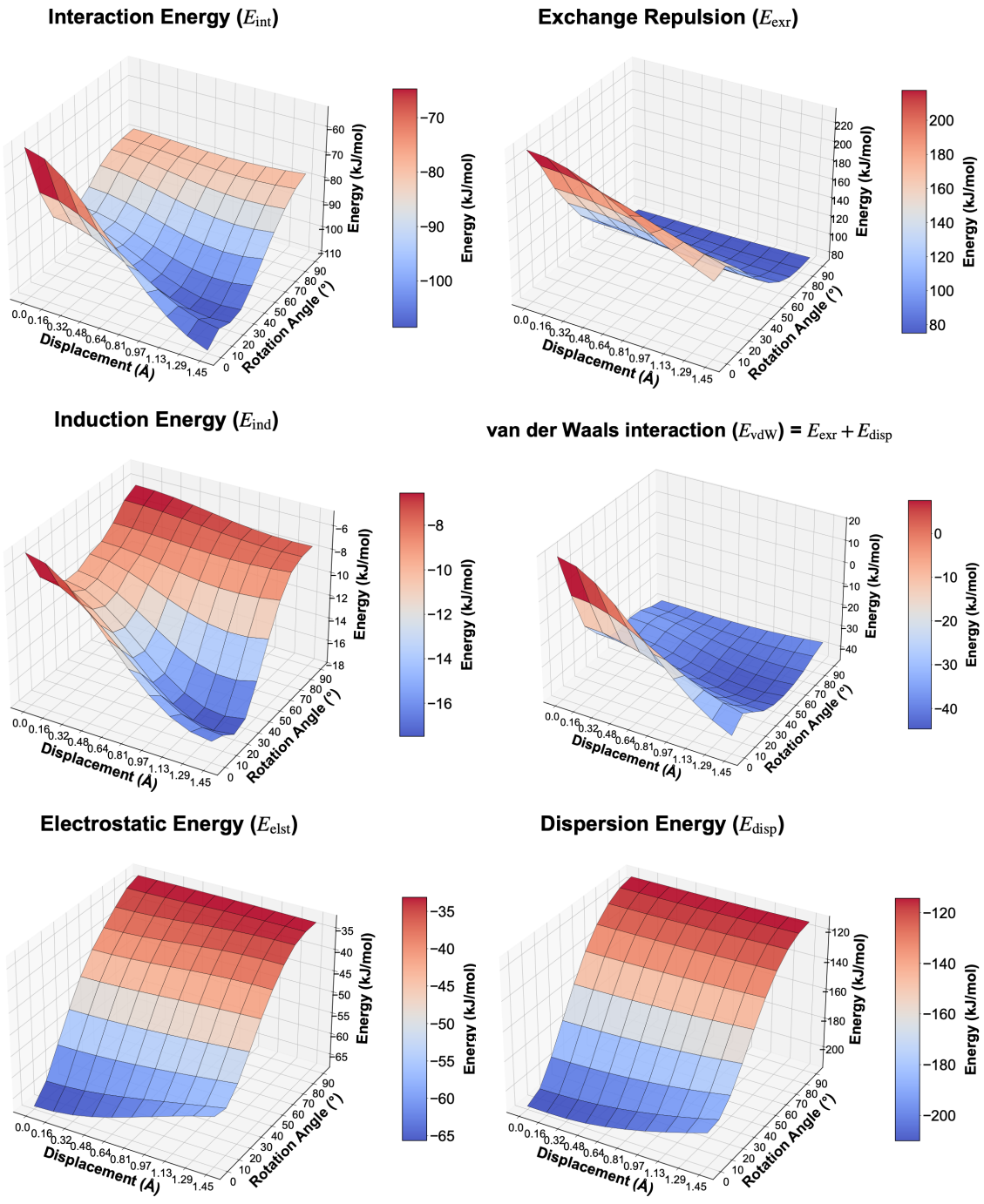


Figure 19: The screening of the interaction energy surface for pentacene dimers with sSAPT0/jun-cc-pVDZ method. The different contributions to the interaction energy (E_{int}) are represented: the exchange-repulsion (E_{exr}), the induction (E_{ind}), the electrostatic (E_{elst}), the dispersion (E_{disp}) contributions. Additionally, $E_{vdW} = E_{exr} + E_{disp}$ was added to this screening of the interaction energy surface.

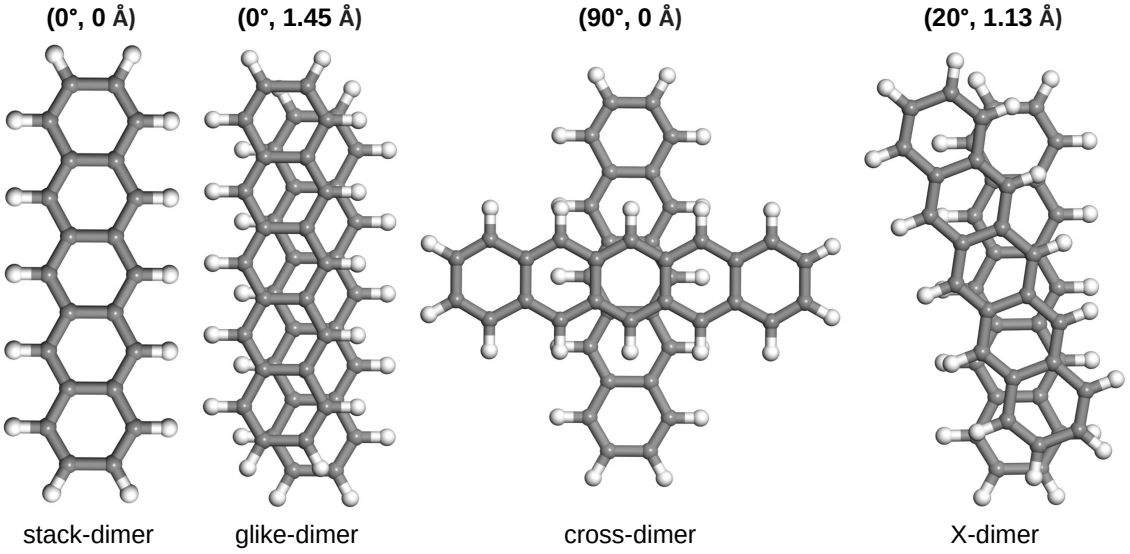


Figure 20: Top view on four selected dimer structures on the interaction energy surface for pentacene system. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

The sum of dispersion and exchange-repulsion contributions, the van der Waals interaction energy ($E_{vdW} = E_{disp} + E_{exr}$) is additionally provided[218, 338]. E_{vdW} is attractive across the entire surface, with the exception of the stack-dimer region. The most attractive energy region of the surface corresponds to geometries with a specific rotational angle (from 10°) and/or large translational displacement (more than 0.16 Å), as in the glike-dimer configuration. However, without the inclusion of E_{elst}, the E_{vdW} contribution alone does not correctly reproduce the behavior of the interaction energy surface. Therefore, E_{elst} also plays a role in determining the minima of the interaction energy surface.

While the X-dimer and the glike-dimer were identified as the most stable dimers on this interaction energy surface, this fact does not imply that other dimer configurations are unstable. For example, the cross-dimer corresponds to an boundary point on this interaction energy surface has been reported as a stable local minimum in the literature[32].

In summary, the X-dimer represents an additional minimum of the interaction energy surface, comparable in stability to the glike-dimer, with a similar E_{int} magnitude. In the following section the relative stability of these dimers using fully DFT optimized geometries, along with a SAPT analysis of the energy contributions is shown.

4.2.2 The SAPT Analysis of the glike-dimer and the X-dimer Trajectory

Firstly, the X-dimer and glike-dimer geometries were optimized for pentacene system with the B3LYP-D3(BJ) method. To further characterize the nature of the intermolecular interactions in pentacene dimers and insights into relative stability, the NCI-plot method[339] was applied, which provides a visual representation of weak NCIs, such as hydrogen bonds and van der Waals interactions, by mapping their spatial distribution and distinguishing their attractive or repulsive character. This approach is widely used to analyze and compare intermolecular interactions in molecular systems. The reduced density gradient function $s(\mathbf{r})$ lies at the core of NCI-plot analysis:

$$s(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|}{2\,(3\pi^2)^{1/3}\,\rho(\mathbf{r})^{4/3}}, \tag{65}$$

where $\nabla\rho(\mathbf{r})$ and $\rho(\mathbf{r})$ are electron density gradient and electron density at point $\mathbf{r}$.

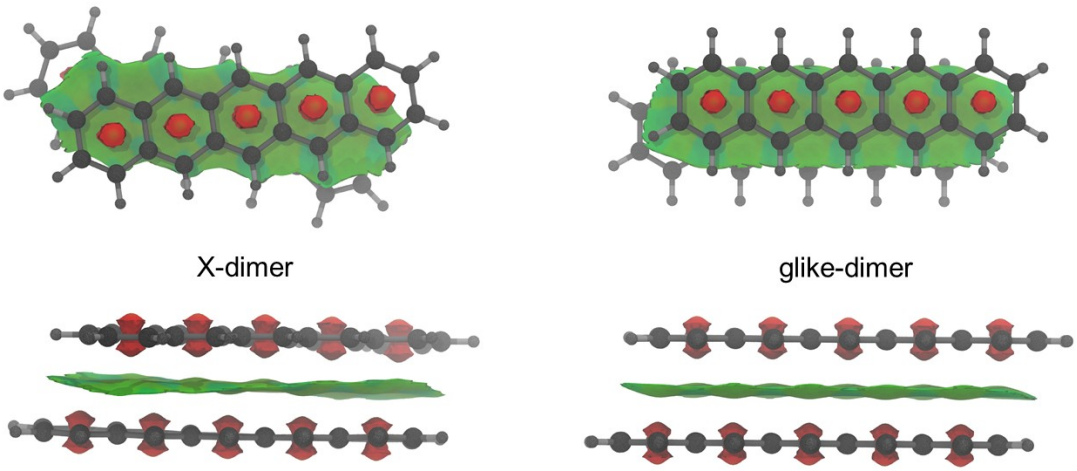


Figure 21: NCI-plot for the X-dimer and the glike-dimer of pentacene. Red zones correspond to strong repulsion, blue zones to strong attraction, and green zones to weak van der Waals interactions. Reprinted with permission from [77]. Copyright 2026 AIP Publishing.

The results for the pentacene dimers are shown in Figure 21. Unfortunately, in the case of the pentacene dimers, no qualitative difference in the non-covalent interaction regions between the X-dimer and the glike-dimer could be observed.

To investigate the relationship between the X-dimer and the glike-dimer, nudged elastic band (NEB) calculations with DFT were performed to estimate the energy barrier of interconversion between the two dimers. This method generates a series of intermediate geometries (images) along the transformation pathway. The NEB trajectory comprised 10 intermediate structures or images. The potential energy surface along this pathway was found to be relatively flat, making it difficult to unambiguously identify an exact transition state. Nevertheless, the DFT-based NEB calculations yielded an energy barrier of approximately 4 kJ/mol.

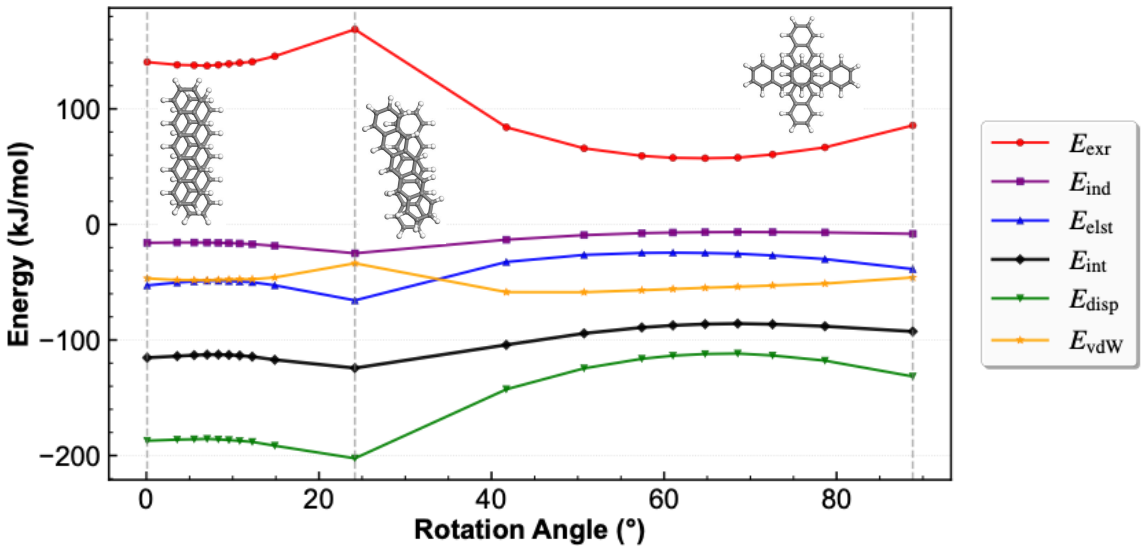


Figure 22: The sSAPT0/jun-cc-pVDZ analysis of the interaction energy contributions in pentacene dimers as a function of the rotation angle between the monomers long axes. The SAPT energy contributions: exchange-repulsion (E_{exr}), induction (E_{ind}), electrostatic (E_{elst}), total interaction energy (E_{int}), dispersion (E_{disp}), and van der Waals (E_{vdW}). Adopted with permission from [77]. Copyright 2026 AIP Publishing.

Based on the geometry of each image, a SAPT analysis was performed for the transformation from the glike-dimer to the X-dimer. The sSAPT0 decomposition of E_{int} allows identification of which energy components contribute most significantly to the changes in E_{int} along the transformation pathway. The cross-dimer was also included as an additional reference point, representing the fully rotated configuration. The transformation between the X-dimer and the glike-dimer predominantly involves rotation, with the glike-dimer representing a near-parallel arrangement of the acene monomers. therefore the sSAPT0 analysis was performed for the full transformation

from the glike-dimer to the X-dimer, and subsequently from the X-dimer to the cross-dimer. The resulting curves are shown in Figure 22.

In Figure 22 the different E_{int} contributions are presented: the exchange-repulsion (E_{exr}), the induction (E_{ind}), the electrostatic (E_{elst}), the dispersion (E_{disp}), and the total interaction energy (E_{int}) that mapped as a function of the rotation angle between the two monomers. The glike-dimer corresponds to a rotation angle of 0.1°, while the X-dimer is located at approximately 24.2°.

Moving along the curves from the glike-dimer to the X-dimer, E_{exr} increases while the other contributions decrease, particularly E_{elst} and E_{disp}. In the region between the glike-dimer and the X-dimer, a minimum in E_{exr} and maxima in the other contributions are observed. The stability of the X-dimer can be attributed to its orientation, which provides additional contributions to E_{int} from E_{disp} and E_{elst}. These contributions not only compensate for the increase in E_{exr}, but overcome it, resulting in a total decrease of E_{int}.

Upon further rotation toward the cross-dimer, E_{exch} reaches its minimum among all three dimer configurations, while the other contributions are slightly more repulsive relatively to the X-dimer and glike-dimer orientations, rendering the cross-dimer less stable than the other two. $E_{vdW} = E_{disp} + E_{exr}$[338] is attractive across the entire rotational range and follows the same trend as E_{exr}, with maxima at the glike-dimer, the X-dimer, and the cross-dimer orientations.

The analysis shows that at each dimer all energy components exhibit a local extremum, most notably E_{exr}. However, the dispersion and electrostatic contributions compensate for the increase in E_{exr}, particularly in the X-dimer, rendering it more stable than the glike-dimer.

Having established the nature of the stabilizing interactions in the pentacene dimers, it is possible to study the trends in dimer stability across the acene series.

4.2.3 DFT and SAPT Interaction Energies for Acene Molecules

In this subsection, acene molecules from naphtalene to pentacene and their corresponding glike-dimer and X-dimer configurations were investigated using DFT and SAPT (see Figure 23). As a first step, all dimer and monomer geometries were optimized at the B3LYP-D3(BJ) level of theory, which was also used to compute the CP-corrected interaction energies (E_{int}). Furthermore, sSAPT0 calculations were performed on all optimized structures to obtain the SAPT interaction energies. Calculations for

acene dimers from naphtalene to tetracene were perfomed by Andrea Buchwald and Professor Reinhold F. Fink (University of Tübingen). The interaction energies for both methods are compiled in Table 6.

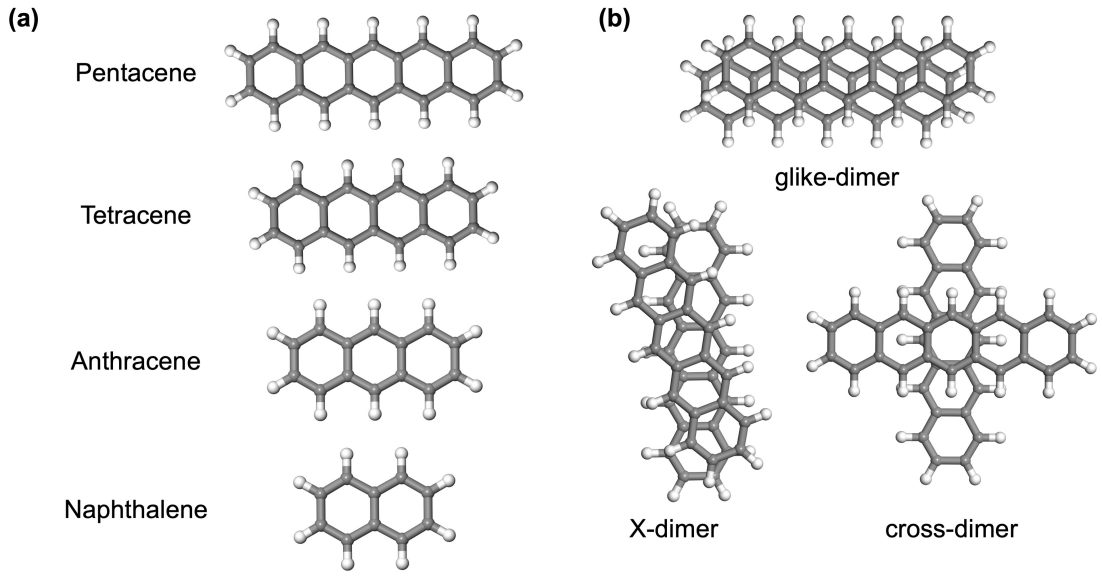


Figure 23: (a) The view of studied acene molecules: naphthalene, anthracene, tetracene and pentacene; (b) The geometries of considered dimeric conformers for the pentacene system.

Table 6: The interaction energy (E_{int}) and difference (ΔE_{int}) for acene dimers calculated with B3LYP-D3(BJ), sSAPT0 and PNO-LCCSD(T*)-F12 methods. All values are given in kJ/mol. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

Method	**B3LYP-D3(BJ)**			**sSAPT0**			**PNO-LCCSD(T*)-F12**		
System	glike	X	ΔE_{int}	glike	X	ΔE_{int}	glike	X	ΔE_{int}
Naphthalene	-29.59	-30.51	-0.92	-34.57	-36.08	-1.51	25.97	-26.63	-0.66
Anthracene	-47.26	-49.96	-2.70	-60.05	-62.68	-2.63	-41.79	-43.30	-1.51
Tetracene	-65.14	-70.48	-5.34	-86.99	-93.42	-6.43	-58.02	-60.92	-2.90
Pentacene	-83.15	-91.42	-8.27	-115.35	-124.29	-8.94	-74.42	-78.01	-3.59

Due to the increase in the size of the conjugated π-system, the stacking interaction energy increases across both methods. For the naphthalene dimers, the DFT interaction energies for the glike- and X-dimers are -29.59 and -30.51 kJ/mol, respectively. For pentacene dimers values are below -83 kJ/mol. The same trend is observed for the sSAPT0 method. However, in all cases it overstimates the interaction energy relative

to DFT. For example, pentacene dimers E_{int} are -115.35 and -124.29 kJ/mol for the glike- and X-dimers, respectively.

Insight into this trend is provided by the buried surface area (BSA) analysis (see Figure 24)[340]. This quantity describes the molecular surface area buried upon complex formation, or, in the present case, upon non-covalent dimer association. BSA was calculated using the following formula:

$$BSA = ASA_A + ASA_B - ASA_{AB} \tag{66}$$

where ASA_A is the solvent-accessible surface area of molecule A in isolation, ASA_B is the solvent-accessible surface area of molecule B in isolation, and ASA_{AB} is the solvent-accessible surface area of the complex formed by A and B. ASA is usually computed with a solvent probe, typically water with radius about 1.4 Å, rolling over the van der Waals surface of the molecule. Both ASA were calculated in Orca 5.0.4 with GEOPOL algorithm[341–343]. In the case of acene dimers $\mathrm{ASA}_A = \mathrm{ASA}_B$ and is called the monomer surface area (S_{mono}). Consequently, BSA provides a quantitative measure of the contact surface area between the two monomers in the dimer.

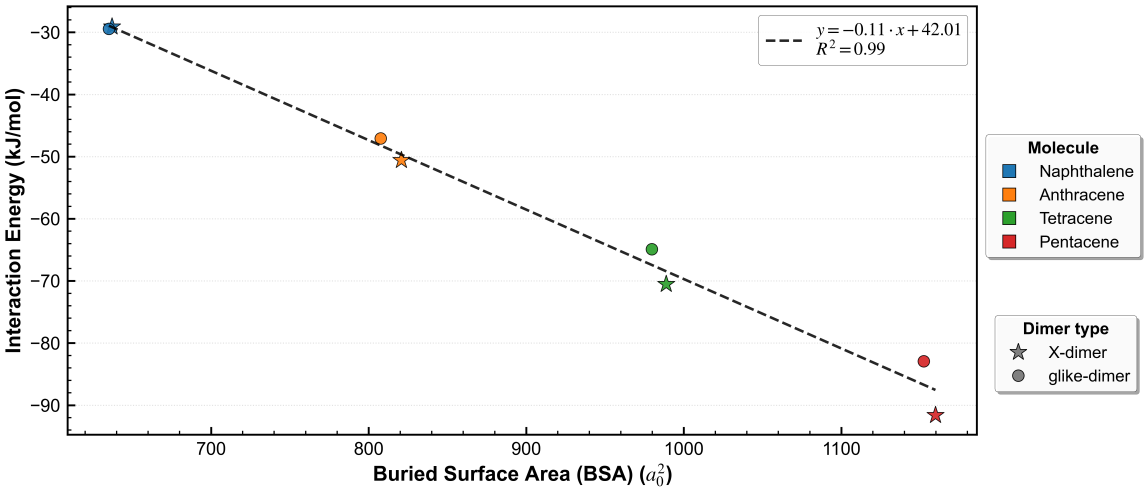


Figure 24: The interaction energy (E_{int} in kJ/mol) as a function of the buried surface area (BSA in a_0^2). Each color corresponds to a different acene molecule and the marker shape denotes the dimer type (the glike-dimer or the X-dimer).

The BSA increases with acene size, from approximately 640 Å^2 for naphthalene dimers to approximately 1150 Å^2 for pentacene dimers (Figure 24). These results demonstrate that the buried surface area is strongly correlated with the interaction energy in these dimers across the acene series ($R^2 = 0.99$). It is well established that London dispersion forces scale with the number of interacting atom pairs across the contact

interface, rendering them inherently proportional to the contact surface area. Larger PAHs with greater molecular surface area present more C–C and C–H pairs in van der Waals contact, thereby amplifying the cumulative dispersion interaction[344, 345]. Electrostatic interactions also increase with the size of the π-system, though their magnitude is more sensitive to the relative orientation of the monomers[218].

ΔE_{int}, defined as the difference between the X-dimer and the glike-dimer interaction energies, indicates that the X-dimer is more stable compared to the glike-dimer for all studied acenes. Furthermore, ΔE_{int} becomes increasingly negative with an increasing number of fused rings (N_{rings}). ΔE_{int} reaches -8.27 and -8.94 kJ/mol for pentacene dimers, while for naphthalene the difference amounts to only approximately -1.0 and -1.5 kJ/mol for the DFT and SAPT methods, respectively.

However, no systematic trend is observed in the BSA difference ($\Delta BSA_{\mathrm{X-glike}}$) between the glike- and X-dimer values. For naphthalene, BSA difference is 1.99 a_0^2, while for anthracene, tetracene, and pentacene the values are 13.08, 9.15, and 7.42 a_0^2, respectively.

Overall, the transformation from the glike-dimer to the X-dimer resulted in a more stable configuration, with the interaction energy difference ranging from approximately 1–2 kJ/mol for naphthalene to 8–9 kJ/mol for pentacene, consistently for SAPT and DFT methods.

Under equilibrium conditions, assuming that the vibrational and entropic contributions are approximately equal for both dimers, the difference in their relative stability can be attributed to ΔE_{int}, as determined from the electronic energies of the optimized structures. Furthermore, the difference is less than 10 kJ/mol that could easily fall within the method's error. Nevertheless, at this moment it is possible to assume that such a difference in values is indeed genuine. It is also permissible to assume the absence of state degeneracy. Under these assumptions, the relative population of each dimer can be estimated using the Boltzmann distribution:

$$P_i = \frac{\exp\left(-\frac{E_i}{k_B T}\right)}{Z}, \tag{67}$$

where the partition function is defined as:

$$Z = \sum_j \exp\left(-\frac{E_j}{k_B T}\right). \tag{68}$$

The ratio of populations of the two dimer conformations can therefore be written as:

$$\frac{P_{\text{glike}}}{P_X} = \exp\left(-\frac{\Delta E_{\text{int}}}{RT}\right), \tag{69}$$

where $T = 298$ K, $R = 8.314$ J mol^{-1} K^{-1}, and $\Delta E_{\text{int}} = 9$ kJ/mol.

Substituting the numerical values yields:

$$\frac{P_{\text{glike}}}{P_X} = \exp\left(-\frac{9000}{8.314 \times 298}\right) = \exp(-3.63) \approx 2.6 \times 10^{-2}. \tag{70}$$

Consequently, at room temperature, the glike-dimer contributes only approximately 2-3% to the equilibrium population in the gas-phase, while the X-dimer accounts for approximately 97-98%, indicating a strong thermodynamic preference for the more stable configuration.

Both methods show consistent trends in ΔE_{int} and E_{int}. However, numerous benchmark studies have highlighted known limitations in the accuracy of DFT and SAPT methods for non-covalent interactions[21, 76, 346, 347]. To reduce these uncertainties, additional single-point calculations were performed at the PNO-LCCSD(T*)-F12 level of theory.

4.2.4 Theoretical Insights from PNO-LCCSD(T*)-F12

In this subsection the results obtained in collaboration with Professor Hans-Joachim Werner (the University of Stuttgart) are presented. First, the quality of the optimized dimer geometries was validated by examining the intermolecular distances as a structural descriptor. The trends in the interaction energy and the relative stability of the X-dimer with respect to the glike-dimer were then calculated. In the end, B3LYP-D3(BJ) and sSAPT0 results were compared with the PNO-LCCSD(T*)-F12 reference.

Geometry Optimization

To validate the geometry of the glike-dimer optimized at the B3LYP-D3(BJ) level, PNO-LCCSD(T*)-F12 single-point energy calculations were performed by Professor Hans-Joachim Werner for a series of structures with systematically varied intermolecular distances (the step is 0.05 Å). A potential energy curve was constructed from these calculations and is shown in Figure 25. Based on polynomial fit of data points, the PNO-LCCSD(T*)-F12 energy minimum occurs at an intermolecular distance only

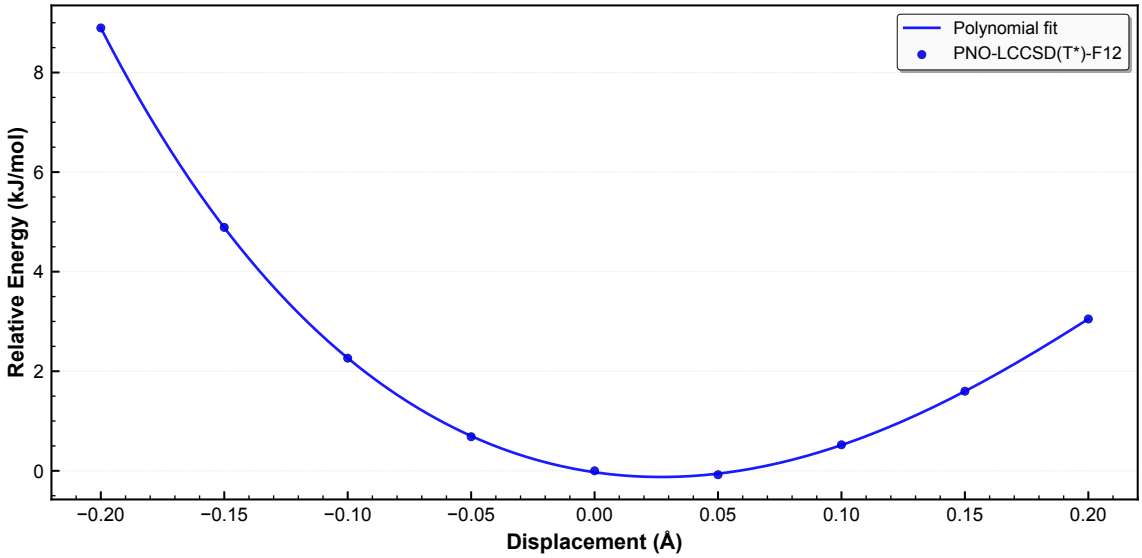


Figure 25: Potential energy curve (in kJ/mol) as a function of the displacement of the monomer intermolecular position from its value in the DFT-optimized glike-dimer geometry (in Å). The energy of the DFT-optimized glike-dimer geometry is taken as the reference point (0 kJ/mol at zero displacement). Adopted with permission from [77]. Copyright 2026 AIP Publishing.

0.027 Å larger than that obtained from the DFT geometry optimization (corresponding to the unshifted glike-dimer geometry in Figure 25), with an energy difference of only approximately 0.12 kJ/mol between the two minima. These results indicate that the B3LYP-D3(BJ) geometry closely approximates the PNO-LCCSD(T*)-F12 minimum.

This result is in line with a series of studies previously reported in the literature. Herman *et al.* demonstrated its reliability for the benzene dimer[348], while Janowski and Pulay showed that it accurately reproduces CCSD(T) geometries for both $\sigma - \sigma$ and $\pi - \pi$ dimers of aromatic and non-aromatic hydrocarbons[273]. Furthermore, in a systematic benchmark against the S66 and S66x8 datasets, Goerigk *et al.* found that B3LYP-D3(BJ) performs well for dispersion-dominated complexes and intermolecular distances[347].

Interaction Energy

The DFT and SAPT interaction energies can thus be compared with the PNO-LCCSD(T*)-F12 values to assess the magnitude of the deviations and whether the trends are preserved.

Table 6 reports the interaction energies and their differences for the PNO-LCCSD(T*)-F12 method. The PNO-LCCSD(T*)-F12 results show the same trends in E_{int} and ΔE_{int} as observed for B3LYP-D3(BJ) and sSAPT0: increasing acene size leads to more

negative interaction energies, and the X-dimer is more stable than the glike-dimer in all cases. However, the increase in magnitude of E_{int} and ΔE_{int} with acene size is smaller than predicted by DFT and SAPT. Whereas for DFT the E_{int} values for the naphthalene and pentacene X-dimer structures are approximately -30 kJ/mol and 91 kJ/mol, respectively, the corresponding PNO-LCCSD(T*)-F12 values are -27 and -78 kJ/mol. Similarly, the increase in ΔE_{int} from naphthalene to pentacene is more pronounced with DFT (-0.92 to -8.27 kJ/mol) than with PNO-LCCSD(T*)-F12 (-0.66 to -3.59 kJ/mol).

Consequently, comparing the equilibrium populations of the pentacene dimers at room temperature (see Eq. 70), and with $\Delta E_{int} = 4$ kJ/mol, yields the following expression:

$$\frac{P_{\text{glike}}}{P_X} = \exp\left(-\frac{4000}{8.314 \times 298}\right) = \exp(-1.614) \approx 0.199. \tag{71}$$

At the PNO-LCCSD(T*)-F12 level of theory, the glike-dimer accounts for approximately 20% of the equilibrium population, with the remaining 80% attributed to the X-dimer. This differs notably from the DFT result, where the X-dimer population exceeded 97%.

In summary, DFT-optimized geometry are in good agreement with PNO-LCCSD(T*)-F12 single-point interpolated minimum, particularly with respect to the intermolecular distances in the dimer structures. All three methods show consistent trends across two interaction energy descriptors: the total interaction energy and the interaction energy difference, indicating that DFT and SAPT reliably describe the energetics of non-covalent interactions in acene dimers.

4.2.5 Comparison of Different Methods to PNO-LCCSD(T*)-F12

A key question in any theoretical study is how accurately different methods reproduce the reference trends and values. To address this, single-point interaction energy calculations were performed on the B3LYP-D3(BJ)-optimized geometries with different DFT-based, semi-empirical and NNP-based methods (see Section 4.1.2). The mean absolute deviations and the mean signed deviations from the PNO-LCCSD(T*)-F12 reference interaction energies were evaluated. In addition, the ability of each method to reproduce the correct trend in the relative stability of the two dimer types was checked.

A comparison of the absolute interaction energies across all tested methods shows that they can be broadly classified into three categories based on their MAD values (see Figure 26). The first category exhibits the poorest performance, with MADs exceeding 10 kJ/mol. This category includes B3LYP-D4 (10.2 kJ/mol), M06-D3 (10.7 kJ/mol), and sSAPT0 (25.6 kJ/mol), all of which systematically overestimate the interaction energy relative to the PNO-LCCSD(T*)-F12 reference values (MSDs are negative).

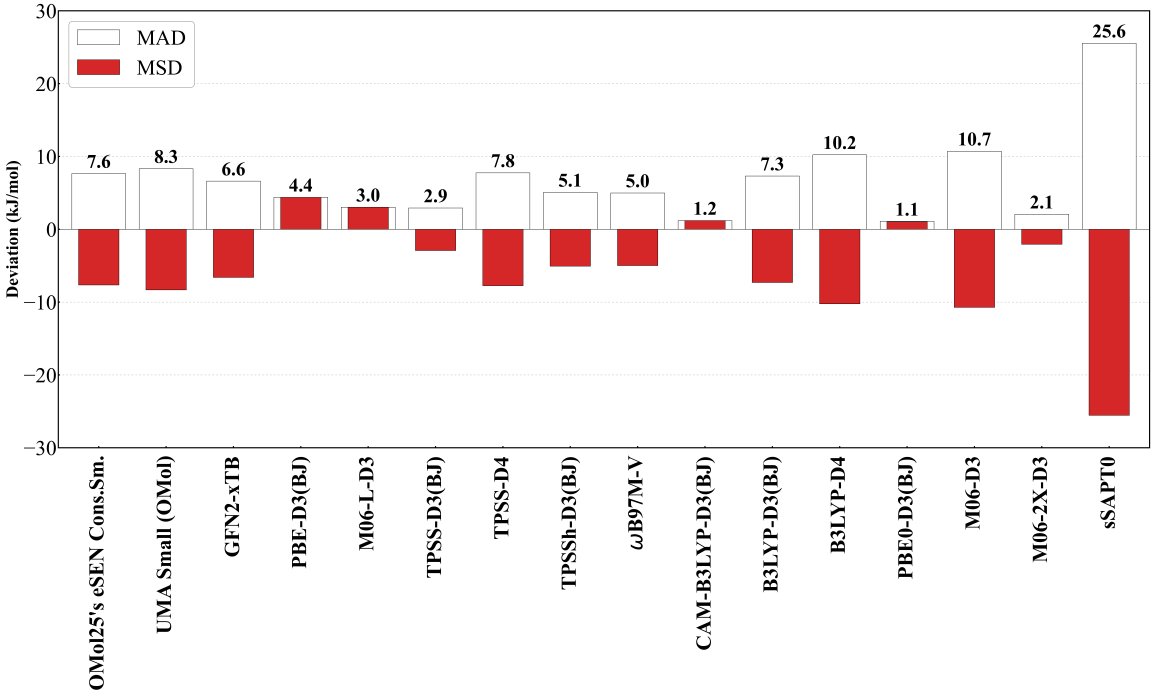


Figure 26: Mean absolute deviations (MAD, white boxes) and mean signed deviations (MSD, red boxes) of the interaction energy (in kJ/mol) for each tested method with respect to the PNO-LCCSD(T*)-F12 reference values, averaged over all studied complexes. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

The second category demonstrates intermediate accuracy, with MAD values in the range of 5–8 kJ/mol, comprising NNP's UMA Small (8.3 kJ/mol), TPSS-D4 (7.8 kJ/mol), OMol25 eSEN Cons. Sm. (7.7 kJ/mol), B3LYP-D3(BJ) (7.3 kJ/mol), GFN2-xTB (6.6 kJ/mol), TPSSh-D3(BJ) (5.1 kJ/mol), and ωB97M-V (5.0 kJ/mol), all of which tend to overestimate the interaction energy.

The third and most accurate category comprises methods with MAD values in the range of 1–5 kJ/mol, including PBE-D3(BJ) (4.4 kJ/mol), M06-L-D3 (3.0 kJ/mol), TPSS-D3(BJ) (2.9 kJ/mol), M06-2X-D3 (2.1 kJ/mol), CAM-B3LYP-D3(BJ) (1.2 kJ/mol), and PBE0-D3(BJ) (1.1 kJ/mol). Unlike the previous two categories, these

methods do not only overestimate the interaction energy; in certain cases, a slight underestimation is observed (PBE, M06-L, CAM-B3LYP, PBE0).

Interestingly, the fraction of Hartree–Fock exchange incorporated into the functional does not exhibit a consistent trend with respect to accuracy. For instance, the inclusion of HF exchange reduces the MAD when going from PBE to PBE0, whereas the opposite trend is observed in the TPSS family, where TPSSh yields a higher MAD than pure TPSS. In the M06 family, a non-monotonic behavior is observed: the MAD increases from M06-L to M06, but decreases again for M06-2X, falling below the M06-L value.

An additional question is how dispersion corrections behave in terms of deviations. Dispersion correction D3(BJ) provides a more accurate description of the interaction energy than D4 for both TPSS (2.9 and 8.7 kJ/mol) and B3LYP (7.3 and 10.2 kJ/mol). The ωB97M-V method with VV-10 dispersion correction[349] also showed relatively good results for MAD (5.0 kJ/mol).

To assess the contribution of each benzene ring in the acene to the deviation of the interaction energy from the reference method, the MAD and MSD were additionally computed on a per ring basis (e.g., naphthalene contains two fused rings, while pentacene contains five fused rings). The resulting MAD and MSD per ring are presented in Figure 52 (see Appendix for Chapter 4, Section 4.2). It is important to note that, with the exception of sSAPT0, whose MAD per ring reached 7 kJ/mol, all other methods remained within 3 kJ/mol per ring. This suggests that the large deviations observed for certain methods are driven primarily by the size of the system itself, rather than solely by the treatment of the electronic effects within the dimer. For example, the deviation for pentacene X-dimer and glike-dimer in E_{int} are 41 and 46 kJ/mol, respectively with sSAPT0.

While MAD and MSD values provide valuable insight into the accuracy of the tested methods for reproducing absolute interaction energies, a more critical question concerns their ability to correctly reproduce the relative stability of the two dimer types. Specifically, it is important to estimate whether the methods correctly identify the X-dimer as the ground state and whether they reproduce the trend of increasing stability difference with acene size. To address this, ΔE_{int} was evaluated for each acene system considered (see Table 7).

The NNP methods predict the glike-dimer to be slightly more stable than the X-dimer in some cases, with small differences and no clear trend across the acene series. In contrast, all other methods correctly identify the X-dimer as the ground state for each acene. The trend of increasing ΔE_{int} with the size of the acene, reflecting the increasing

stabilization of the X-dimer, is reproduced by sSAPT0 and all DFT functionals except semi-empirical GFN2-xTB and M06-D3. A separate analysis of the effect of dispersion correction revealed that the choice between D3(BJ) and D4 has a negligible effect on these trends for both TPSS and B3LYP, with ωB97M-V producing similar results.

Table 7: The energy difference between the X-dimer and the glike-dimer interaction energies (in kJ/mol) of all studied acenes for each tested method. First row is the PNO-LCCSD(T*)-F12 reference. Reprinted with permission from [77]. Copyright 2026 AIP Publishing.

Method	**Naphthalene**	**Anthracene**	**Tetracene**	**Pentacene**
PNO-LCCSD(T*)-F12	**-0.66**	**-1.51**	**-2.90**	**-3.59**
OMol25 eSEN Cons. Sm.	0.14	0.71	-0.60	-1.82
UMA Small (OMol)	-0.24	0.42	0.35	1.53
GFN2-xTB	-1.57	-0.91	-2.51	-2.75
PBE-D3(BJ)	-0.49	-2.46	-4.93	-8.55
M06-L-D3	-0.69	-2.28	-3.01	-5.42
TPSS-D3(BJ)	-0.88	-2.68	-5.81	-9.68
TPSS-D4	-1.11	-2.54	-6.17	-10.10
TPSSh-D3(BJ)	-0.88	-2.63	-5.46	-8.87
ωB97M-V	-0.96	-1.80	-4.28	-5.58
B3LYP-D3(BJ)	-0.92	-2.70	-5.34	-8.27
B3LYP-D4	-1.13	-2.31	-5.19	-7.90
CAM-B3LYP-D3(BJ)	-0.49	-1.98	-3.00	-4.36
PBE0-D3(BJ)	-0.49	-2.34	-4.41	-7.18
M06-D3	-0.48	-3.80	-2.89	-4.38
M06-2X-D3	-0.63	-1.29	-3.40	-5.06
sSAPT0	-1.51	-2.63	-6.43	-8.94

In conclusion, B3LYP-D3(BJ) performed within the average category in terms of mean absolute deviation; however, all trends were consistent with the reference data, as was the case for most other DFT functionals and sSAPT0. This indicates that B3LYP-D3(BJ) and sSAPT0 are suitable for an adequate description of such systems and can provide a qualitative assessment of the relative stability of the dimers.

4.2.6 Electronic Structure and Optical Properties

Having established the relative stability of the dimers, in this subsection the effect of non-covalent interaction on the electronic and optical properties of the dimers is investigated. Specifically, the frontier molecular orbitals (HOMO and LUMO), HOMO-LUMO gap, ionization potential, electron affinity, and UV-vis absorption spectra are analyzed.

Some of these features are defined as following:

$$E_{H-L} = E_{\text{LUMO}} - E_{\text{HOMO}}, \tag{72}$$

$$\text{vIP} = E^{+1} - E^{0}, \tag{73}$$

$$\text{vEA} = E^{0} - E^{-1}, \tag{74}$$

$$\text{vFG} = vIP - vEA, \tag{75}$$

where E_{H-L} is the HOMO-LUMO gap, E_{LUMO} and E_{HOMO} are the energies of the lowest unoccupied and highest occupied molecular orbitals, respectively, vIP is the vertical ionization potential, vEA is the vertical electron affinity, vFG is the vertical fundamental gap, and E^{+1}, E^{0}, E^{-1} are the total energies of the cationic, neutral, and anionic states, respectively.

Electronic Properties

First, changes in the HOMO and LUMO orbitals of the monomer and the two dimers for anthracene were examined (see Figure 27). These orbitals were calculated with the B3LYP-D3(BJ) level of theory. Upon formation of a non-covalent dimer from PAH or acene molecules, a fundamental reorganization of the frontier orbitals occurs: each molecular orbital of the monomer splits into two, one bonding and one antibonding, analogously to the interaction of atomic orbitals upon chemical bond formation. This overlap arises because the π-electron clouds in PAHs and acenes are oriented perpendicular to the molecular plane and, in a co-facial arrangement, directly face one another, as demonstrated for the benzene dimer[350].

As can be seen upon dimer formation in the case of anthracene, for the glike-dimer only the LUMO exhibits overlap, whereas for the X-dimer both the HOMO and LUMO π-orbitals overlap between the monomers. For instance, the HOMO and LUMO orbital energies of the anthracene monomer are -5.4 and -1.9 eV (Literature values are 5.7 and 1.7 eV[351]), respectively. For the glike-dimer HOMO and LUMO orbital energies are

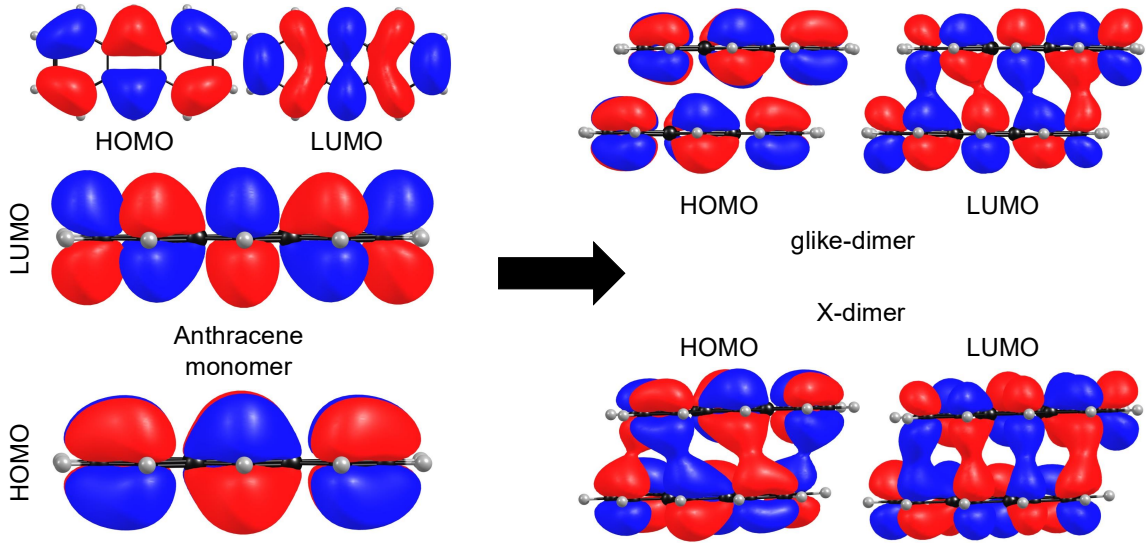


Figure 27: The view of HOMO and LUMO orbitals for (left) anthracene monomer (top and side view) and (right) two dimers (side view, top is the glike-dimer, bottom is the X-dimer) at an iso-contour value of $\sigma = 0.03\, a_0{}^{-3}$. Method: B3LYP-D3(BJ)/def2-TZVP.

-5.2 and -1.8 eV, and for the X-dimer -5.3 and -1.8 eV, respectively. The energy of both orbitals are decreased within dimers. It is well established that the HOMO-LUMO gap generally decreases with increasing number of molecules in PAH clusters, consistent with the behavior of an indirect band gap material[352]. It is therefore expected that this trend should also hold for non-covalently bound acene dimers. Indeed, this is observed for the anthracene glike-dimer (E_{H-L} = 3.35 eV). However, for the X-dimer E_{H-L} remains nearly unchanged relative to the monomer (3.52 eV).

It is well established in the literature that across the acene series from naphthalene to pentacene, the following trends are observed with increasing N_{rings}: the HOMO–LUMO gap decreases, the ionization potential decreases, the electron affinity increases, and the fundamental gap decreases. These trends arise from the progressive delocalization of the acene π-system, which lowers the barrier for oxidation and reduction, reduces the stability of the system due to a decrease in the excitation energy of low-lying excited states, and increases the polarizability of the system[351, 353, 354].

To confirm these trends and to compare the electronic properties of the monomers with those of the corresponding dimers, the following properties were calculated for monomers and dimers of all acenes considered in this work: HOMO–LUMO gap

(E_{H-L}), vertical ionization potential (vIP), vertical electron affinity (vEA), and vertical fundamental gap (vFG). These data can be found in Table 8.

Table 8: Comparison of selected electronic properties trends for the monomer, the X-dimer, and the glike-dimer across the acene series: HOMO–LUMO gap (E_{H-L}), vertical ionization potential (vIP), vertical electron affinity (vEA), and vertical fundamental gap (vFG). Every feature is in kJ/mol. Method: B3LYP-D3(BJ)/def2-TZVP.

Feature	**Type**	**System**			
		Naphthalene	**Anthracene**	**Tetracene**	**Pentacene**
E_{H-L}	monomer	456.19	339.08	261.04	206.43
	glike-dimer	438.41	324.13	253.42	200.54
	X-dimer	439.39	339.43	263.69	211.68
vIP	monomer	757.82	679.57	626.87	589.04
	glike-dimer	695.56	627.88	581.87	548.36
	X-dimer	695.94	636.39	588.43	554.04
vEA	monomer	-49.12	31.79	87.86	128.48
	glike-dimer	-21.86	54.24	106.69	143.84
	X-dimer	-21.85	51.76	102.87	139.60
vFG	monomer	806.94	647.78	539.01	460.56
	glike-dimer	717.42	573.64	475.19	404.51
	X-dimer	717.79	584.63	485.56	414.43

All observed trends for the monomer are consistent with those reported in the literature. Notably, the same trends are observed for both dimers: the HOMO–LUMO gap and the ionization potential decrease, the electron affinity increases, and the fundamental gap decreases. However, the dimers exhibit certain distinctions relative to the monomer. For instance, while the HOMO–LUMO gap of the glike-dimer is consistently lower across all acenes, the X-dimer gap gets values closest at anthracene monomer and then gets bigger. The ionization potential of both dimers is lower than that of the respective monomers across the acene series. Similarly, the electron affinity is significantly higher

for both dimers compared to the monomer. Both effects indicate a stabilization of charged states in the dimer relative to the isolated monomer. This is consistent with previous report on positively charged acene dimers, where a stabilization relative to the neutral system[25], that can be attributed to the stabilization with charge-transfer[355]. Therefore charge delocalization can decrease the energy. Furthermore, the reduction in the vertical fundamental gap from the monomer to the dimers may indicate intermolecular polarization effects. That is, the creation of a charge-hole pair in the non-covalent dimer is energetically less costly than in the isolated molecule.

Optical Properties

Previous studies have demonstrated that the acenes in both crystalline and thin-film phases undergo a pronounced bathochromic shift[228–230], which suggests that dimer formation may similarly induce a red shift in the absorption spectra. However, before an investigation of acene non-covalent complexes can be conducted, it is important to understand the quality of proposed method. Therefore, the validation of TD-CAM-B3LYP/def2-TZVP was performed on several acene monomers and compared against other methods, experimental data, and previously reported theoretical results in the literature.

Table 9: Comparison of different computational methods for excited-state calculations (first four excited states) with literature data[285] for the pentacene monomer. All values are in eV. TD-CAM-B3LYP is denoted as TD-DFT. The basis set for TD-DFT is def2-TZVP; the basis set for TD-DFT/GW+BSE, CC2, and ADC(2) is def2-TZVPPD. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

Exc. st.	TD-DFT	TD-DFT/GW+BSE	CC2	ADC(2)	CC2[285]	Exp.[285]
1	2.22	2.00	2.32	2.31	2.32	2.21
2	3.40	3.28	3.26	3.26	3.24	-
3	3.50	3.29	4.62	4.68	3.50	-
4	3.87	3.65	4.80	4.87	4.41	-

The first validation was performed for the pentacene monomer using the following methods: TD-CAM-B3LYP, TD-CAM-B3LYP/GW+BSE, CC2, and ADC(2). The first four excited states and the energies of these excitations in eV are reported in Table 9 and compared against gas-phase literature values[285] (experimental and CC2). The most relevant for UV-Vis spectra is the first excitation. CC2 and ADC(2) methods yield results consistent with the CC2 literature data (2.32, 2.31 eV), whereas the

experimental first excitation energy is 2.21 eV. TD-CAM-B3LYP/def2-TZVPPD and GW+BSE performs qualitatively worse (2.00 eV), while TD-CAM-B3LYP/def2-TZVP yields a value nearly equal to the experimental one (2.22 eV). Of course, there is extensive literature discussing the limitations of TD-DFT[356–358], various benchmarks[359–363], and methods such as TD-DFT/GW+BSE[364] or CC2/ADC(2)[365, 366] may provide higher accuracy. However, in the case of pentacene, TD-CAM-B3LYP/def2-TZVP demonstrates good accuracy.

Table 10: Comparison of calculated first excited state with experimental data for studied acene monomers, from naphthalene to pentacene. All values are in eV. TD-CAM-B3LYP is denoted as TD-DFT. Basis set is def2-TZVP.

Acene Molecule	TD-DFT	Experiment
Naphthalene	4.62	4.45[367]
Anthracene	3.47	3.43[368]
Tetracene	2.76	2.77[369]
Pentacene	2.22	2.21[285]

Then a comparison was performed for the studied acenes and benchmarked against gas-phase experimental data (see Table 10). It can be seen that the TD-CAM-B3LYP method reproduces the experimental values reasonably well, with the largest deviation observed for naphthalene: 4.62 eV versus 4.45 eV. For the remaining acenes, the deviations are well below 0.1 eV. Nearly identical values are obtained for tetracene and pentacene, with a difference of only 0.01 eV.

Table 11: Comparison of different computational methods for excited-state calculations (first four excited states) for the anthracene monomer. All values are in eV. Method: TD-CAM-B3LYP. Implicit solvent model calculations were performed using the COSMO method[370] for tetrahydrofuran (THF) and dichloromethane (DCM). The def2-TZVP basis set is denoted as TZ; The def2-TZVPPD basis set is denoted as TZPPD. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

Exc. st.	TZ	TZPPD	TZPPD/GW+BSE	TZPPD/THF	TZPPD/DCM
1	3.47	3.52	3.33	3.47	3.47
2	3.55	4.03	3.90	4.03	4.03
3	3.81	5.15	4.99	5.11	5.11
4	3.87	5.35	5.15	5.13	5.13

The third validation test of TD-DFT methods was done for the anthracene monomer. Typically, procedure for UV-Vis spectroscopy is carried out in solution[371]. Therefore, it is important to investigate the effect of solvent models on excitations. For this reason, TD-CAM-B3LYP with def2-TZVP and def2-TZVPPD basis sets was compared, as well as TD-CAM-B3LYP/GW+BSE and TD-CAM-B3LYP/def2-TZVPPD with the COSMO implicit solvent model[370] for tetrahydrofuran (THF) and dichloromethane (DCM). The results are presented in Table 11.

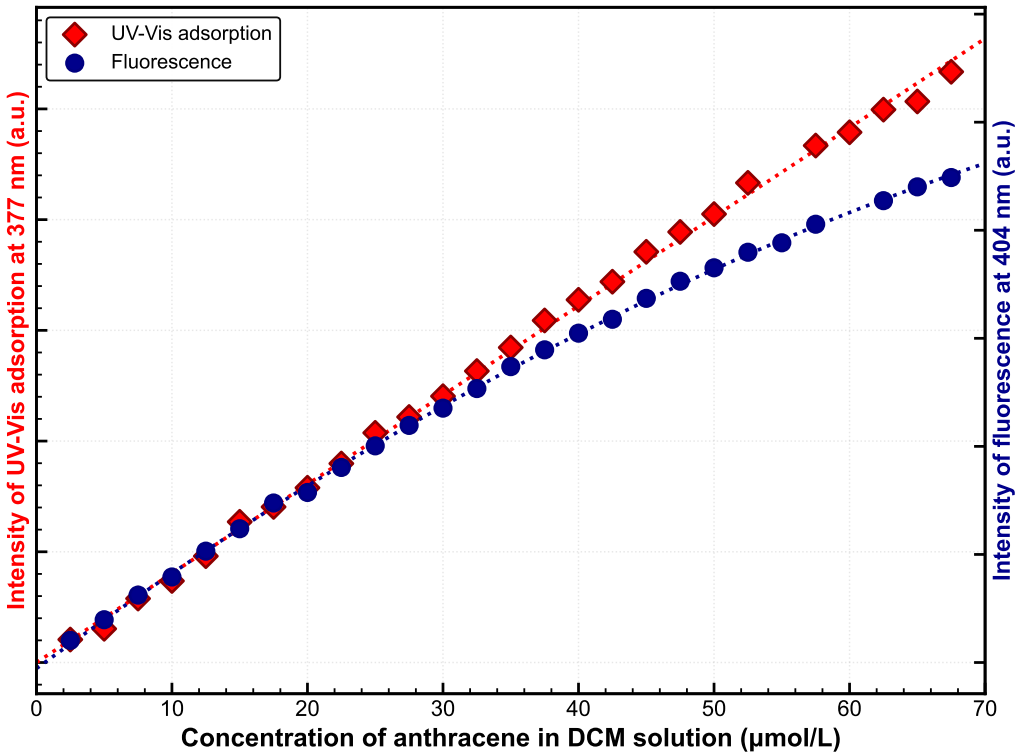


Figure 28: Intensity of UV-Vis absorption (red diamonds) and of fluorescence (blue circles) as functions of anthracene concentration in dichloromethane (DCM). Intensities are given in arbitrary units, and scaled in a way that the tangent to the fluorescence parabola at zero concentration overlaps with the UV-Vis linear trend-line. Adopted with permission from [77]. Copyright 2026 AIP Publishing.

Neither a larger basis set nor the inclusion of GW and BSE corrections improves the accuracy, since TD-CAM-B3LYP with the smaller basis set (def2-TZVP) and with the larger basis set (def2-TZVPPD) combined with implicit solvent models yield the same first excitation energy (3.47 eV). This indicates that, for the systems under consideration, there is no need to employ solvent models or higher-level corrections. A full benchmark of TD-DFT is beyond the scope of this work, as the primary goal is

to obtain a qualitative picture of the UV-Vis spectra, specifically, how the spectral features change upon $\pi - \pi$-stacking and how the peak positions shift. This qualitative understanding is what is relevant within the scope of this thesis.

Consequently, the TD-CAM-B3LYP/def2-TZVP method can be considered validated for acene-type molecules. Therefore, this method was chosen for all subsequent calculations to provide theoretical insights into the optical properties of non-covalent complexes. To investigate the effect of stacking (or aggregation) in different dimers on optical properties, fluorescence and UV-Vis spectra were measured for anthracene. The UV-Vis spectra (280–900 nm) and the fluorescence spectra were recorded, with the fluorescence excitation wavelength set to 375 nm for anthracene. Due to the limited solubility of higher acenes (tetracene and pentacene), it was not possible to access a sufficient window of concentrations in solution. Experiments were carried out by Mikhail Kalinin from the Martin-Luther-University Halle-Wittenberg. Tetrachloromethane and dichloromethane (DCM) were used as solvents in the experiments to prevent $\pi - \pi$ stacking of anthracene with solvent molecules. First, two types of anthracene spectra were compared in the low-concentration regime. Intensities of UV-Vis and fluorescence spectra were compared at different concentrations of anthracene in DCM solution (Figure 28). The concentration was varied from 0 to 70 μmol/L in steps of 2.5 μmol/L.

Although the positions of the UV absorption peaks did not change (Figure 29) and their intensity scales linearly with concentration, the intensity of fluorescence clearly deviates from linear behavior with increasing concentration. This can be considered evidence for the formation of anthracene aggregates, including dimer structures. Fluorescence quenching is well established during the aggregation process[238, 372]. In Figure 29, theoretical UV-Vis spectra are also shown for the two dimer conformers and the monomer in the low-concentration regime. The difference between the dimer peaks is only a few nanometers, close to the monomer peak position. This confirms why, experimentally, when the concentration is increased from 27.5 to 67.5 μmol/L, no shift of the UV-Vis peak is observed. Overall, there is good agreement between the first peak position of the experimental and simulated UV-Vis spectra.

To investigate why no significant shift was observed, the natural transition orbital (NTO) analysis was performed to examine the nature of the electronic transition upon excitation. This approach allows one to establish the underlying physics of the process—namely, whether the transition corresponds to a local excitation (LE) or a charge transfer (CT), when applied to acene aggregates[373]. Figure 30 presents the

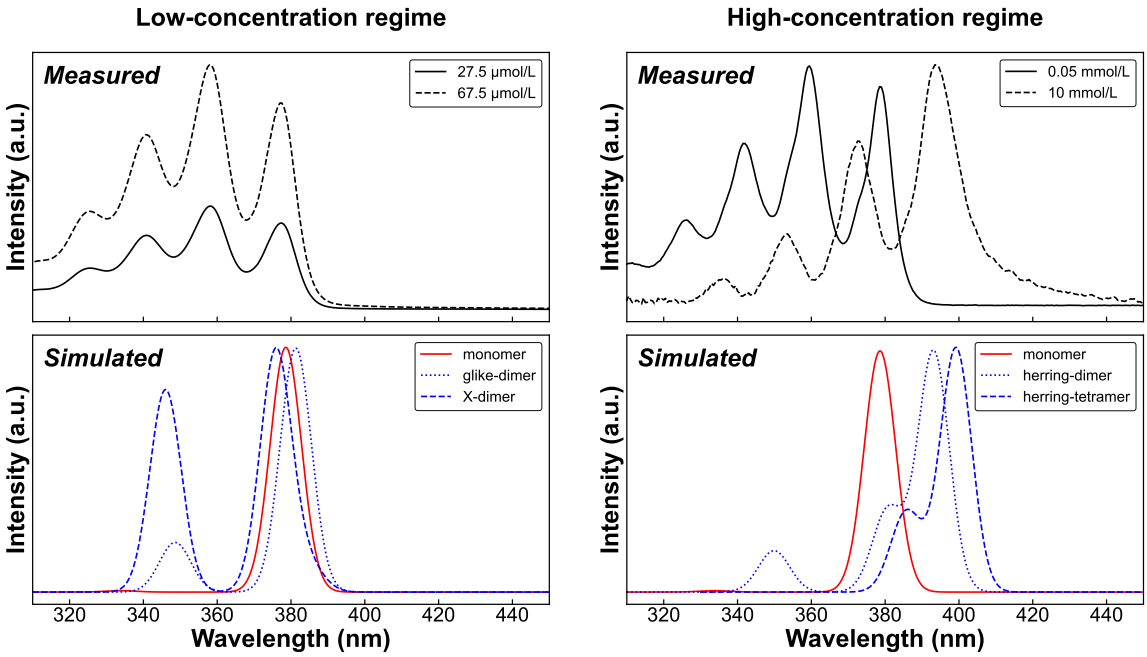


Figure 29: **Low-concentration regime:** UV-Vis spectra of anthracene measured at concentrations of 27.5 and 67.5 µmol/L in dichloromethane (upper left), as well as simulated spectra for the monomer and two dimer structures (bottom left). The simulated spectra are shifted by 25 nm to the right so that the monomer peak aligns with the experimental peak in solution.
High-concentration regime: UV-Vis spectra of anthracene measured at concentrations of 10 and 0.05 mmol/L in tetrachloromethane (upper right), as well as simulated spectra for the monomer, the herringbone tetramer, and the herringbone dimer extracted from the tetramer (bottom right). The simulated spectra (bottom right) are shifted by 26 nm to the right so that the monomer peak aligns with the experimental peak in solution.
Adopted with permission from [77]. Copyright 2026 AIP Publishing.

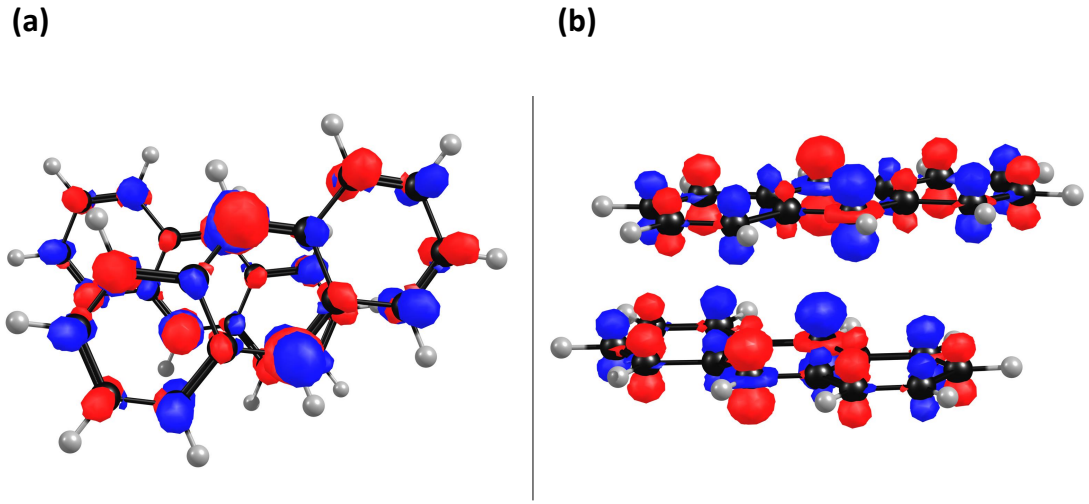


Figure 30: Natural transition orbitals analysis at an iso-contour value of $\sigma = 0.0025\, a_0{}^{-3}$ for the X-dimer of anthracene system: (a) side view, (b) top view. The red color is denoted to positive and blue to negative transition density regions. Method: TD-CAM-B3LYP/def2-TZVP.

natural transition orbitals analysis for the X-dimer of anthracene (the glike-dimer exhibits a qualitatively similar picture, that can be found in Appendix for Chapter 4, Section 4.2). One can observe that the positive (red) and negative (blue) transition density regions are localized on the same fragment of the molecule, indicating that the excitation occurs within a single molecule. In conclusion, this excitation is a LE process and the non-covalent interaction does not change the resulted LE qualitatively.

For high concentrations, tetrachloromethane was used as the solvent. Two experimental concentrations were measured: 0.05 mmol/L and 10 mmol/L. As one can see in Figure 29 for the high-concentration regime, the peak exhibits a bathochromic shift at higher concentration. Since no changes in the UV-Vis peaks were observed for the $\pi - \pi$-stacked dimers, additional structures of the herringbone type were considered for anthracene. As discussed previously, the herringbone motif is the most common crystal structure packing for acene molecules in the solid state[260, 261]. Therefore, a herringbone tetramer and a herringbone dimer were selected for calculations (see Figure 31). First, the anthracene tetramer was optimized using the same method as for the X-dimer and glike-dimer. Subsequently, the herringbone dimer was extracted from the optimized tetramer structure without further optimization, and UV-Vis spectra were calculated for both structures at the TD-CAM-B3LYP/def2-TZVP level. Indeed, the simulation of the anthracene tetramer in the herringbone arrangement showed a

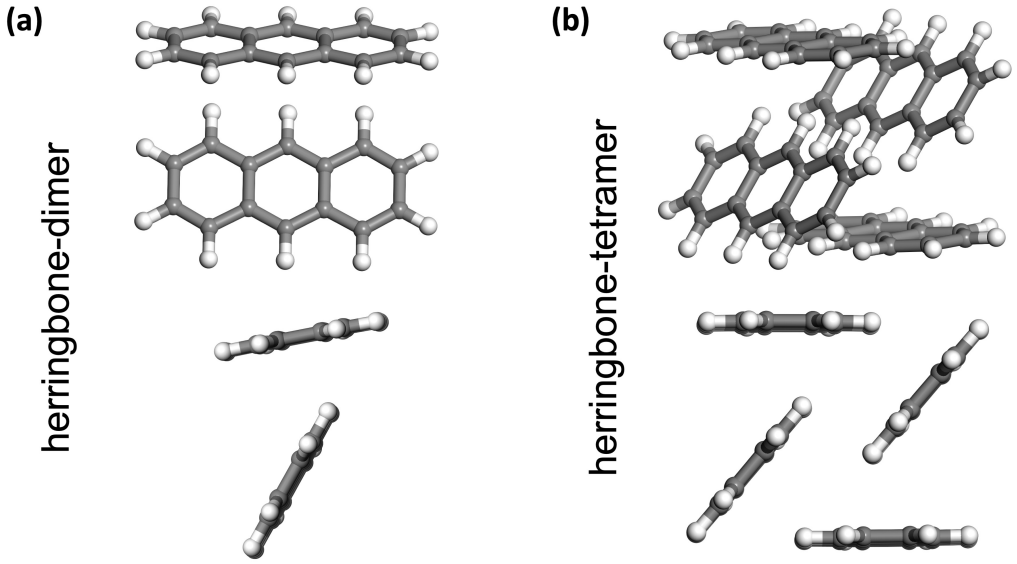


Figure 31: Examples of structures for the (a) dimer and (b) tetramer herringbone stacking of anthracene system (Top view and side view).

strong bathochromic shift, relatively to the glike- and X-dimers. By simulating a dimer extracted from such a tetramer (a structure that does not correspond to a minimum by itself), resulted in UV-Vis spectra demonstrating that it is not the cluster size but the relative orientation in space that is the origin of the observed shift (see Figure 29). Thus, larger aggregates such as the tetramer, or structural motifs derived from it, can qualitatively alter the UV-Vis peaks in high-concentration regime, whereas the glike-dimer and X-dimer cannot.

4.2.7 Discussion

In summary, two minima were identified on the interaction energy surface by screening different rotations and shifts of pentacene monomers relative to each other using the sSAPT0 method. Both minima were then separately analyzed in terms of the different contributions to their stabilization. It was shown, using pentacene as an example, that the X-dimer is more stable than the glike-dimer owing to a reduction in dispersion and electrostatic contributions that outweigh the contribution of exchange-repulsion. Subsequently, both dimers were systematically investigated for systems with varying numbers of benzene rings, from naphthalene to pentacene, using the B3LYP-D3(BJ) and sSAPT0 methods. Two trends were established: a growth in the interaction energy and an increase in the relative stabilization of the X-dimer with respect to the glike-dimer. These trends were confirmed with the reference PNO-LCSSD(T*)-F12

method and were also shown for the most of considered DFT functionals and the semi-empirical GFN2-xTB. Furthermore, some insights of the electronic properties of anthracene dimers were demonstrated, which lead to changes in the HOMO–LUMO gap through orbital modifications. It was also shown that, as in the monomer, many electronic properties follow the same trends. However, unlike the monomer, the dimers exhibited enhanced stability toward changes in charge (as illustrated by vIP and vEA). Finally, it was demonstrated that non-covalent dimers can, in principle, shift the positions of UV-Vis peaks, as exemplified by the unstable herringbone dimer; however, the stable X-dimer and glike-dimer showed no changes in the UV-Vis peak positions. Only certain changes in fluorescence were observed, which can be interpreted as possible evidence for the onset of acene aggregation. This process should logically begin with dimer formation.

4.3 Chemical Substitution Effects on Anthracene Dimers

Looking ahead, it is of interest to examine, using anthracene as a model system, how various atomic substitutions within the molecular structure affect dimer formation and their stability. And which properties reflect the strength of the intermolecular interaction in the dimer. Such phenomena will be explored through two examples involving the carbon skeleton and hydrogen atoms of anthracene: fluorine substitution (F-substitution) of hydrogen atoms and boron-nitride substitution (BN-substitution) at different carbon positions.

In this section the effects of two chemical modifications on monomeric and dimeric properties of anthracene are examined: (i) BN-substitution, replacing carbon pairs with boron–nitrogen (BN) units, and (ii) peripheral fluorination, replacing all hydrogen atoms with fluorine. For the monomers, key electronic properties: the HOMO–LUMO gap, dipole moment (μ), polarizability (α), vertical ionization potential (vIP), and vertical electron affinity (vEA). The analysis is extended to non-covalent dimers in two stacking orientations (X-dimer and glike-dimer), where the influence of substitution on the intermolecular interaction energy is compared. Finally, changes in the UV-Vis absorption spectra for selected systems are discussed. s The functionalization of hydrocarbon molecules (through peripheral substitution or internal heteroatom doping) constitutes a versatile strategy for modulating non-covalent interactions and the associated electronic properties of π-conjugated systems[374–377]. For example, a clear correlation between the vertical ionization energy and the change in molecular

dipole moment has been established for substituted benzene dimers, demonstrating that even simple substituent variations produce quantifiable shifts in the electrostatic landscape governing intermolecular binding[378]. Importantly, the classical picture in which substituents influence π-stacking by globally polarizing the ring π-density has been challenged by Wheeler and Houk, who showed that substituent effects are better rationalized as direct, local through-space interactions between the substituent and second monomer[379].

Heteroatom doping introduces additional complexity. Nitrogen substitution into PAHs, red-shifts the UV absorption bands, reflecting a systematic perturbation of the frontier orbital energetics[380]. At the intermolecular level, Do *et al.* demonstrated that the influence of a heteroatom on stacking energetics is governed by the coupling of its local bond dipole with the electric field generated by the opposing aromatic system: alignment of the heteroatom dipole along the field lines stabilizes the complex, whereas anti-alignment leads to destabilization[377]. Furthermore, heteroatom incorporation can open competing interactions. Experimental and theoretical studies of benzene–pyridine, pyridine–pyridine and benzene–pyrimidine dimers revealed that nitrogen atoms create hydrogen-bond acceptor sites that compete with π-stacking, thereby reshaping the potential energy surface and altering the preferred dimer geometries[381–383].

It is also noteworthy that different classes of NCIs respond to substitution in qualitatively distinct ways. In contrast to parallel-displaced π-stacking, C–H$\cdots\pi$ interactions, exemplified by benzene-acetylene complexes, exhibit an opposite substituent trend: electron-donating groups strengthen the binding, while electron-withdrawing groups weaken it, consistent with the dominant role of electrostatics between the partially positive hydrogen atom and the π-electron cloud of the acceptor[384]. Together, these findings underscore that a unified description of substituent effects on NCIs must account for the specific interaction topology, the local electrostatic environment of the substituent, and the possible emergence of competing binding motifs upon chemical modification.

In this work, two types of chemical modifications were investigated using anthracene as a representative model system: (i) complete peripheral fluorination, which yields perfluoroanthracene and (ii) substitution with BN, in which a pair of carbon atoms is replaced by a pair of boron-nitrogen at various positions within the aromatic framework.

Fluorination

The synthesis of perfluorinated acenes is well established, and a substantial body of research has been devoted to their structural modifications and applications[385–388], including perfluoroanthracene in particular[389–391]. It is well known that the introduction of the strongly electronegative substituent effectively converts organic p-type semiconductors into n-type materials[385, 387, 392], and, unlike their parent hydrocarbons benzene and naphthalene, perfluorinated analogues exhibit positive electron affinities[393]. Beyond the electronic effects, comprehensive computational screenings of the stability and properties of fluorinated PAHs[394–397] have revealed that perfluorination can drastically alter the relative stability of PAH isomers, in some cases completely reversing the stability ordering[397]. The effect of fluorination on $\pi - \pi$ stacking has also been investigated in the context of mixed fluorinated/non-fluorinated aromatic pairs[397]. Such interactions exhibit enhanced binding affinity compared to conventional $\pi - \pi$ contacts and can be tuned for selected self-assembly processes for functional materials fabrication.

Hetero BN-Substitution

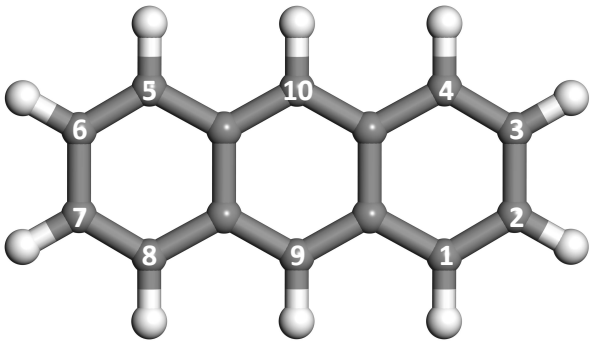


Figure 32: Atom numbering system of carbon atom positions for BN-substitution in anthracene monomer.

For convenience, a systematic labeling scheme for the BN-substitution positions within the anthracene skeleton was adopted from [398]. The scheme of positions is represented in Figure 32). Among the structures considered, the types of (9,10)-BN[399] and (2,3)-BN[400, 401] have been previously synthesized and modified. For example, the analogues of [1,4]azaborines have been prepared with various substituents and explored as electronic components[399, 402], in particular spintronic materials[403] and UV-emissive organic luminophores[404]. While the remaining BN-anthracene isomers have not yet been reported experimentally, the rapid progress in the chemical synthesis of BN-substituted PAHs is encouraging[405–407]. Furthermore, a growing body of theoretical studies and computational screenings[408–410] indicates significant

potential for BN-doped acene-based molecules in organic electronics[411] and as bright singlet-fission chromophores[398]. Notably, it has been shown theoretically for the case of pentacene that the electronic singlet-fission coupling in BN-pentacene is approximately five times stronger than in its pristine carbon analogue. Moreover, thermal distortions affect singlet fission differently in doped and undoped crystals, suggesting that BN-substitution could facilitate the multiplication of photogenerated excitons and thereby enhance the efficiency of photovoltaic devices[398].

4.3.1 Monomer Properties of Substituted Anthracenes

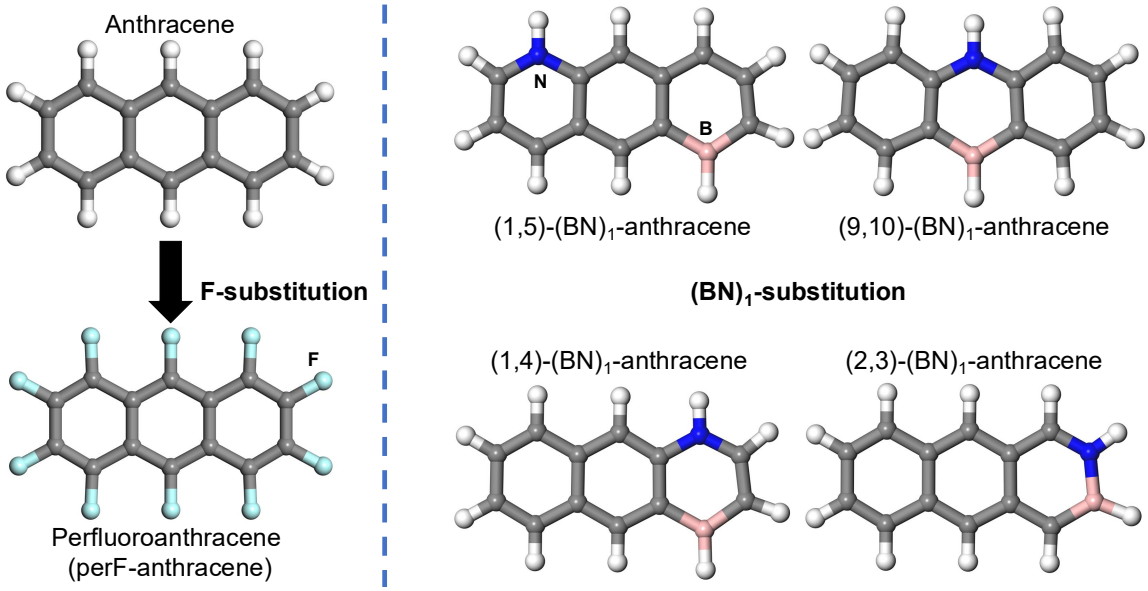
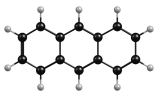
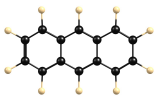
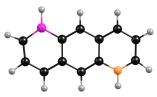
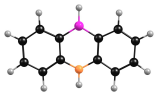
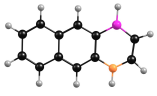
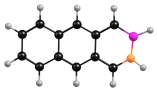


Figure 33: All considered anthracene substitutions: (left) F-substition; (right) BN-substitution: four types.

The full F-substitution and BN-substitution positions were selected to systematically probe the effect of the relative placement of boron and nitrogen atoms on the monomeric properties and, more importantly, on the non-covalent interactions in the corresponding dimers. The chosen configurations represent a progressive variation of the B and N positions within the anthracene framework: first, the two atoms are placed in separate rings at the (1,5) positions ((1,5)-BN-anthracene). Next, they occupy the central ring at the (9,10) positions, analogous to [1,4]azaborines ((9,10)-BN-anthracene). Then, both atoms are located within the same outer ring at the (1,4) positions ((1,4)-BN-anthracene). Finally, they are placed at the peripheral (2,3) positions ((2,3)-BN-anthracene).The selection of positions is determined by their placement on one or more rings of the molecule, in order to illustrate trends in the

arrangement of substituents. All resulting structures are shown in Figure 33. The perfluorinated counterpart is straightforward—all hydrogen atoms are replaced by fluorine. Following the selection of the target systems, geometry optimizations were performed using the same computational approach as for anthracene in the preceding section (see Section 4.1.2). The vertical ionization potential, vertical electron affinity, as well as vertical fundamental gap, HOMO-LUMO gap, dipole moment, and isotropic polarizability were evaluated for each monomer; the resulting values are summarized in Table 12.

Table 12: Calculated electronic properties of the anthracene monomer and its derivatives. Short description of BN-substituted monomers is (X,Y)-BN. Properties: HOMO–LUMO gap (E_{H-L}, kJ/mol), vertical ionization potential (vIP, kJ/mol), vertical electron affinity (vEA, kJ/mol), vertical fundamental gap (vFG, kJ/mol), dipole moment (μ, D), isotropic polarizability (α, a.u.).

System	Structure	E_{H-L}	vIP	vEA	vFG	μ	α
Anthracene		339.08	679.57	31.79	647.78	0.00	176.11
perF-anthracene		315.77	757.79	130.51	627.27	0.00	184.94
(1,5)-BN		183.23	591.94	109.00	482.94	8.83	203.05
(9,10)-BN		388.25	704.44	12.74	691.70	2.83	168.73
(1,4)-BN		358.70	689.26	26.99	662.27	3.32	173.09
(2,3)-BN		264.64	635.66	66.34	569.32	2.78	183.26

Anthracene and perfluoroanthracene (perF-anthracene) both possess zero dipole moment based on their symmetry, and their isotropic polarizabilities are comparable,

differing by only 9 a.u. However, the HOMO–LUMO gap is somewhat larger for anthracene than for its perfluorinated counterpart (339 vs 316 kJ/mol). The situation is different for the BN-substituted anthracene isomers. All four structures exhibit non-zero dipole moments: (9,10)-BN-anthracene, (1,4)-BN-anthracene, and (2,3)-BN-anthracene have values of approximately 3 D, whereas (1,5)-BN-anthracene stands out with a significantly higher dipole moment of roughly 9 D. The HOMO-LUMO gap of (1,5)-BN-anthracene is also substantially reduced relative to other species. Additionally, (1,5)-BN-anthracene displays the highest isotropic polarizability among all monomers (203 a.u.). Among the remaining BN-substituted isomers, (2,3)-BN-anthracene is also noteworthy, possessing the second smallest HOMO–LUMO gap at 265 kJ/mol. The isotropic polarizability scales in the following order: (1,5)-BN > perF-anthracene > (2,3)-BN > anthracene > (1,4)-BN > (9,10)-BN. Given the established correlations between dipole moment and polarizability on the one hand and the strength of NCI on the other[98, 412–414], (1,5)-BN-anthracene is expected to exhibit the most distinct behavior in the dimer configurations, and all BN-substituted dimers are anticipated to show stronger interaction energies compared to both perfluoroanthracene and anthracene.

4.3.2 Dimer Construction

For F-substituted anthracene, the geometries of the X-dimer and glike-dimer for anthracene were taken and reoptimized. Both dimers are stable; their structures can be found in Figure 34, and there is no significant difference in orientations compared to the unsubstituted case. The intermolecular distances between monomers for the anthracene X-dimer and glike-dimer are approximately 3.32 Å and 3.36 Å, while for perfluoroanthracene the values are 3.32 Å and 3.29 Å, respectively. However, for perfluoroanthracene, the angle of rotation in the X-dimer is larger (51.22° versus 34.13° for anthracene X-dimer).

BN-Substituted Anthracene

For BN-substituted monomers, the situation is more complex. To identify all possible combinations, four types of dimers were considered (Figure 35). The first type is called translation **(N)**: this second monomer is translated from the plane of the first by a distance of approximately 3.4 Å along the Z-axis without any additional operations. The rotated **(R)** dimer is obtained when the translation is combined with a rotation around the X-axis (180°). The twisted **(T)** dimer is formed when the translation is accompanied by a rotation around the Z-axis. Finally, the combination of both

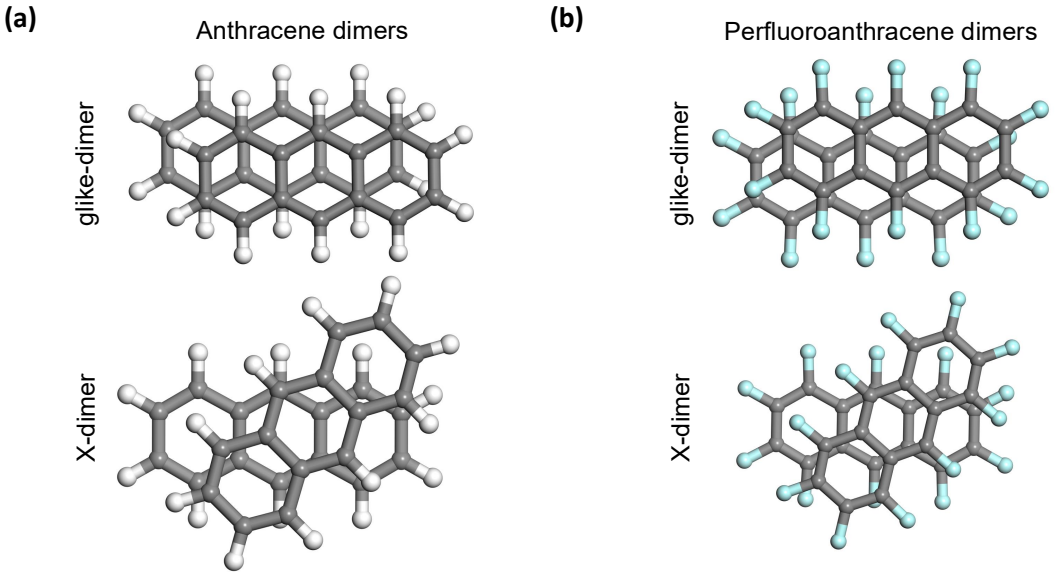


Figure 34: Anthracene and perfluroanthracene dimers: glike-dimer (C_i point group) and X-dimer (C_2 point group) structures.

rotation and twisting yields the twist-rotated **(TR)** dimer, where the monomer is first translated, then rotated around the X-axis, and subsequently around the Z-axis.

Following this, a geometry optimization of the initial dimer structures was performed. Unfortunately, not all combinations yielded successful optimizations. Some structures were unstable. Based on these results, Table 13 summarizes the stable and unstable cases. Ranked from the highest to the lowest number of stable dimers, the order is as follows: (1,4)-BN (6 dimers) > (2,3)-BN (5 dimers) > (1,5)-BN (4 dimers) > (9,10)-BN (3 dimers). The most populated category is Twisted (T) (6 dimers), while the remaining three categories each contain the same number of dimers (4 species). The structures of all dimers for each monomer are shown in Figure 36. Out of all 18 stable dimers, 12 belong to the X-dimer type and 6 to the glike-dimer type, yielding a ratio of 2:1.

4.3.3 Dimer Interaction Energies

For all aforementioned dimers of F-substituted and BN-substituted anthracene molecules, E_{int} and BSA were calculated for comparison with unsubstituted anthracene. The results of their correlation are shown in Figure 37 and all values are presented in Appendix for Chapter 4, Section 4.3. in Table 17. It can be seen that the majority of dimers of different types are located in approximately the same region, with E_{int} rang-

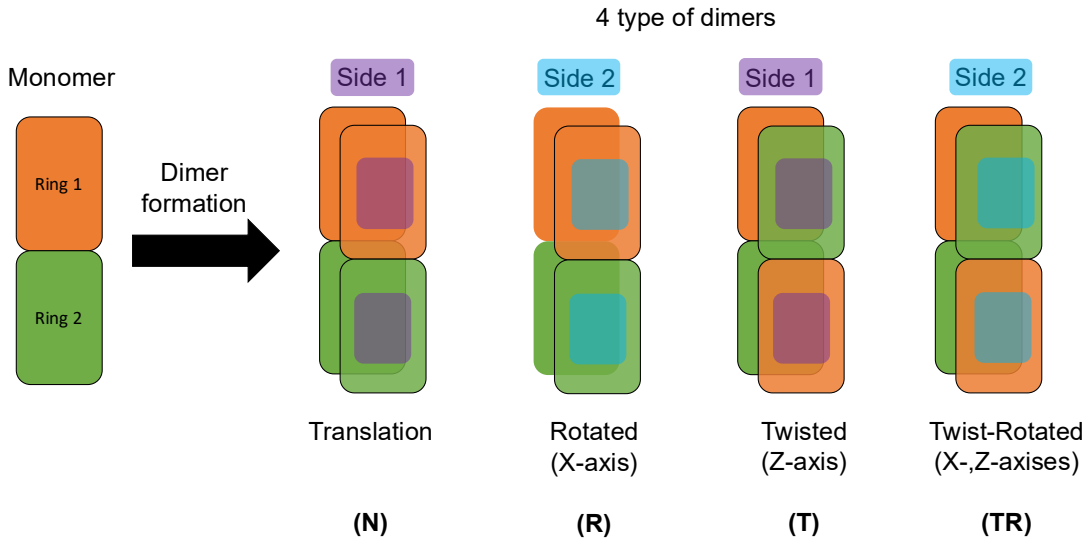
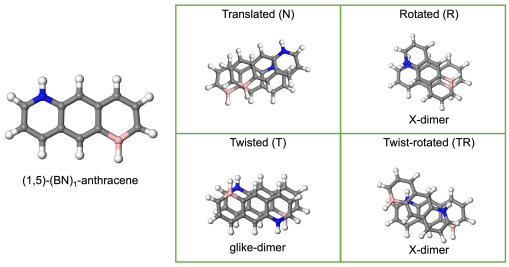


Figure 35: Schematic representation of four types of dimer for BN-substituted anthracene and their short abbreviations.

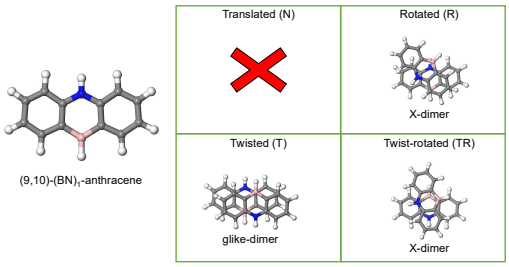
Table 13: Stability of dimer conformers for BN-substituted anthracene structures. ✓ — stable minimum; ✗ — unstable (different minimum found).

Structure	**Dimer type**	**Configuration**			
		Translation	**Rotated**	**Twisted**	**Twist-rotated**
(1,5)-BN	X-dimer	✓	✓	✗	✓
	glike-dimer	✗	✗	✓	✗
(9,10)-BN	X-dimer	✗	✓	✗	✓
	glike-dimer	✗	✗	✓	✗
(1,4)-BN	X-dimer	✓	✓	✓	✓
	glike-dimer	✓	✗	✓	✗
(2,3)-BN	X-dimer	✓	✗	✓	✓
	glike-dimer	✗	✓	✓	✗

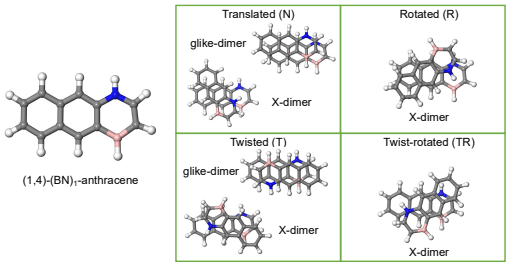
ing from -40 to -70 kJ/mol and BSA from 800 to 870 a_0^2. Both anthracene dimers and most of the BN-substituted dimers fall within this region. The majority of BN-dimers in this area exhibit lower E_{int} values, indicating greater thermodynamic stability. The perF-anthracene dimers stand out separately: their E_{int} is approximately -60 kJ/mol, but the BSA is considerably higher, around 1050 a_0^2. Two additional outliers belong to (1,5)-BN-anthracene, which possesses the highest dipole moment and polarizability among the studied monomers. Only two out of four existing dimers of (1,5)-BN-



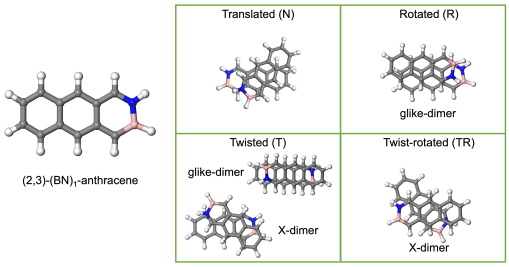
(a) (1,5)-BN dimers.



(b) (9,10)-BN dimers.



(c) (1,4)-BN dimers.



(d) (2,3)-BN dimers.

Figure 36: The view of dimers for all BN-substituted monomers: (a) for (1,5)-BN-anthracene monomer; (b) for (9,10)-BN-anthracene monomer; (c) for (1,4)-BN-anthracene monomer; (d) for (2,3)-BN-anthracene monomer. Each case has four types of dimers and stable dimers are represented inside them.

anthracene have E_{int} below -100 kJ/mol: one is the X-dimer (TR) and the other is the glike-dimer (T). Presumably, the enhanced stability of these dimers arises from the inherent instability of the monomer itself, which has the smallest HOMO-LUMO gap. As can be seen from the dimer structures (Figure 36a), only in these two dimers do the B and N centers of the interacting molecules converge geometrically. This explains the significant difference in E_{int} compared to other molecules, where the B and N centers are located within the same ring, rendering the monomers more stable.

To examine how the average E_{int} values compare across the studied substitutions, one can refer to Figure 38, which shows the beeswarm distributions of E_{int}. For unsubstituted anthracene, the mean value is approximately -50 kJ/mol. For the fluorinated derivative, the mean is lower, around -60 kJ/mol. On average, all BN-substituted derivatives exhibit lower E_{int} values compared to pure anthracene. Among them, (1,5)-BN stands out with a mean around -80 kJ/mol. The remaining three sets of BN-dimers fall within the range of -55 to -65 kJ/mol. Both (9,10)-BN and (1,4)-BN have lower average E_{int} than unsubstituted anthracene, but slightly higher than the fluorinated derivative. The (2,3)-BN derivative is lower in energy than both the fluorinated derivative and unsubstituted anthracene. Thus, ranking the species

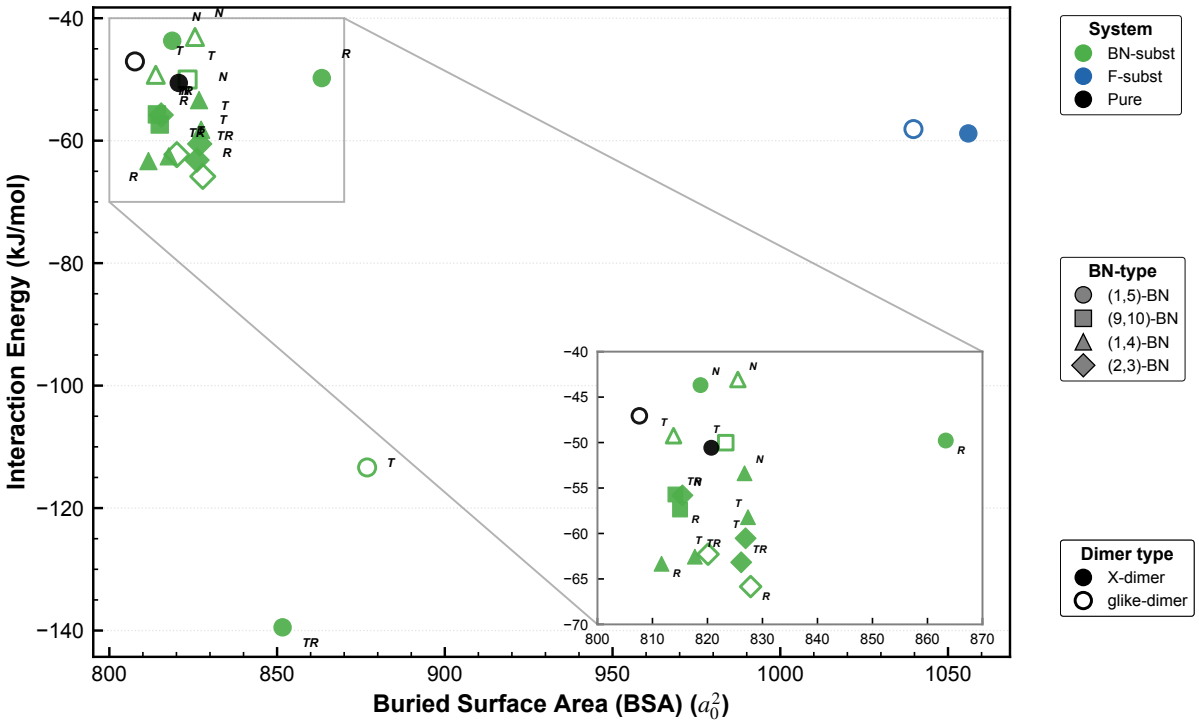


Figure 37: The interaction energy (in kJ/mol) as a function of the buried surface area (BSA in a_0^2). Each color corresponds to a different substitution of anthracene monomer and the marker shape denotes special monomer of BN-substituted anthracene. Filled and empty marker shows the type of dimer. The type of BN-substituted dimer is also considered with N/R/T/RT abbreviation.

in order of increasing average E_{int} yields the following sequence: (1,5)-BN < (2,3)-BN < F-subst < (1,4)-BN < (9,10)-BN < pure anthracene. This ordering does not fully coincide with the trends observed for isotropic polarizability or dipole moment. However, the prediction based on isotropic polarizability that F-subst and (2,3)-BN should exhibit lower E_{int} values is confirmed.

It is also necessary to examine how the stabilities of the X-dimer and the glike-dimer compare for each monomer. As shown above, both types of dimer were found to be stable for the fluorinated derivative. For the BN-substituted molecules, the situation is different: only in certain cases the same operations yielded two stable dimer types, namely for (1,4)-BN-anthracene (N), (1,4)-BN-anthracene (T), and (2,3)-BN-anthracene (T). Therefore, for anthracene, fluorinated anthracene, and selected BN-substituted monomers, it is worth examining how the relative stability varies depending on the monomer. The E_{int} and BSA values for the glike-dimer and the X-dimer pairs are presented in Table 14. As discussed above, fluorinated anthracene dimers exhibited lower E_{int} and larger BSA values than unsubstituted anthracene.

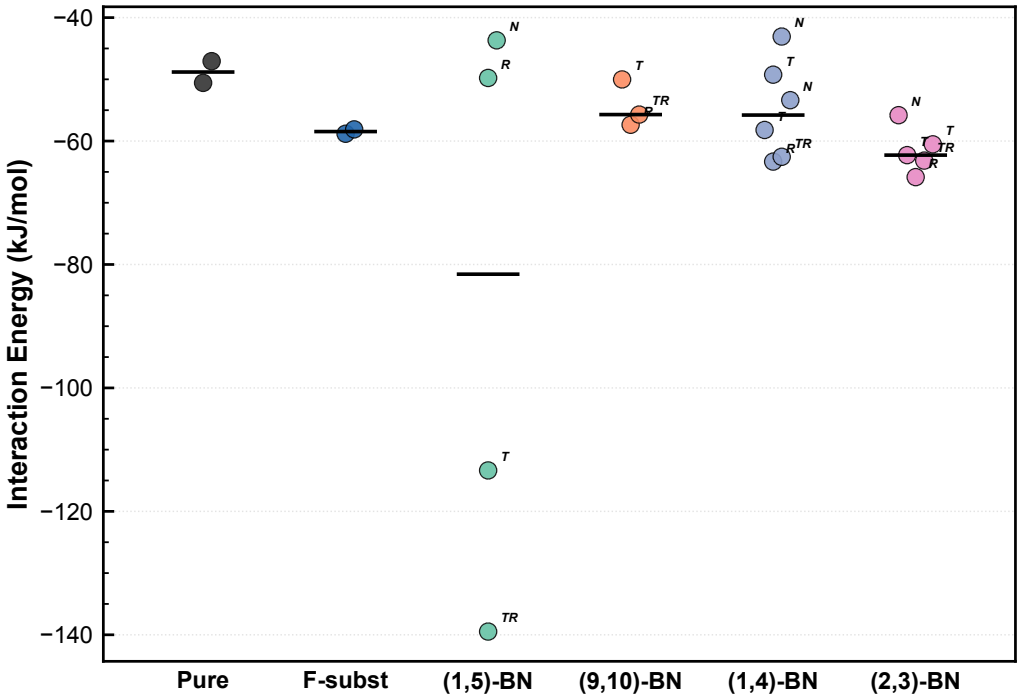


Figure 38: The interaction energy (E_{int} in kJ/mol) distribution for all considered anthracene dimers and substituted dimers. Each color corresponds to a different substitution of anthracene monomer (black is for pure anthracene). The type of BN-substituted dimer is also considered with N/R/T/RT abbreviation (see Figure 35).

Table 14: The interaction energy (E_{int} in kJ/mol) and the buried surface area (BSA in a_0^2) for X-dimer and glike-dimer orientations for anthracene, perfluoroanthracene and selected BN-substituted anthracene molecules.

Molecule	E_{int}, kJ/mol			BSA, a_0^2	
	glike-dimer	X-dimer	ΔE_{int}	glike-dimer	X-dimer
Anthracene	−47.26	−49.96	−2.70	807.62	820.69
Perfluroanthracene	−58.11	−58.82	−0.71	1039.70	1056.05
(1,4)-BN-anthracene (N)	−43.07	−53.38	−10.31	825.53	826.74
(1,4)-BN-anthracene (T)	−49.25	−58.21	−8.96	813.84	827.37
(2,3)-BN-anthracene (T)	−62.28	−60.53	1.75	820.11	826.93

The interaction energy is lower by approximately 10 kJ/mol for both dimer types. At the same time, the difference between the X-dimer and the glike-dimer is only -0.71 kJ/mol. For (1,4)-BN, (N) dimers have higher E_{int} than anthracene, but the difference between the X-dimer and the glike-dimer is to -10.31 kJ/mol. (T) dimers

exhibit lower E_{int} than anthracene, and the difference between the two dimer types is approximately 9 kJ/mol. For (2,3)-BN (T), E_{int} is the lowest among the pairs, and the glike-dimer is more stable than the X-dimer (by 1.75 kJ/mol). The BSA values for the BN-substituted and anthracene dimers are nearly identical, falling within the range of 800–830 a_0^2.

It can be concluded that substitutions within the anthracene structure alter both monomeric and dimeric properties. Furthermore, a correlation between these two sets of properties is not consistently observed. Although some dimers were found to be higher in energy than those of unsubstituted anthracene dimers, on average substitution enhances the interaction energy. In addition, the X-dimer configuration remains the most stable.

4.3.4 Optical Properties of Substituted Dimers

To address the question of substitution effect on optical properties of dimers, TD-CAM-B3LYP calculations of UV-Vis spectra were performed for all monomers and dimers. Four different cases of UV-Vis spectra are presented in Figure 39: (a) anthracene monomer and dimers, (b) F-substituted anthracene and dimers, (c) (1,4)-BN-anthracene, the X-dimer and the glike-dimer (N), and (d) (1,5)-BN-anthracene, X-dimer (TR), and the glike-dimer (T).

As discussed in Section 4.2.6, there is no significant change in peak positions going from monomer to dimers, all appear near 350 nm for anthracene. The F-substituted systems exhibit a bathochromic shift. The monomer shows two closely spaced peaks, while both dimers share the same first peak position as the monomer. For (1,4)-BN-anthracene, all structures display a peak near 340 nm, slightly blue-shifted compared to anthracene. What is particularly interesting is the behavior of (1,5)-BN-anthracene. Both the monomer and dimers deviate significantly from the peaks observed in the other cases. The (1,5)-BN-anthracene monomer peak appears near 700 nm, while both dimers are equal to each other but blue-shifted by approximately 100 nm relative to the monomer.

With the exception of (1,5)-BN-anthracene, all other systems exhibit nearly the same first excitation peak. Only (1,5)-BN-anthracene shows a distinct optical response. As discussed before this monomer is also the most exceptional in terms of interaction energy, dipole moment, polarizability, and HOMO–LUMO gap. Therefore, these substitutions do not produce unique spectral shifts for each dimer. However, substitution

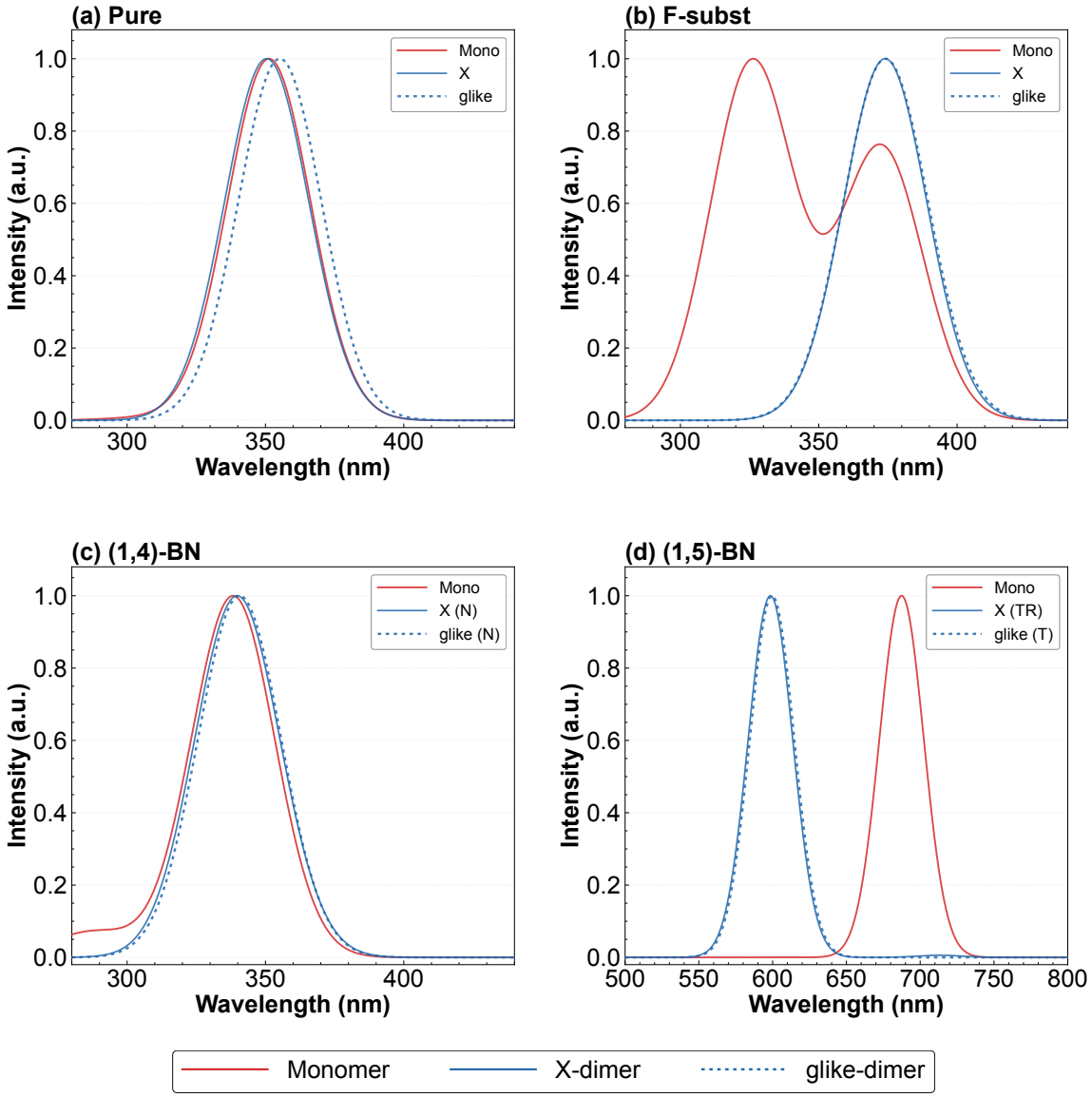


Figure 39: The UV-Vis spectra for selected monomers and dimers of anthracene calculated with TD-CAM-B3LYP/def2-TZVP. The following colors and lines are chosen: for the monomer is solid red, for the X-dimer blue solid line and for the glike-dimer blue dotted line in each case. (a) anthracene monomer and dimers; (b) F-subsituted anthracene and dimers; (c) (1,4)-BN-anthracene monomer and dimers; (d) (1,5)-BN-anthracene monomer and dimers.

polarizing the molecules yields a significant shif the UV-Vis absorption peak for both monomer and dimers, simultaneously, as shown for (1,4)-BN-anthracene and fluorinated anthracene.

4.3.5 Discussion

In this section, the influence of F- and BN-substitutions on the strength of $\pi - \pi$ stacking was investigated, along with the relative stability of the X-dimer and the glike-dimer orientations and the effect on UV-Vis spectra. It was shown that the presence of a large dipole moment and high polarizability in the monomer is reflected in the exceptionally high stability of the (1,5)-BN dimer relative to its monomers. Although some individual dimers exhibit higher E_{int} values than unsubstituted anthracene, on average both BN- and F-substitutions lower E_{int}, resulting in stronger $\pi - \pi$ interactions. The BSA values for BN-substituted dimers largely fall in the same range as unsubstituted anthracene (near 820 a_0^2), whereas fluorinated dimers possess a notably larger BSA (near 1000 a_0^2) due to the larger van der Waals radius of fluorine. It is worth noting that among the BN-substituted dimers considered, the majority of the more stable structures adopt the X-dimer orientation; pairwise comparison of BN-substituted X-dimers and the glike-dimers shows that in two out of three cases the X-dimer is more stable. For the fluorinated derivative, the difference between the two dimer orientations is -0.71 kJ/mol, which is negligibly small.

Subsequently, the influence of substitution on the UV-Vis spectra of monomers and dimers was examined. With the exception of (1,5)-BN-anthracene, all other cases, including the fluorinated derivative, show that the dimer absorption peak coincides with that of the monomer, and any spectral shift originates from the monomer itself. Fluorine substitution induces a bathochromic (red) shift, while (1,4)-BN-anthracene causes a hypsochromic (blue) shift. For (1,5)-BN-anthracene, the monomer exhibits a strongly red-shifted peak (near 700 nm), whereas both dimers are blue-shifted relative to the monomer (near 600 nm).

While substitution offers one route to tuning dimer properties, another dimension is extension of the π-system. It is possible to explore how far these trends extend beyond pentacene.

4.4 Extended Linear Acenes

$\pi - \pi$ stacking in acene dimers ranging from naphthalene to decacene (Figure 40a), with the aim of establishing how the interaction trends evolve with increasing system size are examined in this section. Already within the naphthalene-to-pentacene series, a progressive narrowing of the HOMO–LUMO gap is observed as the number of fused

rings increases[415]. Recent theoretical studies have demonstrated that beyond a hexacene length, DFT treatments based on a closed-shell reference fail to capture the correct physics[416–418].

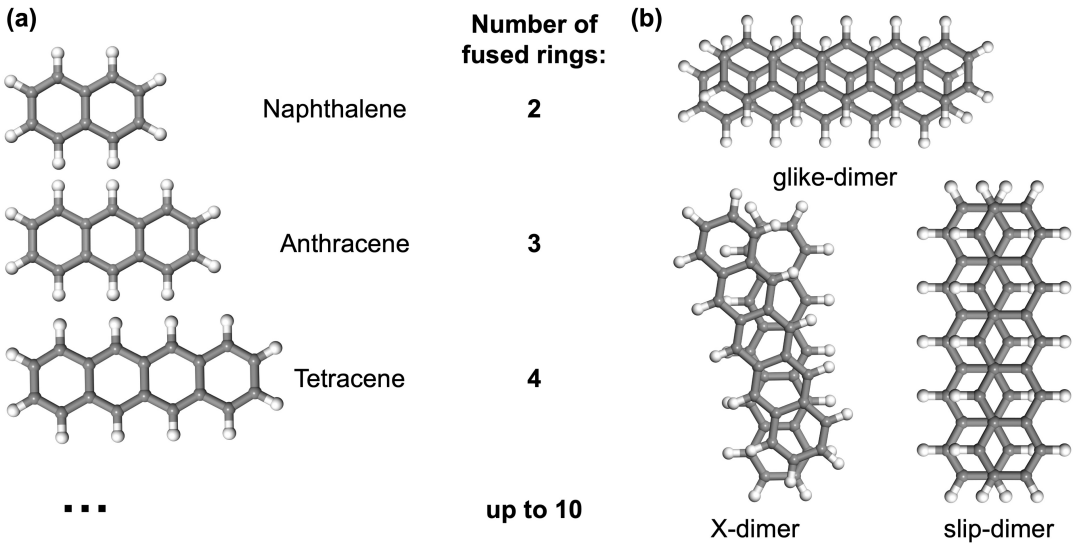


Figure 40: Studied (a) acene monomers up to 10 fused rings (decacene) and (b) dimer types.

The extremely small singlet–triplet gaps and low ionization potentials of higher acenes give rise to pronounced chemical reactivity, manifested in their susceptibility to photoinduced oxidation and spontaneous dimerization. These limitations restricted for a long time the study of such systems to matrix-isolation conditions[255, 291, 419]. However, advances in on-surface synthesis and hydrogen-protected precursor strategies have not only confirmed the existence of acenes up to fifteen fused rings, but also enabled detailed visualization of their electronic resonances by means of atomic force microscopy[420]. The combination of exceptional charge-transport properties, relevant to field-effect transistors, and the capacity for efficient singlet-fission renders longer acenes as a promising platform for next-generation organic optoelectronics and quantum materials with tunable magnetism[255, 291].

In this context, several questions naturally arise:

(i) Does $\pi - \pi$ stacking scale with increasing acene length.

(ii) How adequately can systems beyond pentacene be described within standard DFT-based methods.

(iii) Which dimer configurations continue to exhibit linear trends in their interaction energies and which deviate.

(iv) What happens to the spin-state in both monomers and dimers as the chain grows.

These questions are addressed in the present section.

To provide a comprehensive picture, a third dimer configuration, the slip-dimer, was included alongside the two geometries discussed above. This arrangement, identified as the second most stable stacking motif in the systematic study by Cabaleiro-Lago and co-workers[32], differs from the the glike-dimer in that the displacement of the second monomer relative to the first occurs along a single direction only (see Figure 40b).

4.4.1 Interaction Energy Trends

First, interaction energy calculations were performed for the three dimer configurations across systems containing 2 to 10 fused rings. The resulting dependence of the interaction energy on the number of fused rings is presented in Figure 41a. A common trend is immediately apparent: the interaction energy decreases steadily with increasing system size. This effect is most pronounced for the X-dimer. Whereas for naphthalene the energy differences among the three configurations are small, the curves diverge progressively as the acene chain grows. The X-dimer consistently exhibits the strongest binding, followed by the glike-dimer and the slip-dimer, respectively. More importantly, beyond seven fused rings (heptacene) this trend breaks down. The deviation is particularly evident for the slip-dimer, while the X-dimer curve also begins to depart from the near-linear behavior that persists for the glike-dimer. On the basis of these observations, a linear and predictable scaling of the interaction energy can be identified for all three dimer types within the range of 2 to 7 fused rings; beyond this point, the trend can no longer be considered systematic. This raises a natural question: what is the origin of this breakdown?

To address this question, the interplanar distance of the inner six carbon atoms in the middle ring, denoted o_6, was calculated. This parameter describes the distance between the centroids of the six central carbon atoms (or near center) measured along the direction of the plane normal (i.e., the direction normal to the molecular plane). The script for computing o_6 of dimers was provided by Andrea Buchwald and Professor Reinhold F. Fink from the University of Tübingen. The resulting curves are presented in Figure 41b. Interestingly, the onset of the trend breakdown identified in the 7-to-10 ring range is already apparent for the heptacene X-dimer, which exhibits a substantial variation of o_6 from approximately 3.2 to 2.9 Å. By contrast, the glike-dimer and

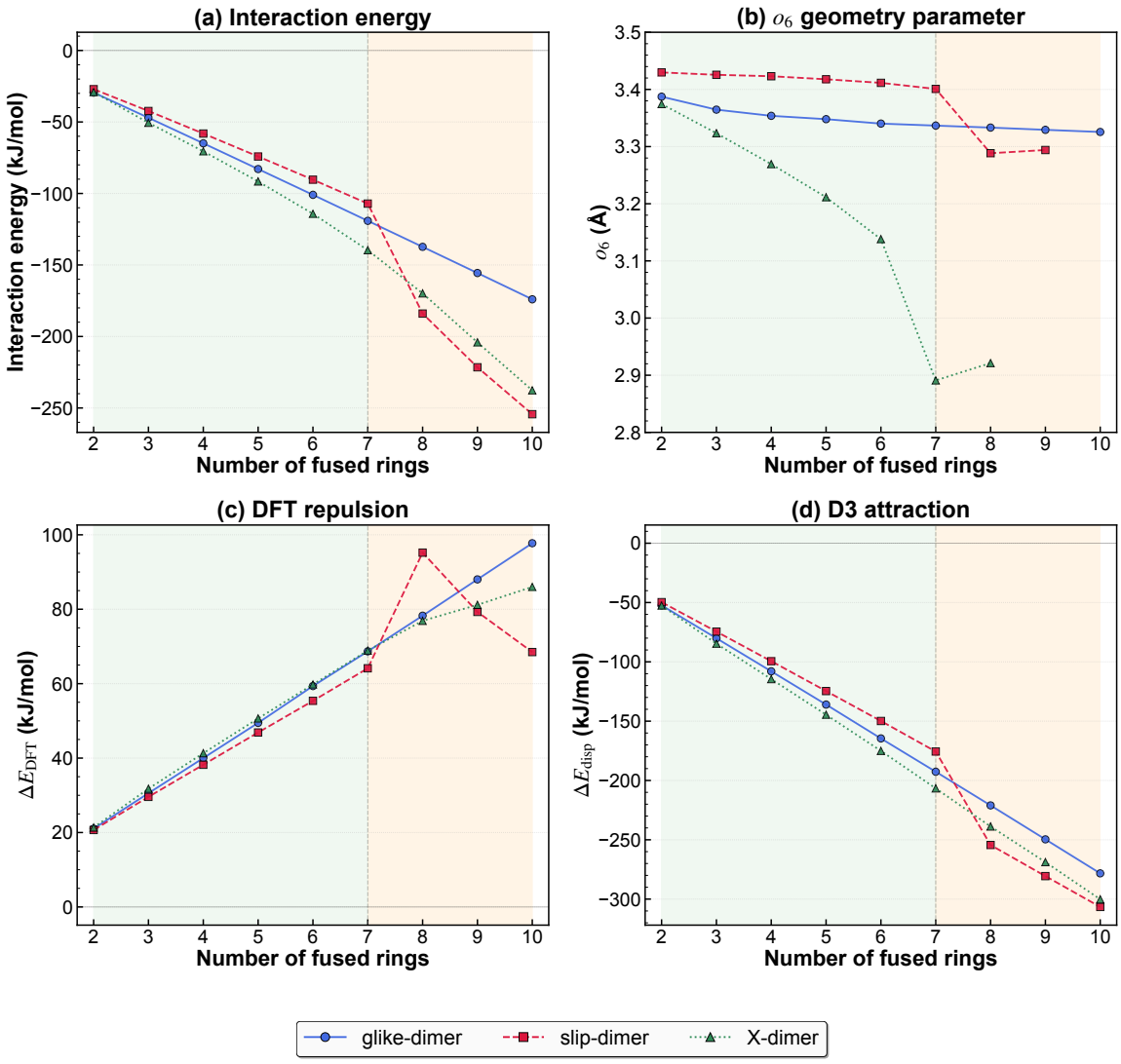


Figure 41: (a) The interaction energy (in kJ/mol) as a function of the number of fused rings. Normal trend shows linearity of E_{int}, broken trend shows non-linear behavior of the interaction energy. (b) The interplane distance of the inner six carbon atoms (o_6 in Å) as a function of the number of fused rings. (c) and (d) Energy components of B3LYP-D3(BJ) interaction energy in kJ/mol: (c) B3LYP (DFT) energy (repulsion); (d) D3(BJ) dispersion energy component (attraction). Each color and marker shape denotes the dimer type (the glike-, slip-, and X-dimers).

slip-dimer show a modest decrease in o_6 with increasing system size. For the slip-dimer, a large o_6 distortion is observed only upon going from heptacene to octacene, whereas the glike-dimer maintains a linear trend up to decacene. Unfortunately, for the slip-dimer at the decacene size and for the X-dimer at the nonacene and decacene sizes, several geometrical distortions of the molecular planes preclude a meaningful determination of o_6 (examples are provided in Appendix for Chapter 4, Section 4.4). Consequently, these data points are absent from the diagram.

A further question arises: given the geometrical distortions observed in the larger acenes. And which contributions to the interaction energy are more significantly affected, the dispersion component (attractive effect on E_{int}) or the DFT energy (repulsive effect on E_{int}). Both contributions are plotted in Figure 41c,d. Notably, the X-dimer and the glike-dimer exhibit linear changes in both energy components across the entire size range. In particular, although variations in o_6 for the X-dimer were already evident from heptacene onward, no signatures of non-linear behavior are found in either the B3LYP or the dispersion component D3(BJ) of the B3LYP-D3(BJ) energy. Only the slip-dimer displays a deviation from linearity in both components beyond heptacene: the B3LYP energy increases and the dispersion contribution D3(BJ) decreases upon going from heptacene to octacene. Beyond this point, however, the slip-dimer recovers a linear trend that persists from octacene through decacene.

Comparing BSA with the interaction energy reveals several noteworthy observations (Figure 42). First, despite the geometrical distortions and the breakdown of regular trends beyond the heptacene size, the linearity of the interaction energy with respect to BSA is preserved across all three dimer types ($R^2 = 0.92$). This suggests that that BSA serves as a universal feature for the considered systems. Second, while the interaction energy trends of the three dimer configurations are nearly identically monotonic for acenes up to heptacene, they diverge beyond this point into two distinct branches: one corresponding to the glike-dimer and the other to the X-dimer and the slip-dimer. Third, after heptacene size the X-dimer and the slip-dimer follow a nearly coincident trend of increasing interaction energy, with only minor deviations between the two configurations.

4.4.2 Electronic Structure Evolution in Extended Acenes

The following aspect to be addressed is the evolution of the electronic structure with increasing acene size. To this end, the following properties were systematically evaluated: the singlet–triplet gap (ΔE_{S-T}), vIP, vEA, vFG, and E_{H-L}. As discussed

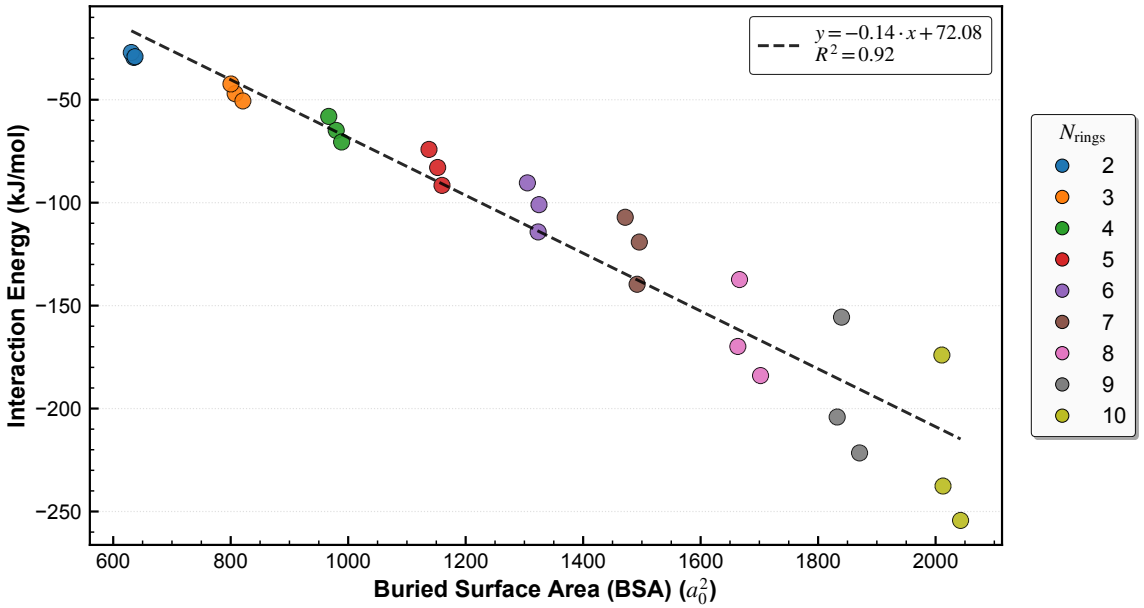


Figure 42: The interaction energy (E_{int} in kJ/mol) as a function of the buried surface area (BSA in a_0^2). Each color corresponds to a different number of fused rings (N_{rings}).

in the preceding sections, both theoretical and experimental evidence point to the emergence of open-shell character and possible polyradical ground states in higher acene homologues. Accordingly, the singlet–triplet energy gap was computed for all dimer configurations as well as for the isolated monomers, enabling a direct comparison of the trends.

The resulting curves are presented in Figure 43. Across the entire size range, monomer, the glike-dimer, and the X-dimer exhibit monotonic trends up to decacene. The slip-dimer deviates from regularity beyond heptacene. Notably, at the heptacene size the singlet–triplet gap is near or above 50 kJ/mol for all systems. For the glike-dimer and the monomer, the singlet–triplet gap becomes positive beyond octacene, while for the slip-dimer and the X-dimer it does not change, a consequence of the geometrical distortions discussed above.

Trends in vIP and vEA were also examined and compared with the corresponding monomer values (see Figure 44a,b). For all dimer configurations up to heptacene, the difference in vIP relative to the monomer remains approximately constant. Beyond heptacene, the slip-dimer and X-dimer exhibit higher vIP values than the glike-dimer, and by decacene the vIP of the slip-dimer approaches that of the isolated monomer.

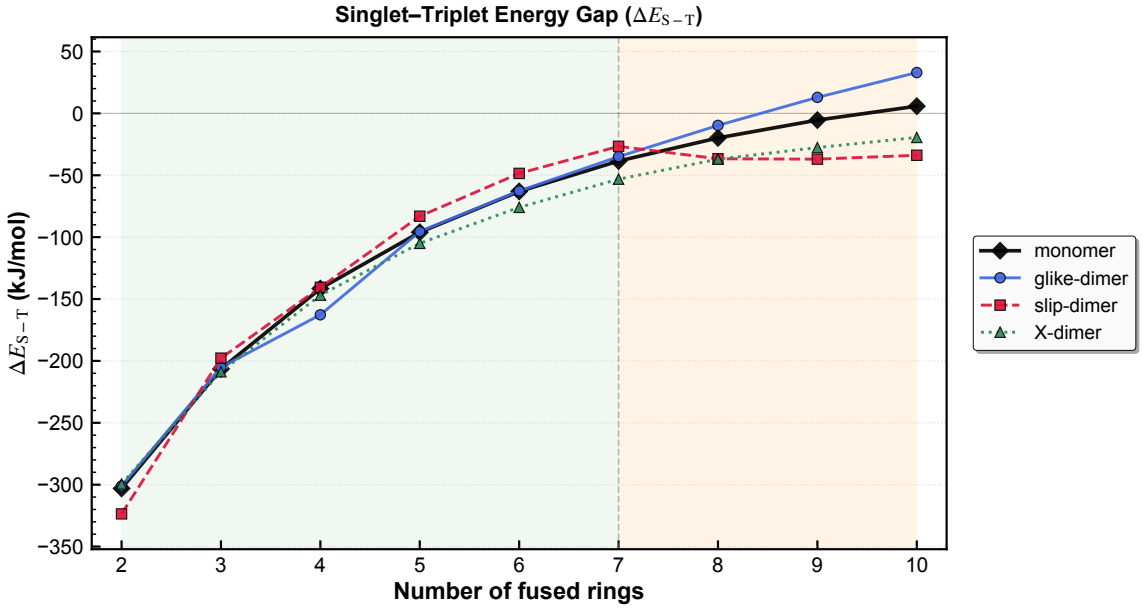


Figure 43: Singlet-triplet energy gap ($\Delta E_{S-T} = E_S - E_T$ in kJ/mol) for different dimers and monomers based on the number of fused rings.

Regarding vEA, all three dimers display higher values than the monomer up to heptacene. Within this range, the glike-dimer and the X-dimer follow nearly identical trends, while the slip-dimer consistently shows a somewhat higher vEA. Beyond heptacene, the regular trend breaks down. For instance, vEA values of the slip-dimer and X-dimer converge toward those of the monomer, effectively erasing the dimer stabilization observed for shorter acenes.

Finally, the HOMO-LUMO gap and vFG were examined (see Figure 44c,d) as these quantities provide direct insight into the electronic stability of the systems and their potential suitability as organic semiconductors. Up to heptacene, the monomer, the glike-dimer, and the X-dimer exhibit nearly identical HOMO–LUMO gaps with minor deviations, whereas the slip-dimer falls outside this trend, displaying consistently lower values. Beyond octacene, the behavior of the slip-dimer and the X-dimer deviate from the preceding trend. Their HOMO–LUMO gaps become larger than those of the glike-dimer. Monomer and the glike-dimer, in turn, remain so close that they are practically indistinguishable.

Regarding the vFG, the monomer consistently exhibits higher values than all dimer configurations. The X-dimer and the glike-dimer behave nearly identically up to heptacene, while the slip-dimer again shows lower values. This is a consequence of its

higher vEA, given that vIP was approximately the same for all three dimers in this size range (Figure 44a,b).

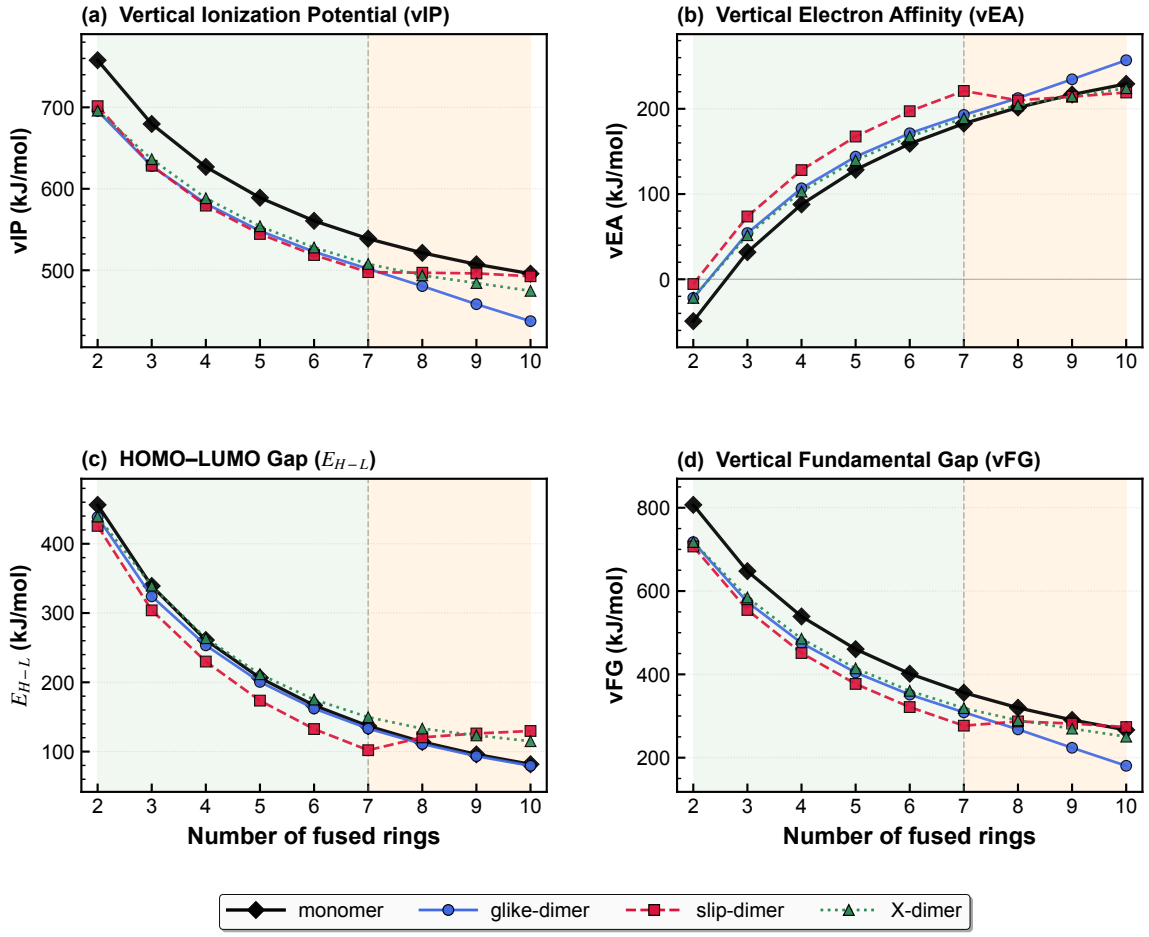


Figure 44: Different electronic features for dimers and monomers of the acene series (in kJ/mol): vIP (a); vEA (b); E_{H-L} (c); vFG (d). vIP (left) and vEA (right) for dimers and monomers of the acene series (in kJ/mol).

4.4.3 Discussion

The results presented above reveal a consistent pattern: nearly all electronic and structural properties of acene dimers undergo a qualitative change around heptacene. This threshold likely marks the point beyond which a closed-shell hybrid DFT description becomes inadequate, and more sophisticated multireference approaches, such as CASSCF/CASPT2 or DMRG-based methods, would be required to capture the correct properties of these systems[418].

Importantly, this breakdown in regularity is observed only for the X-dimer and the slip-dimer, whereas the glike-dimer exhibits considerably more predictable behavior, with no significant deviations in the property trends even for the largest acenes considered. Probably, the displaced geometry of the glike-dimer is inherently less susceptible to the emerging open-shell character than the geometries of the X-dimer and slip-dimer (see Appendix for Chapter 4, Section 4.4). Notably, despite the geometrical distortions observed for the latter two orientations, the linearity of the interaction energy with respect to BSA is preserved, implying that even geometrically distorted dimers retain a well-defined interaction energy for a given molecular size, a finding that underscores the robustness of BSA as a transferable descriptor.

The observation that the HOMO–LUMO gap, vIP, vEA, and vFG of certain dimers converge toward the corresponding monomer values at large acene sizes is physically significant. This convergence suggests a progressive increase in the cooperative electronic stabilization within the dimer, possibly linked to the localization of frontier orbitals driven by the increasing polyradical character of the ground state.

In conclusion, the breakdown of regular trends beyond heptacene may partially reflect intrinsic limitations of the B3LYP-D3(BJ) framework when applied to systems with a substantial multireference character. These findings motivate future investigations employing multiconfigurational methods and, equally importantly, help delineate the acene size range within which B3LYP-D3(BJ) can be reliably applied for the description of $\pi - \pi$ stacking interactions.

4.5 *cata*-Condensed Polybenzoid Isomers

The following section is devoted to the most stable dimers of acene isomers, examining how their properties are correlated with one another. In particular, the question of whether dimer properties can be predicted from monomer descriptors alone, or whether certain dimer properties are interdependent. This issue will be addressed, along with related aspects of structure–property relationships across the isomer series.

In the era of big data in chemistry, the preparation of specific datasets that can be further applied to screen properties and structures, as well as to develop new research methods, for example those based on machine learning, has become increasingly important[421–424]. This trend has naturally extended to non-covalent complexes. To date, numerous databases have been developed and are widely used for benchmarking various *ab initio* methods, training modern machine-learned potentials[425], and

screening specific properties of complexes[293]. Among the most well-known are NCIE[426], S22[40, 427], GMTKN30[41], A^2ID[428], S66[429], S12L[430], L7[431, 432], A24[433], as well as NENCI-2021[434] (a database of non-equilibrium interactions). The atlas of some of NCI databases is available[435]. But mostly they don't touch $\pi-\pi$ stacking interaction. Given the above, there is a clear need for new databases of non-covalent interactions, particularly in the domain of polycyclic aromatic compounds, where $\pi-\pi$ stacking is the dominant interaction[436].

It is worth noting separately that PAH species are relevant not only for potential applications in organic electronics but also for their presence in cosmic environments[207, 437, 438]. For instance, it has been demonstrated that PAHs are the carriers of the unidentified infrared emission bands observed in astronomical spectra[437–440]. Entire databases, such as NASA PAHdb[441], have been created for PAH identification, and the presence of non-covalent complexes would further complicate the picture in the absence of reference databases for comparison.

Building upon the publicly available **COMPAS Project database**[293], originally designed for screening cata-condensed polybenzoid topologies with the aim of enabling rational design of functional aromatic molecules for organic electronics. This section introduces its extension: **The COMPAD Project** (**COMputational database of Polycyclic Aromatic Dimers: the *cata*-Condensed Polybenzenoid Hydrocarbons Case**). The primary objectives of this new database are to enable screening of the properties of $\pi-\pi$ dimers of cata-condensed polybenzenoid hydrocarbons, to serve as a benchmark for various *ab initio* methods, and to facilitate the training of new machine learning (ML) approaches. In this section, a new database is presented, that database contains only the most stable dimers for each monomer, ranging in size from benzene to all isomers of hexacene. However, it will continue to be expanded over time to increase its scope and utility. This chapter provides the analysis of existed trends and correlations between monomer and dimer properties.

4.5.1 Construction of the COMPAD Project Database

Initially, all monomers and their geometries for the range of one to six fused rings were taken from the COMPAS Project database[293]. This yielded a total of 58 monomers. The distribution of the number of monomers as a function of size is presented in Table 15. For example, all monomers of tetracene are shown in Figure 45.

The construction of this database involved the following steps:

Table 15: The number of *cata*-condensed polybenzoid hydrocarbon monomers as a function of the number of fused rings.

Number of Fused Rings	Number of Monomers
1	1
2	1
3	2
4	5
5	12
6	37
Total	58

(i) Generation of initial dimer structures. Initial structures were generated by duplicating the monomer coordinates with a translation along the axis perpendicular to the monomer plane.

(ii) Conformational search and filtering. The initial dimers were subjected to metadynamics-based conformational sampling using CREST with GFN2-xTB for non-covalent complexes under default parameters. The six most stable conformers from the sorted CREST output were selected for further optimization.

(iii) Two-stage geometry optimization. The six selected conformers were first optimized at the B3LYP-D3(BJ)/def2-SVP level of theory. The most stable structure among them was then reoptimized using the larger def2-TZVP basis set.

(iv) Single-point energy calculations. For each final dimer geometry, two single-point calculations were performed: one at the TD-CAM-B3LYP/def2-TZVP level to obtain the first excitation energy, and another at the B3LYP-D3(BJ)/def2-TZVPP level to compute the remaining monomer and dimer properties.

The complete data preparation workflow is shown in Figure 45.

4.5.2 Monomer and Dimer Electronic Properties

The following electronic properties were obtained for monomers and their dimers: the HOMO–LUMO gap, first excitation energy, vertical ionization potential, vertical electron affinity, dipole moment, and isotropic polarizability. As previously demonstrated for acene molecules, dimeric electronic properties differ only slightly from monomeric

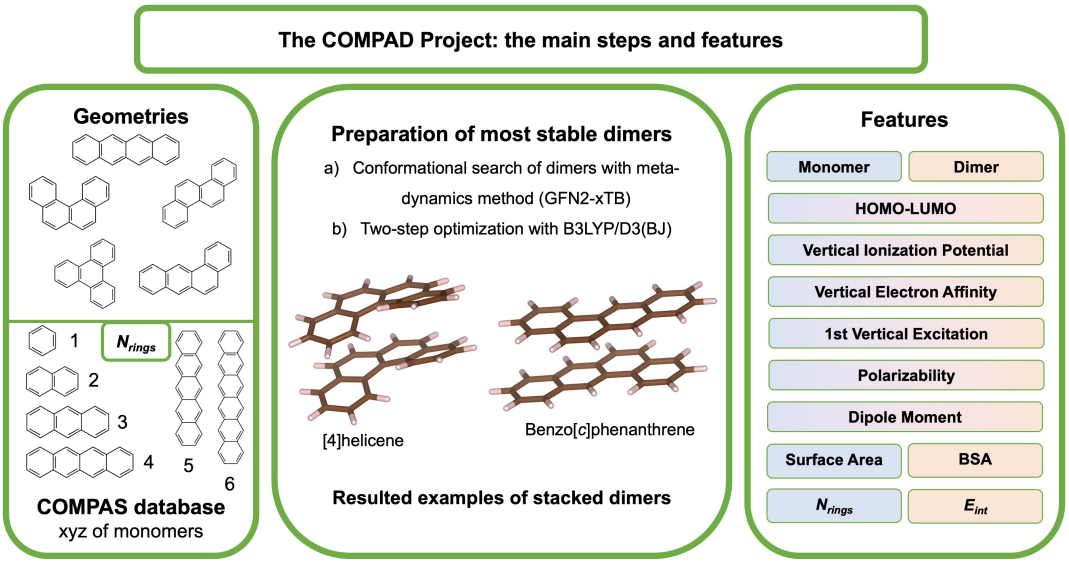


Figure 45: The full scheme of the COMPAD database with main steps: (left) considered monomer sizes that taken for database and examples of tetracene isomers; (center) two main steps of preparation final dimer geometries - conformational search and DFT optimization; (right) all features of database for monomers (blue left) and dimers (orange right). Some of them exist for both structures.

ones while following the same trends. Therefore, the correlation between monomer and dimer properties was examined first, as shown in Figure 46.

The HOMO–LUMO gap, consistent with earlier findings for acenes, is slightly reduced for dimers relative to monomers, but otherwise the two sets of values are strongly correlated, that well observed in Figure 47. For vIP and vEA, a clear linear relationship is also observed. However, dimers exhibit lower vIP and higher vEA compared to monomers, suggesting that the charge in ionized states is more effectively delocalized over the larger system, thereby stabilizing both positively and negatively charged species. In the case of E_{exc}, perfect agreement between dimer and monomer values is observed, with negligibly small deviations. Isotropic polarizability shows also a linear correlation, with dimer values approximately twice those of the corresponding monomers. The dipole moment, however, displays no discernible trend: the data points are widely scattered, likely owing to the small magnitudes involved, with a large fraction of monomer values equal or close to zero.

In summary, most electronic properties of monomers are well correlated with those of dimers, differing primarily by a scaling coefficient, as exemplified by α, which roughly

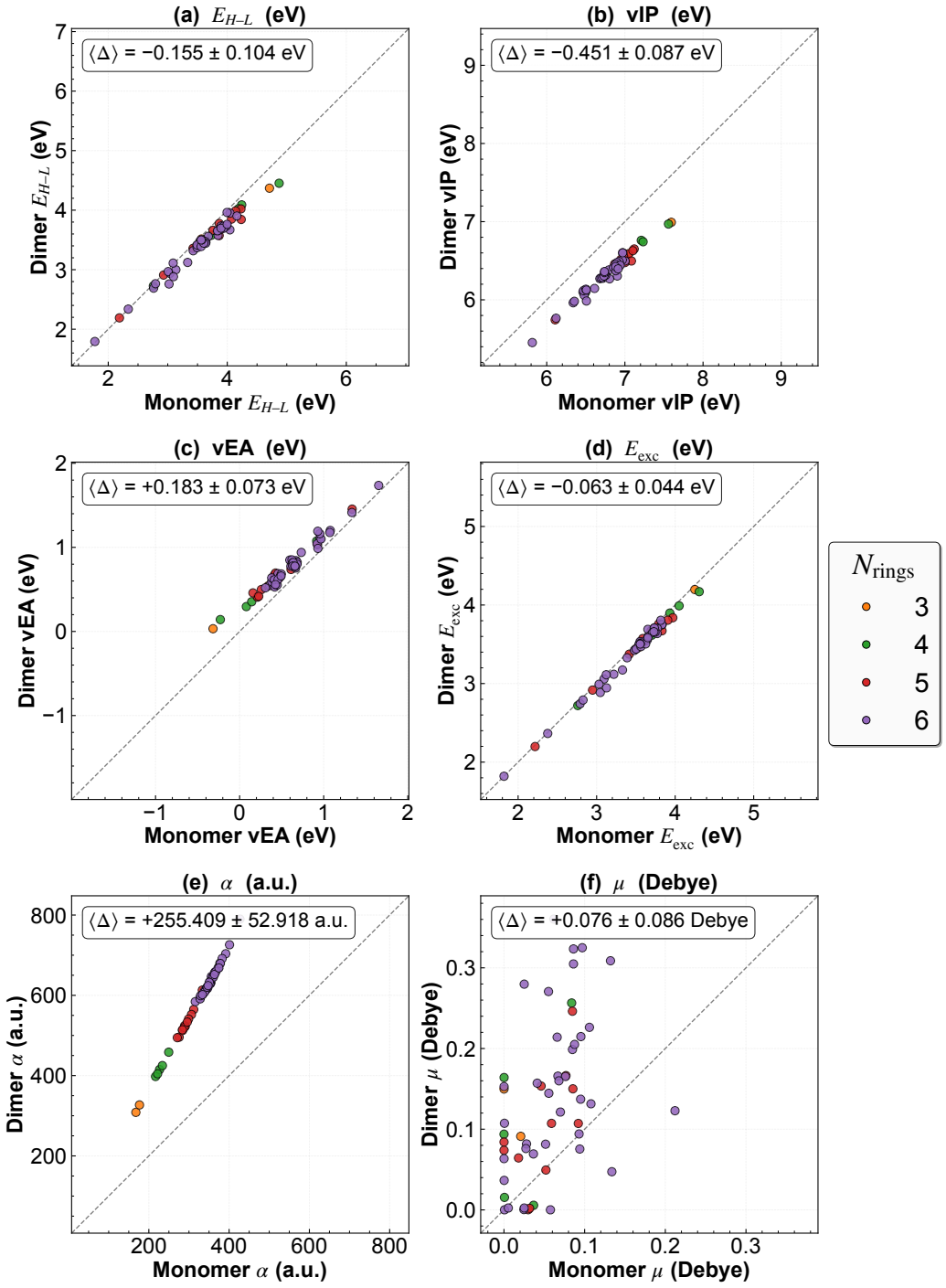


Figure 46: The correlation between monomeric and dimeric electronic properties for all isomers composed of 3-6 fused rings. On axis X are monomeric electronic properties, on axis Y are dimeric electronic properties. There are the following features: (a) HOMO–LUMO gap E_{H-L} (eV); (b) vertical ionization potential vIP (eV); (c) vertical electron affinity vEA (eV); (d) first excitation energy E_{exc} (eV); (e) isotropic polarizability α (a.u.), (f) dipole moment μ (D).

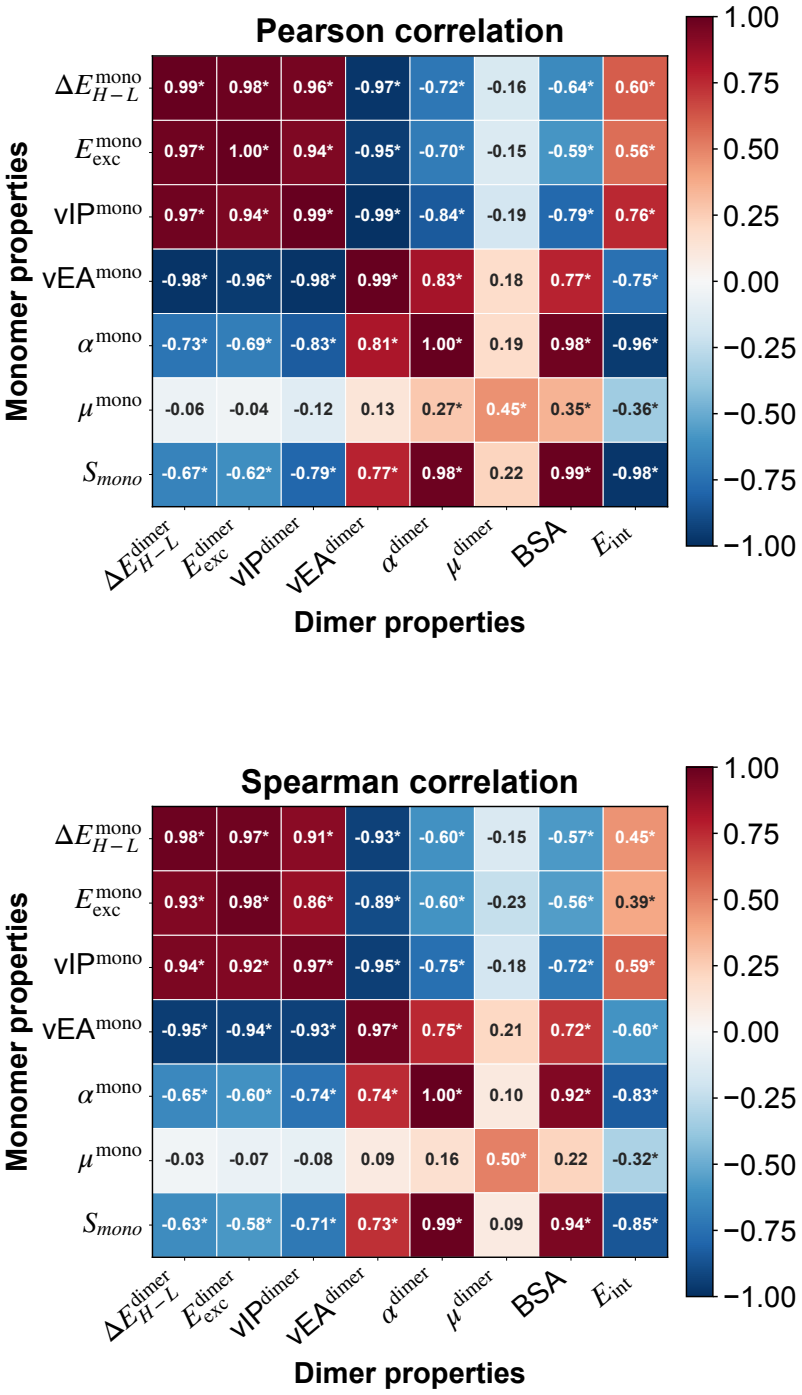


Figure 47: The view of (top) Pearson Correlation (linear), (bottom) Spearman Correlation (non-linear) matrices. Two correlation matrices between the dimeric (columns) and monomeric (rows) features, including interaction energy (E_{int}), buried surface area (BSA) of dimers and surface are of monomers (S_{mono}): HOMO–LUMO gap (ΔE_{H-L}), first excitation energy (E_{exc}), vertical ionization potential (vIP), vertical electron affinity (vEA), isotropic polarizability (α), dipole moment (μ).

doubles, and by vIP and vEA, which shift systematically toward greater charge-state stability.

4.5.3 Interaction Energy Correlation with BSA and Isotropic Polarizability

The examination of properties unique to dimers is of particular interest. Especially, the interaction energy (E_{int}) and BSA. How do they correlate with each other? Whether do and to what extent they correlate with individual monomer properties? These questions are discussed here.

To address these questions, Pearson and Spearman correlation heatmaps[442] were computed, as presented in Figure 47. Both matrices include all electronic and structural properties of the monomers alongside two dimer-specific properties, E_{int} and BSA. Also, two matrices look identical, that means the connection between features is near linear. It can be concluded that, to varying degrees, all properties correlate with E_{int} and BSA, with the exception of the dipole moment. E_{int} and BSA are strongly correlated with each other (-0.98 by Pearson), consistent with earlier findings for extended acenes: the larger the BSA, the lower the E_{int} and the stronger the interaction in the case of π–π stacking. The dependence of E_{int} on BSA for number of fused rings in monomer $N_{rings} = 3 - 6$ is shown in Figure 48, yielding $R^2 = 0.93$. The mean absolute error (MAE) and the root mean square error (RMSE) are 3.41 and 4.26 kJ/mol, respectively. A clear shift along the regression line is observed with increasing number of fused rings, toward lower E_{int} and larger BSA, as each additional ring participates in the stabilization of the non-covalent interaction.

Notably, several outlying points deviate from the general trend, lying further toward the upper left of the plot. These correspond to three isomers with five and six fused rings: one tetracene isomer, one pentacene isomer and one hexacene isomer. All three points exhibit higher E_{int} (i.e., weaker interaction) and lower BSA values compared to other members of their respective families. Inspection of their structures, shown in Figure 49 and their properties in Table 16, reveals that all three isomers possess a pronounced structural curvature, which can reasonably be expected to hinder dispersion and electrostatic interactions, thereby reducing the overall interaction energy due to a smaller contact surface area.

Returning to the monomer properties correlated with E_{int}, isotropic polarizability also exhibits a very strong correlation (-0.96 in Pearson correlation). It is well established that polarizability is one of the most fundamental molecular descriptors governing

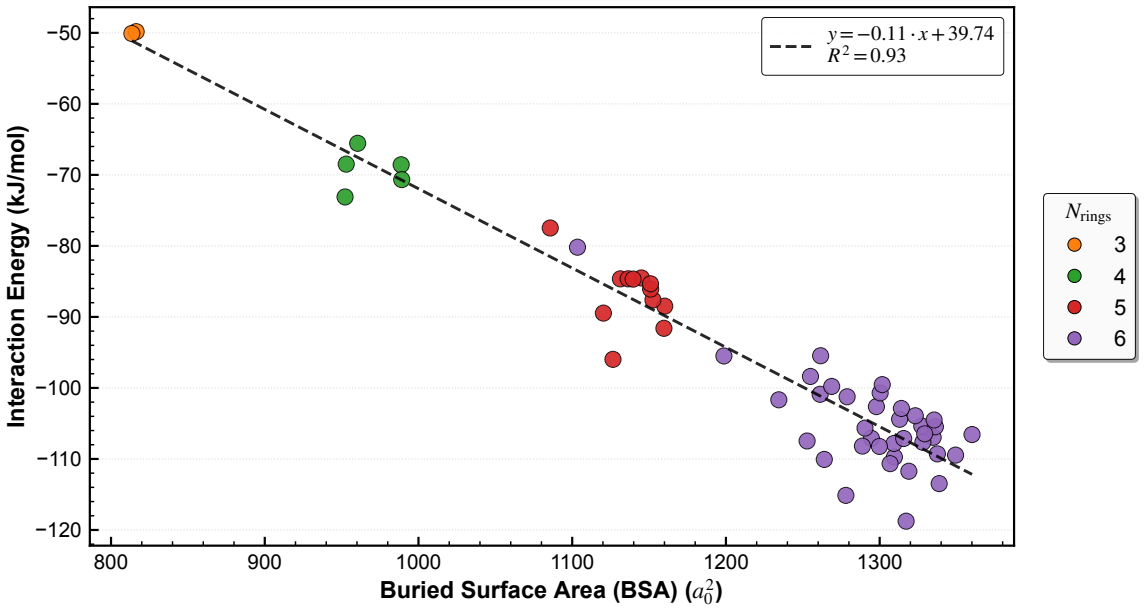


Figure 48: The interaction energy (E_{int} in kJ/mol) as a function of the buried surface area (BSA in a_0^2) for all isomers composed of 3-6 fused rings. Each color corresponds to a different number of fused rings (N_{rings}).

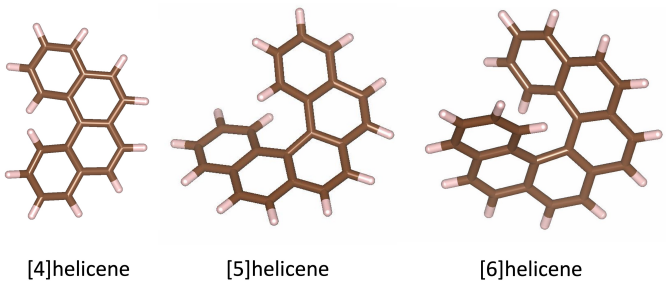


Figure 49: The examples of helicene molecules with four, five and six fused rings. (From the left to the right is the order of how they are located in Figure 48).

Table 16: Interaction energy (E_{int} in kJ/mol), buried surface are (BSA in a_0^2) and isotropic polarizability (α in a.u.) for helicene molecules with four, five and six fused rings.

Molecule	E_{int}, kJ/mol	BSA, a_0^2	α, a.u.
[4]helicene	-65.56	960.45	221.75
[5]helicene	-77.48	1085.71	270.95
[6]helicene	-80.19	1103.43	315.77

the strength of non-covalent interactions, particularly van der Waals forces[345]. It is therefore logical that these two quantities are strongly correlated: the higher the polarizability, the stronger the interaction and the lower the value of E_{int}. Plotting E_{int} against monomeric isotropic polarizability yields $R^2 = 0.86$ (see Figure 50). The MAE and the RMSE are 4.70 and 5.82 kJ/mol, respectively. Remarkably, within each group the highest polarizability and the more linear-dependent E_{int} correspond to the acene dimers (from anthracene to hexacene, inclusive). In every case, the most stable configuration was the previously studied X-dimer, which did not stand out notably in terms of BSA but ranked first in isotropic polarizability among all *cata*-condensed polybenzoids of the same ring count.

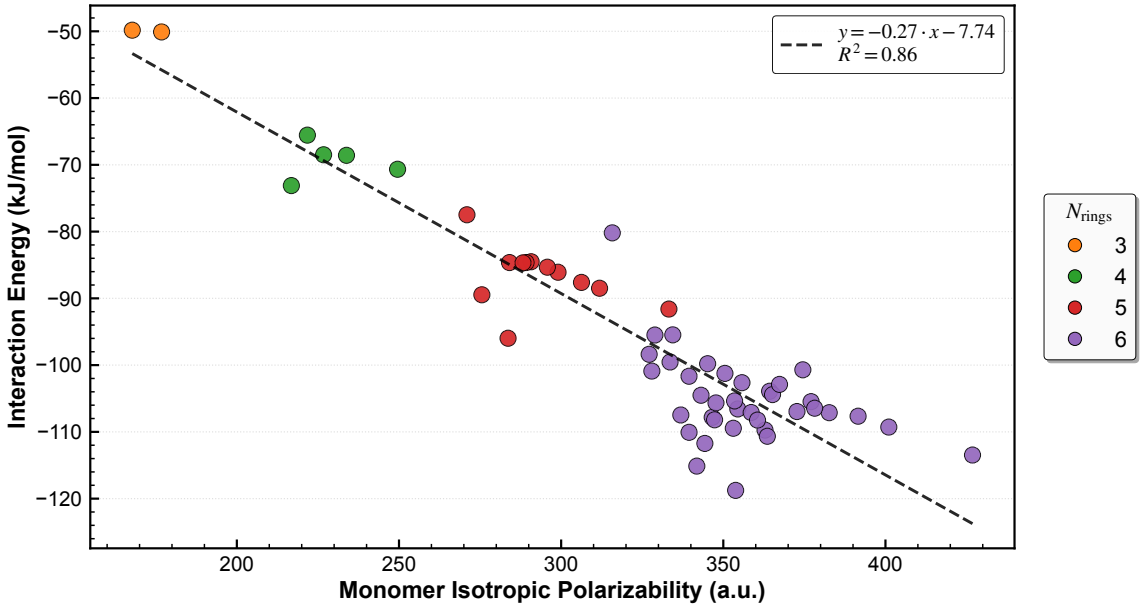


Figure 50: The interaction energy (E_{int} in kJ/mol) as a function of the isotropic polarizability (α in a.u.) for all isomers composed of 3-6 fused rings. Each color corresponds to a different number of fused rings (N_{rings}).

As demonstrated above, E_{int} correlates with many monomer properties, particularly structural descriptors and isotropic polarizability. The correlations with vIP and vEA were also relatively high (0.76 and -0.75, respectively). A natural follow-up question arises: can this results of interaction prediction be systematically improved by considering the combination of different features? And how will different subset of parameters perform in this prediction? To address this question, a machine learning study was carried out to establish which set of features is most predictive. This study is presented in the Appendix for Chapter 4, Section 4.5. as a supplementary addition

to the thesis. It is not reflected in the conclusions, but provides additional insight into how an ML-based analysis can be performed on this dataset.

4.5.4 Discussion

The resulting dataset was analyzed by constructing correlation matrices and training ML methods with different feature subsets. Several previously proposed ideas were reaffirmed, including the universality of BSA as a descriptor reflecting E_{int}. The relationships between monomer and dimer electronic properties were also examined, revealing strong correlations and, in the case of E_{exc}, near-exact agreement between monomer and dimer values. Separately, monomer polarizability was shown to be very strongly correlated with E_{int}.

4.6 Chapter Conclusions

In this chapter, different aspects of non-covalent dimers across a wide range of structures, including acenes, their derivatives, and isomers were systematically investigated. The main methodology was first implemented for acene dimers and subsequently expanded to encompass different substitution patterns, system sizes, and topologies.

Firstly, unexplored systematically the X-dimer was investigated with different methods for four acene molecules. The SAPT analysis revealed the dominant components of the interaction energy that render the X-dimer more stable than the glike-dimer. For acene dimers, the relative stability of the glike-dimer and the X-dimer was examined, and the trend of increasing stability with system size was explored. The buried surface area was identified as a suitable descriptor associated with the interaction energy, though not with the relative stability of different dimer configurations. The geometries and interaction energies obtained at the B3LYP-D3(BJ)/def2-TZVP level were validated against the CC values. A comparison of several NNP-based, semi-empirical, and DFT-based methods was alsoperformed using PNO-LCCSD(T*)-F12 as a reference. The results indicate that most DFT methods reproduce the correct stability trends for dimers, suggesting that they provide a qualitatively accurate picture. The electronic properties of dimers follow the same trends as those of the corresponding monomers, albeit with systematic shifts that stabilize the charged states. The UV-Vis spectra of the acene dimers were shown to be essentially identical to those of the monomers due

to the local excitation character of the transitions. Only sophisticated aggregates as herringbone one can induce bathochromic shifts in the UV-Vis spectra.

Subsequently, it was shown that substitutions with BN and F have a favorable effect on the interaction energy, lowering it further. The strength of the interaction was associated with certain monomeric electronic properties, such as polarizability and dipole moment. The X-dimer was found to be more stable orientation compared to the glike-dimer for nearly all cases. The monomers of substituted acenes exhibited unique UV-Vis spectral features, whereas the spectra of their dimers, with the exception of (1,5)-BN-anthracene, were indistinguishable from those of the monomers.

Extension to larger acene sizes confirms that open-shell and polyradical effects become increasingly pronounced with system size. For example, the singlet-triplet energy gap of octacene falls below 50 kJ/mol for all dimers and the corresponding monomer. As a consequence, the DFT electronic structure description becomes unreliable beyond heptacene at the B3LYP-D3(BJ) level. Significant geometry distortions, manifested as out-of-plane deformations and central pinching of the monomers, are observed starting from heptacene-sized systems for the X-dimer and the slip-dimer. This affects the final interaction energy and other electronic properties, breaking the visible trends observed for acenes smaller compared to heptacene. The interaction energy trends remain consistent with those found in the naphthalene-pentacene range for the glike-dimer and for the X-dimer and the slip-dimer up to heptacene. The interaction energy decreases and BSA increases linearly with size. The X-dimer becomes increasingly more stable relative to the glike-dimer and the slip-dimer.

The database of *cata*-condensed polybenzoids was compiled and analyzed. The buried surface area was shown to be a universal descriptor for non-substituted *cata*-condensed polybenzoids that reliably captures the ordering of interaction energies for geometrically distinct acene isomers. In addition, the monomer polarizability was identified as an effective feature for describing the interaction energy. Most monomeric and dimeric electronic properties are well correlated, with the exception of the dipole moment.

Overall, this chapter demonstrates the unique dimer type for acene systems. The stability of dimers was investigated and confirmed with different substitutions and molecule sizes for acene systems. Then the strong correlation between the interaction energy and different monomeric properties are shown on different acene isomers. Also, it was shown that B3LYP-D3(BJ) and other DFT methods correctly reproduce the stability trends up to certain size of dimers and can adequately describe additional properties such as UV-Vis spectra. The results provide well-founded answers regarding

the relative stabilities of selected dimers and how non-covalent interactions affect electronic and optical properties. These insights can guide future investigations and serve as a basis for screening molecular structures and tuning them toward desired properties.

5 Summary and Conclusion

In the present work, the energy landscapes associated with catalytic systems and non-covalent complexes were investigated using various quantum chemical methods. The selected cases served not only to demonstrate the capabilities of modern computational approaches, but also to critically examine their limitations and define the boundaries of their applicability in the context of chemical modeling.

In the first part, a particular emphasis was placed on the role of using a QM:QM embedding scheme in catalysis. The hierarchical cluster correction scheme was illustrated in the context of ethylene dimerization in the Ni-NU-1000 MOF catalyst. An accurate description of the reaction energy profile was first established at the DLPNO-CCSD(T) level and subsequently used to validate DFT methods for the hierarchical cluster correction scheme, suitable for modeling the full ethylene oligomerization mechanism. Crucially, the CC-corrected TOF of 260 h^{-1} was found to be close to the experimental value of 252 h^{-1}, providing direct experimental support for the predictive power of the scheme. Uncorrected periodic DFT using the PBE functional was found to deviate from the CC-corrected reference by more than 20 kJ/mol for adsorption and desorption steps and by up to 15 kJ/mol for energy barriers, underlining the necessity of such a correction in the first place. Among the tested functionals within the hierarchical cluster correction scheme, TPSSh and M06-L emerged as the most reliable, reproducing the coupled cluster reference for ethylene dimerization with a mean absolute deviation of 3–4 kJ/mol. The ability to extend validated methods to larger and more complex systems is central to the practical value of this approach. Ultimately, this correction scheme provided kinetic insights for the olefin oligomerization in the Ni-NU-1000 catalyst.

Kinetic analysis across different computational methods revealed substantial differences in product distribution, further highlighting the necessity of such benchmarking studies. For instance, with M06-L no ethylene conversion was observed, whereas with TPSSh more than half of the ethylene was converted (63%), yielding oct-1-ene as the primary product with 37% selectivity in the trigonal pore. Complete ethylene conversion

was predicted using the PBE functional, illustrating how uncorrected DFT can lead to qualitatively misleading conclusions. Furthermore, the experimentally observed catalytic activity of specific pore types toward the formation of products heavier than hex-1-ene was confirmed and rationalized by microkinetic analysis based on TPSSh. The trigonal pore was found to have a chain growth up to oct-1-ene, while the hexagonal pore did not produce olefins beyond hex-1-ene, that confirms an experimental evidence for preferential chain growth in the trigonal pore.

In the second part of the work, focused on non-covalent PAH dimers, the relative stability of the previously unexplored X-dimer was established for acene systems and their derivatives. The SAPT analysis identified the dominant interaction energy components, such as dispersion and electrostatics, that render the X-dimer more stable compared to the competing glike-dimer, providing a physically grounded rationale for its preference. It was shown how the choice of computational methods critically affects the predicted stability of the complexes. Most benchmarked DFT methods were nevertheless able to reproduce the correct stability trends. Beyond energetics, particular attention was to the influence of molecular conformation on both the electronic structure and the optical properties of the non-covalent complexes. The UV-Vis spectra of anthracene-based dimers were found to closely match those of the corresponding monomers, in agreement with experimental data, while anthracene substitutions were shown to enhance the interaction energy and to modulate electronic and optical properties. The interplay between molecular geometry, system size, and substitution pattern was analyzed in terms of their individual effects on dimer stability and on the electronic and optical response, with the buried surface area and the monomer polarizability emerging as universal descriptors of the interaction energy across the studied set, providing a rational basis for future computational screening and materials design. The limitations of the DFT description were also identified, with open-shell effects becoming significant beyond heptacene and rendering a reliable DFT-based description of larger acene dimers unfeasible.

In summary, the hierarchical cluster correction scheme offers a transferable framework that can be applied to related catalytic systems to improve the reliability of kinetic modeling, as illustrated quantitatively by the shift from complete (uncorrected) to more than half (corrected) ethylene conversion and by the recovery of the experimentally observed pore-dependent product distribution in the Ni-NU-1000 catalyst. Similarly, the methodological insights gained for non-covalent acene-based systems lay the groundwork for the computationally guided design of functional organic materials. The results can help with quantified stability and tailored electronic and optical

properties. Especially, this design patterns can be supported by compact universal descriptors and a clearly defined DFT applicability domain extending up to heptacene. In the end, these two case studies lead to the same conclusion: reliable predictions of chemical observables require an accurate energy.

A Appendix

A.1 Appendix for Chapter 3, Section 3.2

The Ethylene Dimerization Mechanism

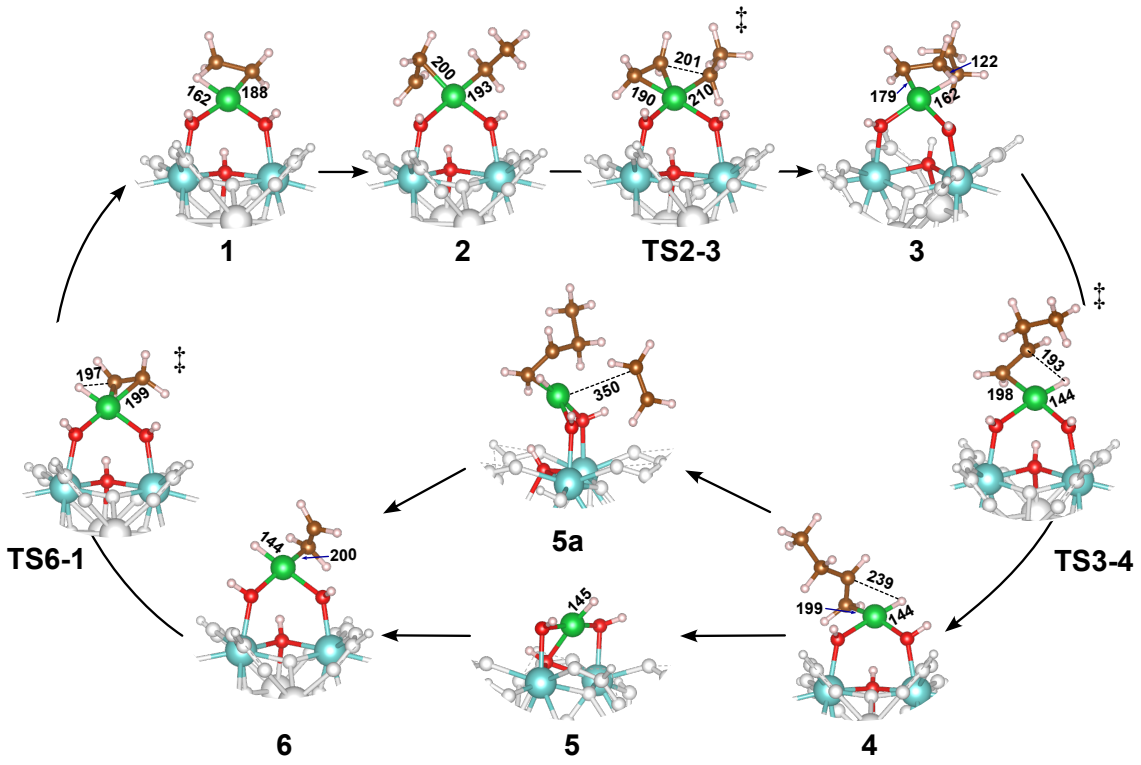


Figure 51: The ethylene dimerization mechanism representation as key stationary structures calculated at the PBE-D3(BJ)/def2-TZVP level using the formate-cluster model. In the visualization, chemical elements are represented by specific colors: Zr (cyan), Ni (green), O (red), C (brown), and H (light pink). The bond length and interatomic distances are represented in pm. Reprinted with permission from [111]. Copyright 2025 American Chemical Society.

A.2 Appendix for Chapter 3, Section 3.3

Reaction Equations for Olefin Oligomerization Catalytic Reaction Network

This section presents the main reactions within the catalytic reaction network operative on Ni-NU-1000 for olefin oligomerization (from Figure 11B and Figure 12). (**gp**) subscript denotes the gas-phase molecule.

1. **Dimerization Process:**

   a) First ethylene adsorption

   $$[\text{Ni-H}] + \text{C}_2\text{H}_{4\,\mathbf{(gp)}} \rightarrow [\text{et-Ni-H}] \tag{76}$$

   b) Ethylene hydrogenation (β-agostic intermediate formation)

   $$[\text{et-Ni-H}] \rightarrow [\text{Ni-ethyl}] \tag{77}$$

   c) Ethylene adsorption (Initial structure formation)

   $$[\text{Ni-ethyl}] + \text{C}_2\text{H}_{4\,\mathbf{(gp)}} \rightarrow [\text{et-Ni-ethyl}] \tag{78}$$

   d) C-C coupling (Dimerization)

   $$[\text{et-Ni-ethyl}] \rightarrow [\text{Ni-butyl}] \tag{79}$$

   e) β-H elimination

   $$[\text{Ni-butyl}] \rightarrow [\text{but-Ni-H}] \tag{80}$$

2. **Trimerization Process:**

   a) Ethylene adsorption on [Ni-butyl]

   $$[\text{Ni-butyl}] + \text{C}_2\text{H}_{4\,\mathbf{(gp)}} \rightarrow [\text{et-Ni-butyl}] \tag{81}$$

   b) C-C coupling (Trimerization)

   $$[\text{et-Ni-butyl}] \rightarrow [\text{Ni-hexyl}] \tag{82}$$

c) β-H elimination

$$[\text{Ni-hexyl}] \rightarrow [\text{hex-Ni-H}] \tag{83}$$

3. **Tetramerization Process:**

a) Ethylene adsorption on [Ni-hexyl]

$$[\text{Ni-hexyl}] + \text{C}_2\text{H}_{4\,\mathbf{(gp)}} \rightarrow [\text{et-Ni-hexyl}] \tag{84}$$

b) C-C coupling (Tetramerization)

$$[\text{et-Ni-hexyl}] \rightarrow [\text{Ni-octyl}] \tag{85}$$

c) β-H elimination

$$[\text{Ni-octyl}] \rightarrow [\text{octyl-Ni-H}] \tag{86}$$

4. **Desorption Process:**

a) Direct but-1-ene desorption

$$[\text{but-Ni-H}] \rightarrow [\text{Ni-H}] + \text{C}_4\text{H}_{8\,\mathbf{(gp)}} \tag{87}$$

b) Substitutive but-1-ene desorption (1)

$$[\text{but-Ni-H}] + \text{C}_2\text{H}_{4\,\mathbf{(gh)}} \rightarrow [\text{et-Ni-but-H}] \tag{88}$$

c) Substitutive but-1-ene desorption (2)

$$[\text{et-Ni-but-H}] \rightarrow [\text{et-Ni-H}] + \text{C}_4\text{H}_{8\,\mathbf{(gh)}} \tag{89}$$

d) Direct hex-1-ene desorption

$$[\text{hex-Ni-H}] \rightarrow [\text{Ni-H}] + \text{C}_6\text{H}_{12\,\mathbf{(g)}} \tag{90}$$

e) Substitutive hex-1-ene desorption (1)

$$[\text{hex-Ni-H}] + \text{C}_2\text{H}_{4\,\mathbf{(gh)}} \rightarrow [\text{et-Ni-hex-H}] \tag{91}$$

f) Substitutive hex-1-ene desorption (2)

$$[\text{et-Ni-hex-H}] \rightarrow [\text{et-Ni-H}] + \mathrm{C_6H_{12}}_{\mathbf{(gh)}} \tag{92}$$

g) Direct oct-1-ene desorption

$$[\text{oct-Ni-H}] \rightarrow [\text{Ni-H}] + \mathrm{C_8H_{16}}_{\mathbf{(gp)}} \tag{93}$$

h) Substitutive oct-1-ene desorption (1)

$$[\text{oct-Ni-H}] + \mathrm{C_2H_4}_{\mathbf{(gh)}} \rightarrow [\text{et-Ni-oct-H}] \tag{94}$$

i) Substitutive oct-1-ene desorption (2)

$$[\text{et-Ni-oct-H}] \rightarrow [\text{et-Ni-H}] + \mathrm{C_8H_{16}}_{\mathbf{(gh)}} \tag{95}$$

A.3 Appendix for Chapter 4, Section 4.2

Statistical Deviations for Acene Dimers per Ring

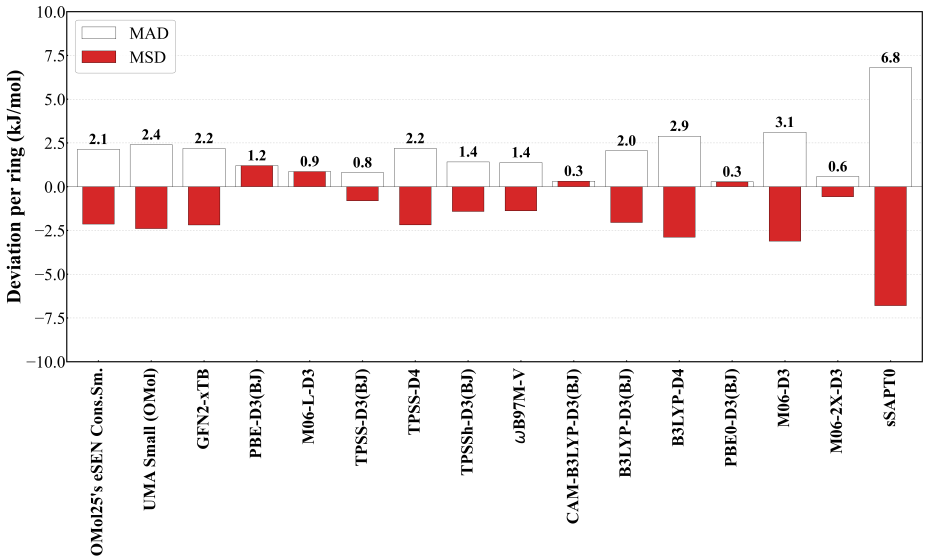


Figure 52: Mean absolute deviations (MAD, white boxes) and mean signed deviations per ring of the system (MSD, red boxes) of the interaction energies (in kJ/mol) for each tested method with respect to the PNO-LCCSD(T*)-F12 reference values, averaged over all studied complexes.

Natural Transition Orbital Analysis

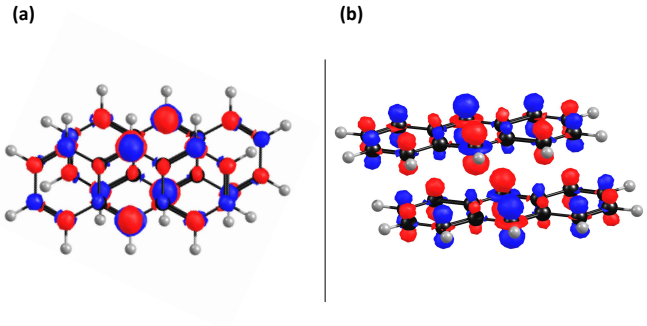


Figure 53: Natural transition orbital analysis at an iso-contour value of $\sigma = 0.0025\, a_0{}^{-3}$ for the glike-dimer of anthracene system: (a) top view, (b) side view. The red color is denoted to positive and blue to negative transition density regions. Method: TD-CAM-B3LYP/def2-TZVP.

A.4 Appendix for Chapter 4, Section 4.3

Interaction Energy and Buried Surface Area for Different Acene Dimers and Their Derivatives

Table 17: Interaction energy (in kJ/mol) and buried surface area (in a_0^2) for studied anthracene and its derivatives dimers.

System	Substitution	Dimer Type	Operation	E_{int}	BSA
Pure		X		-49.96	820.70
Pure		glike		-47.26	807.62
F		X		-58.82	1056.05
F		glike		-58.11	1039.70
BN	(1,5)	X	N	-43.68	818.74
BN	(1,5)	X	R	-49.78	863.32
BN	(1,5)	X	TR	-139.48	851.64
BN	(1,5)	glike	T	-113.37	876.84
BN	(9,10)	X	R	-57.36	815.02
BN	(9,10)	X	TR	-55.71	814.17
BN	(9,10)	glike	T	-50.01	823.33
BN	(1,4)	X	N	-53.38	826.74
BN	(1,4)	X	R	-63.33	811.65
BN	(1,4)	X	T	-58.21	827.37
BN	(1,4)	X	TR	-62.55	817.71
BN	(1,4)	glike	N	-43.07	825.53
BN	(1,4)	glike	T	-49.25	813.84
BN	(2,3)	X	N	-55.81	815.45
BN	(2,3)	X	T	-60.53	826.93
BN	(2,3)	X	TR	-63.17	826.13
BN	(2,3)	glike	R	-65.84	827.85
BN	(2,3)	glike	T	-62.28	820.11

A.5 Appendix for Chapter 4, Section 4.4

Distorted Geometries of Selected Acene Dimers

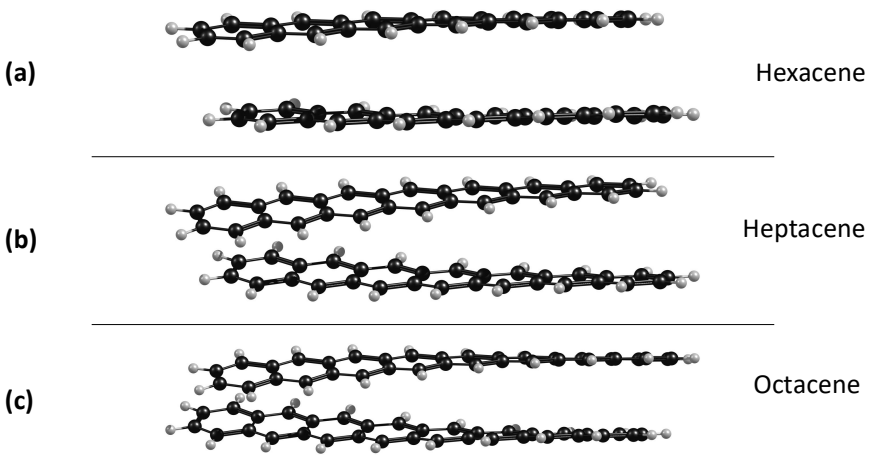


Figure 54: Side view of X-dimers for hexacene, heptacene, and octacene. As the number of fused rings increases, the planarity of the dimer is progressively distorted, with one monomer beginning to wrap around the other.

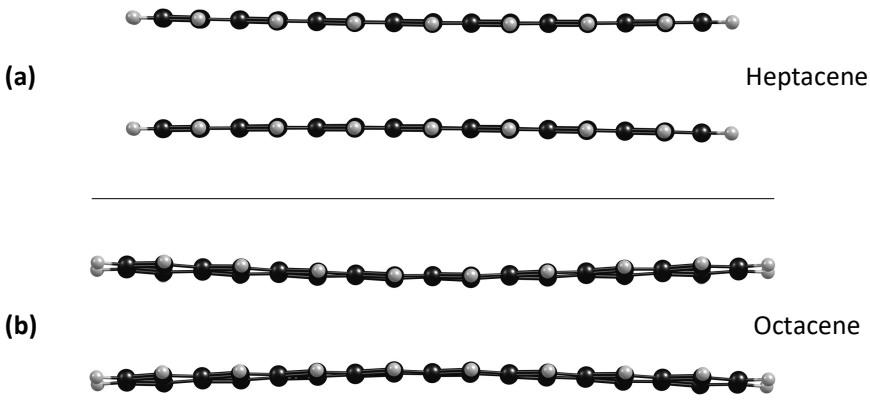


Figure 55: Side view of slip-dimers for heptacene and octacene. For the octacene dimer, a progressive distortion of planarity is observed, accompanied by central pinching of the two monomers toward each other.

A.6 Appendix for Chapter 4, Section 4.5

Machine Learning Study of Interaction Energy Prediction with Different Methods

Features of Dataset

Two categories of monomer features were considered: (i) structural features, including the number of aromatic rings N_{rings} and the monomer molecular surface area S_{mono}; and (ii) electronic features, including the HOMO–LUMO gap (E_{H-L}), first excitation energy E_{exc}, vertical ionization potential (vIP), vertical electron affinity (vEA), isotropic polarizability (α), and dipole moment (μ). The interaction energy E_{int} of dimer was chosen as a feature that can be predicted through monomer properties.

Three models

To systematically assess the relative contributions of structural and electronic descriptors to the prediction of E_{int}, three feature subsets were defined (Table 18). Model A serves as the full baseline. Models B and C are ablation studies that isolate the predictive power of structural against electronic features.

Table 18: Definition of the three models (A, B, C), their feature subsets and purpose. The subset with structural and electronic features denotes baseline.

Model	Feature subset	Purpose
A	Structural + Electronic (8 features)	All properties (baseline)
B	Structural only (2 features)	Evaluate role of electronic properties
C	Electronic only (6 features)	Evaluate role of structural properties

Regression Algorithms

Four regression algorithms spanning different method families were employed to ensure that conclusions are not algorithm-specific:

Ridge Regression

An L_2-regularized linear model that minimizes the sum of squared residuals plus a penalty term $\alpha||\omega||^2$ on the coefficient magnitudes. The regularization parameter was set to $\alpha = 1.0$. Ridge regression provides a linear baseline and is robust to

multicollinearity among features, which is relevant given the correlation between structural and electronic descriptors in the *cata*-condensed polybenzoid data.

Support Vector Regression (SVR)

A kernel-based method that constructs an ϵ-insensitive tube around the regression function, penalizing only predictions outside this margin. A radial basis function kernel was employed (C = 10, $\epsilon = 0.1$), enabling the capture of nonlinear structure–property relationships.

Random Forest (RF)

An ensemble of 200 decision trees, each trained on a bootstrapped subsample with random feature subsets. Predictions are averaged across trees, reducing variance. The maximum tree depth was restricted to 5 to mitigate overfitting on the limited dataset.

Gradient Boosting (XGBoost)

A sequential ensemble of 200 shallow decision trees (max depth = 3), where each successive tree corrects the residual errors of the preceding ensemble. A learning rate of 0.1 was used to prevent overfitting by limiting the contribution of each individual tree.

All ML methods were implemented in Python using scikit-learn (version $\geq$ 1.3).

Feature Preprocessing

All features were standardized to zero mean and unit variance using z-score normalization (StandardScaler). The scaler was fitted exclusively on the training set and applied to the test set to prevent information leakage. Standardization is critical for Ridge Regression and SVR which are sensitive to the scale of input features, and do not affect tree-based methods.

Validation Strategy

Given the limited dataset size (n = 58), a repeated random sub-sampling validation strategy was adopted. The data were split into training (60%) and test (40%) sets, with stratification by N_{rings} to ensure that all molecular sizes were represented in both subsets. Groups with fewer than three members ($N_{rings} \leq 3$) were merged into a single stratum for stratification purposes. This procedure was repeated 50 times with different random seeds, yielding distributions of performance metrics rather than single point estimates.

This approach offers several advantages over a single train/test split or standard k-fold cross-validation: (i) it provides robust estimates of model performance with associated

uncertainty (mean ± standard deviation); (ii) it ensures that conclusions are not artifacts of a particular data partition; (iii) the 40% test set (near 23 samples) is large enough to yield meaningful test-set metrics while retaining sufficient training data (near 35 samples) for model fitting.

Performance Metrics

Three complementary metrics were computed on the held-out test set for each of the 50 repetitions: The coefficient of determination, R^2, quantifies the proportion of variance in E_{int} explained by the model. A value of $R^2 = 1$ indicates perfect prediction; $R^2 = 0$ indicates that the model performs no better than predicting the mean value. Negative values indicate predictions worse than the mean.

The mean absolute error (MAE) provides an interpretable measure of average prediction error in the same units as the target variable (kJ/mol). The root mean square error (RMSE) places greater emphasis on large deviations, serving as a diagnostic for outlier sensitivity.

Feature Importance Analysis

Permutation importance was used to assess the contribution of individual features to model predictions. For each feature, its values in the test set were randomly shuffled while keeping all other features unchanged, and the resulting decrease in R^2 (ΔR^2) was recorded. This procedure was repeated 30 times per feature to obtain stable estimates with associated standard deviations. A large ΔR^2 indicates that the model relies heavily on that feature. A value near zero suggests redundancy or irrelevance.

Permutation importance was chosen over model-specific alternatives because it is model-agnostic, computed on the held-out test set (avoiding optimistic bias), and widely established in the literature. The analysis was performed on the median-performing repetition among the 50 splits to avoid selection bias.

Performance of ML Methods

Initially, four methods were trained on three feature subsets to identify the best-performing model in terms of R^2. The resulting values of R^2, MAE, and RMSE are summarized in Table 19. Ridge Regression consistently outperformed the other methods across all metrics for each task. For models A, B, and C, the R^2 values were 0.946, 0.949, and 0.945, respectively.

Table 19: Comparison of regression method performance (R^2, MAE, $RMSE$) across three models. R^2, MAE, $RMSE$ values are represented in mean and std species (MAE, $RMSE$ in kJ/mol).

Model	Method	R^2_{mean}	R^2_{std}	MAE_{mean}	MAE_{std}	$RMSE_{mean}$	$RMSE_{std}$
A	Ridge	0.946	0.019	3.555	0.545	4.849	0.822
A	SVR	0.533	0.152	7.831	1.510	14.661	3.804
A	RF	0.863	0.037	5.307	0.824	7.914	1.607
A	XGBoost	0.891	0.051	4.970	0.930	6.943	1.897
B	Ridge	0.949	0.023	3.299	0.560	4.673	0.980
B	SVR	0.693	0.112	5.984	1.149	11.884	3.327
B	RF	0.894	0.042	4.830	0.872	6.954	1.945
B	XGBoost	0.881	0.037	5.284	0.949	7.332	1.581
C	Ridge	0.945	0.019	3.769	0.583	4.936	0.749
C	SVR	0.475	0.162	9.091	1.439	15.537	3.848
C	RF	0.844	0.035	5.684	0.838	8.440	1.470
C	XGBoost	0.851	0.064	5.649	0.937	8.075	1.697

Examining the performance of each method separately on the test set over 50 repetitions for each feature subset reveals that Ridge Regression exhibits the lowest standard deviation across all reported metrics, including R^2 (see Figure 56). The second-best performer was XGBoost for subsets A and C, while Random Forest ranked second only for model B. The SVR method consistently showed the weakest performance, with R^2 below 0.7 across all feature subsets. The mean MAE and RMSE values for Ridge Regression did not exceed 4 and 5 kJ/mol, respectively.

It is worth emphasizing once more that for Ridge Regression all three feature subsets yielded approximately equal performance in terms of R^2, whereas for SVR the results varied considerably, with subsets A and B performing notably worse. Random Forest and XGBoost exhibited roughly comparable performance across all subsets (see Figure 56).

In the end, the best method was Ridge Regression. Surprisingly, Model B exhibits noticeably higher mean performance metrics than Models A and C, with the exception of the standard deviations (Table 19). This is noteworthy because, from the perspective of the ML method, the inclusion of electronic properties slightly worsens the mean values of R^2, MAE, and RMSE while simultaneously increasing their standard deviations.

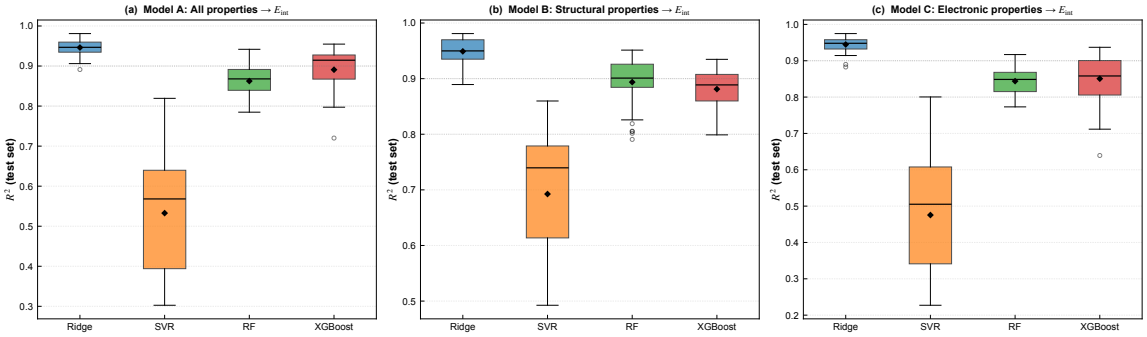


Figure 56: The performance of three different models (A, B, C) for four ML methods based on R^2 for test set. These values were calculate after 50 repeats. Each box represents the inter-quartile range (IQR), with the horizontal line denoting the median. Whiskers extend to 1.5 × IQR; values falling outside this range are displayed as individual data points (open circles), and the mean is indicated by a filled circle.

Probably, the unnecessary features introduce noise that cannot be compensated easily. Ultimately, it follows that E_{int} can be predicted using either structural descriptors, electronic properties, or a combination of both, with only minimal differences in performance.

Permutation Importance of Features

What underlies the comparable performance across three different feature subsets? This question was addressed through permutation importance analysis, presented in Figure 57. For the full monomer feature set (Model A), the most important descriptors were E_{exc}, monomer surface area, vEA, and polarizability. For Model B (structural features only), the dominant descriptor was monomer surface area (S_{mono}), while for Model C (electronic features only), polarizability and vEA were the leading contributors. Since the correlation matrix has shown that nearly all monomer properties, with the exception of the dipole moment, correlate with E_{int}, the absence of any single feature is readily compensated by the remaining ones. Notably, the inclusion of structural descriptors diminishes the contribution of polarizability in Model A, whereas in Model C polarizability becomes the dominant feature, with vEA playing only a minor role.

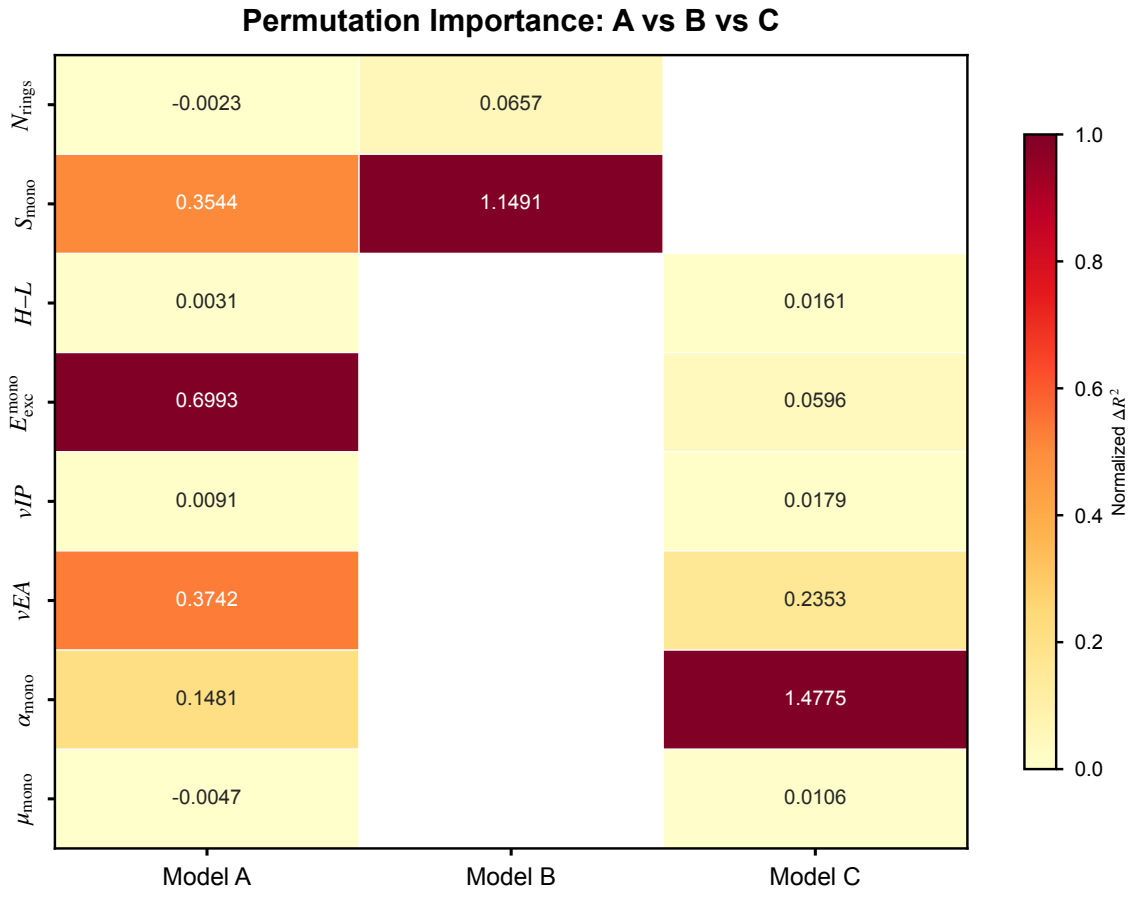


Figure 57: Permutation Importance heatmap (normalized ΔR^2) of three models for all considered monomeric features (from top to bottom): number of fused rings (N_{rings}), monomer surface area (S_{mono}), HOMO–LUMO gap (E_{H-L}), first excitation energy (E_{exc}), vertical ionization potential (vIP), vertical electron affinity (vEA), isotropic polarizability (α), dipole moment (μ).

Acronyms and Abbreviations

ADC(2) Algebraic-diagrammatic construction to second order

ASA Accessible surface area

BSA Buried surface area

BSE Bethe-Salpeter equation

BSIE Basis set incompleteness error

BSSE Basis set superposition error

CAM Coulomb-attenuating method

CASPT2 Complete active space with second-order perturbation theory

CASSCF Complete active space self-consistent field

CBS Complete basis set

CC Coupled cluster

CC2 Approximate coupled cluster singles and doubles

CCSD(T) Coupled cluster singles, doubles, and perturbative triples

CI Configuration interaction

CP Counterpoise correction

CT Charge transfer

D3(BJ) Grimme's dispersion correction with Becke-Johnson damping

DCM Dichloromethane

DFT Density functional theory

DLPNO Domain-based local pair natural orbital

DMRG Density matrix renormalization group

DNA Deoxyribonucleic acid

ev*GW* Eigenvalue-only self-consistent GW method

F12 Explicitly correlated F12 method

FOD Fractional occupation number weighted electron density

GGA Generalized gradient approximation

HF Hartree-Fock theory

HOMO Highest occupied molecular orbital

kMC Kinetic Monte Carlo

KS Kohn-Sham

LAO Localized atomic orbital

LDA Local density approximation

LE Local excitation

LUMO Lowest unoccupied molecular orbital

MAD Mean absolute deviation

MAE Mean absolute error

meta-GGA Meta-generalized gradient approximation

ML Machine learning

MOF Metal-organic framework

MP2 Møller-Plesset perturbation theory of second order

MSD Mean signed deviation

NCI Non-covalent interaction

NEB Nudged elastic band

Ni-NU-1000 Nickel-grafted NU-1000 metal-organic framework

NNP Neural network potential

NTO Natural transition orbital

PAH Polycyclic aromatic hydrocarbon

PAO Projected atomic orbital

PBC Periodic boundary conditions

PES Potential energy surface

PNO Pair natural orbital

RF Random forest

RI Resolution of identity (density fitting)

RMSE Root mean square error

SAC Single atom catalyst

SAPT Symmetry-adapted perturbation theory

SIE Self-interaction error

SVR Support vector regression

TD-DFT Time-dependent density functional theory

THF Tetrahydrofuran

TOF Turnover frequency

TS Transition state

TST Transition state theory

UEG Uniform electron gas

UV-Vis Ultraviolet-visible

vdW van der Waals

vEA Vertical electron affinity

vFG Vertical fundamental gap

vIP Vertical ionization potential

WFT Wave function theory

XC Exchange-correlation

XGBoost Extreme gradient boosting

Bibliography

[1] F. H. Stillinger, T. A. Weber, *Science* **1984**, *225*, 983–989.

[2] J. C. Schön, *The Journal of Chemical Physics* **2024**, *161*, 050901.

[3] D. J. Wales, T. V. Bogdan, *The Journal of Physical Chemistry B* **2006**, *110*, 20765–20776.

[4] C. P. Massen, J. P. K. Doye, *Physical Review E* **2005**, *71*.

[5] D. J. Wales, *Annual Review of Physical Chemistry* **2018**, *69*, 401–425.

[6] C. D. Maranas, C. A. Floudas, *Journal of Global Optimization* **1994**, *4*, 135–170.

[7] R. L. Rowley, C. M. Tracy, T. A. Pakkanen, *The Journal of Chemical Physics* **2006**, *125*.

[8] H. P. G. Thompson, G. M. Day, *Chemical Science* **2014**, *5*, 3173–3182.

[9] H. Frauenfelder, S. G. Sligar, P. G. Wolynes, *Science* **1991**, *254*, 1598–1603.

[10] B. Strodel, *Journal of Molecular Biology* **2021**, *433*, 167182.

[11] H. B. Schlegel, *Journal of Computational Chemistry* **2003**, *24*, 1514–1527.

[12] B. Ensing, M. De Vivo, Z. Liu, P. Moore, M. L. Klein, *Accounts of Chemical Research* **2005**, *39*, 73–81.

[13] K.-M. Weitzel, *ChemPhysChem* **2025**, *26*.

[14] D. G. Truhlar, B. C. Garrett, S. J. Klippenstein, *The Journal of Physical Chemistry* **1996**, *100*, 12771–12800.

[15] S. J. Benkovic, G. G. Hammes, S. Hammes-Schiffer, *Biochemistry* **2008**, *47*, 3317–3321.

[16] E. Tayyebi, K. S. Exner, *ACS Electrochemistry* **2025**, *1*, 910–920.

[17] R. Podeszwa, R. Bukowski, K. Szalewicz, *The Journal of Physical Chemistry A* **2006**, *110*, 10345–10354.

[18] Q. C. Nguyen, Y. S. Ong, H. Soh, J.-L. Kuo, *The Journal of Physical Chemistry A* **2008**, *112*, 6257–6261.

[19] E. Van Dornshuld, G. S. Tschumper, *Journal of Chemical Theory and Computation* **2016**, *12*, 1534–1541.

[20] Y. Chen, Y.-B. Wang, Y. Zhang, W. Wang, *Computational and Theoretical Chemistry* **2017**, *1112*, 52–60.

[21] K. Carter-Fenk, K. U. Lao, J. M. Herbert, *Accounts of Chemical Research* **2021**, *54*, 3679–3690.

[22] S. Kozuch, S. Shaik, *Accounts of Chemical Research* **2010**, *44*, 101–110.

[23] T. C. Lin, E. E. Bickel Rogers, A. Bhan, *ACS Catalysis* **2024**, *15*, 780–788.

[24] V. Van Speybroeck, K. De Wispelaere, J. Van der Mynsbrugge, M. Vandichel, K. Hemelsoet, M. Waroquier, *Chemical Society Reviews* **2014**, *43*, 7326–7357.

[25] R. Bhattacharjee, H. Lischka, M. Kertesz, *ACS Materials Au* **2025**, *5*, 365–376.

[26] M. Thelen, J. F. Henrichsmeyer, A. Buchwald, R. F. Fink, *Journal of the American Chemical Society* **2026**, *148*, 16151–16169.

[27] V. Bernales, A. B. League, Z. Li, N. M. Schweitzer, A. W. Peters, R. K. Carlson, J. T. Hupp, C. J. Cramer, O. K. Farha, L. Gagliardi, *The Journal of Physical Chemistry C* **2016**, *120*, 23576–23583.

[28] J. Ye, L. Gagliardi, C. J. Cramer, D. G. Truhlar, *Journal of Catalysis* **2017**, *354*, 278–286.

[29] J. Ye, L. Gagliardi, C. J. Cramer, D. G. Truhlar, *Journal of Catalysis* **2018**, *360*, 160–167.

[30] T. J. Goncalves, P. N. Plessow, F. Studt, *ChemCatChem* **2019**, *11*, 4368–4376.

[31] P. N. Plessow, F. Studt, *The Journal of Physical Chemistry Letters* **2020**, *11*, 4305–4310.

[32] E. M. Cabaleiro-Lago, J. Rodríguez-Otero, *ACS Omega* **2018**, *3*, 9348–9359.

[33] D. C. A. Valente, M. T. do Casal, M. Barbatti, T. A. Niehaus, A. J. A. Aquino, H. Lischka, T. M. Cardozo, *The Journal of Chemical Physics* **2021**, *154*, 044306.

[34] A. H. Alessa, *The Journal of Physical Chemistry A* **2025**, *129*, 10117–10133.

[35] J. Čížek, *The Journal of Chemical Physics* **1966**, *45*, 4256–4266.

[36] J. Čižek, J. Paldus, *International Journal of Quantum Chemistry* **1971**, *5*, 359–379.

[37] R. J. Bartlett, J. F. Stanton, *Reviews in Computational Chemistry* **1994**, 65–169.

[38] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, *Chemical Physics Letters* **1989**, *157*, 479–483.

[39] J. F. Stanton, *Chemical Physics Letters* **1997**, *281*, 130–134.

[40] P. Jurečka, J. Šponer, J. Černý, P. Hobza, *Physical Chemistry Chemical Physics* **2006**, *8*, 1985–1993.

[41] L. Goerigk, S. Grimme, *Journal of Chemical Theory and Computation* **2010**, *7*, 291–309.

[42] K. E. Riley, P. Hobza, *WIREs Computational Molecular Science* **2011**, *1*, 3–17.

[43] J. Řezáč, K. E. Riley, P. Hobza, *Journal of Chemical Theory and Computation* **2012**, *8*, 4285–4292.

[44] V. K. Prasad, Z. Pei, S. Edelmann, A. Otero-de-la-Roza, G. A. DiLabio, *Journal of Chemical Theory and Computation* **2021**, *18*, 151–166.

[45] J. Sauer, *Journal of Catalysis* **2024**, *433*, 115482.

[46] X.-F. Tan, B. Long, D.-S. Ren, W.-J. Zhang, Z.-W. Long, E. Mitchell, *Physical Chemistry Chemical Physics* **2018**, *20*, 7701–7709.

[47] S. Grimme, M. Steinmetz, M. Korth, *The Journal of Organic Chemistry* **2007**, *72*, 2118–2126.

[48] R. M. Richard, M. S. Marshall, O. Dolgounitcheva, J. V. Ortiz, J.-L. Brédas, N. Marom, C. D. Sherrill, *Journal of Chemical Theory and Computation* **2016**, *12*, 595–604.

[49] A. Karton, *WIREs Computational Molecular Science* **2016**, *6*, 292–310.

[50] Q. Ma, H.-J. Werner, *WIREs Computational Molecular Science* **2018**, *8*.

[51] P. Pulay, *Chemical Physics Letters* **1983**, *100*, 151–154.

[52] S. Saebø, P. Pulay, *Chemical Physics Letters* **1985**, *113*, 13–18.

[53] P. Pulay, S. Saebø, *Theor. Chim. Acta* **1986**, *69*, 357–368.

[54] S. Saebø, P. Pulay, *The Journal of Chemical Physics* **1987**, *86*, 914–922.

[55] S. Saebø, P. Pulay, *The Journal of Chemical Physics* **1988**, *88*, 1884–1890.

[56] S. Saebø, P. Pulay, *Annual Review of Physical Chemistry* **1993**, *44*, 213–236.

[57] Z. L. Seeger, E. I. Izgorodina, *Journal of Computational Chemistry* **2021**, *43*, 106–120.

[58] M. Schwilk, Q. Ma, C. Köppl, H.-J. Werner, *Journal of Chemical Theory and Computation* **2017**, *13*, 3650–3675.

[59] Q. Ma, M. Schwilk, C. Köppl, H.-J. Werner, *Journal of Chemical Theory and Computation* **2017**, *13*, 4871–4896.

[60] C. Riplinger, F. Neese, *The Journal of Chemical Physics* **2013**, *138*, 034106.

[61] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *The Journal of Chemical Physics* **2013**, *139*, 134101.

[62] Q. Ma, H.-J. Werner, *WIREs Computational Molecular Science* **2018**, *e1371*, 1–32.

[63] A. Hansen, P. J. Knowles, H.-J. Werner, *The Journal of Physical Chemistry A* **2025**, *129*, 4812–4833.

[64] A. Altun, C. Riplinger, F. Neese, G. Bistoni, *Journal of Chemical Theory and Computation* **2023**, *19*, 2039–2047.

[65] I. Sandler, J. Chen, M. Taylor, S. Sharma, J. Ho, *The Journal of Physical Chemistry A* **2021**, *125*, 1553–1563.

[66] F. Neese, *WIREs Computational Molecular Science* **2022**, *12*, e1606.

[67] S. Zhong, E. C. Barnes, G. A. Petersson, *The Journal of Chemical Physics* **2008**, *129*.

[68] F. Neese, E. F. Valeev, *Journal of Chemical Theory and Computation* **2010**, *7*, 33–43.

[69] T. Helgaker, W. Klopper, H. Koch, J. Noga, *The Journal of Chemical Physics* **1997**, *106*, 9639–9646.

[70] D. G. Liakos, F. Neese, *The Journal of Physical Chemistry A* **2012**, *116*, 4801–4816.

[71] F. Weigend, R. Ahlrichs, *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.

[72] K. Szalewicz, *WIREs Computational Molecular Science* **2011**, *2*, 254–272.

[73] M. Hapka, K. Patkowski in *Handbook of Electronic Structure Theory*, Elsevier, **2026**, pp. 245–262.

[74] B. Jeziorski, K. Szalewicz, G. Chałasiński, *International Journal of Quantum Chemistry* **1978**, *14*, 271–287.

[75] B. Jeziorski, R. Moszynski, K. Szalewicz, *Chemical Reviews* **1994**, *94*, 1887–1930.

[76] T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, C. D. Sherrill, *The Journal of Chemical Physics* **2014**, *140*, 094106.

[77] N. A. Matsokin, M. Kalinin, A. Buchwald, H.-J. Werner, R. F. Fink, K. Fink, D. I. Sharapa, *The Journal of Chemical Physics* **2026**, *164*, 134312.

[78] A. M. Teale, T. Helgaker, A. Savin, C. Adamo, B. Aradi, A. V. Arbuznikov, P. W. Ayers, E. J. Baerends, V. Barone, P. Calaminici, E. Cancès, E. A. Carter, P. K. Chattaraj, H. Chermette, I. Ciofini, T. D. Crawford, F. De Proft, J. F. Dobson, C. Draxl, T. Frauenheim, E. Fromager, P. Fuentealba, L. Gagliardi, G. Galli, J. Gao, P. Geerlings, N. Gidopoulos, P. M. W. Gill, P. Gori-Giorgi, A. Görling, T. Gould, S. Grimme, O. Gritsenko, H. J. A. Jensen, E. R. Johnson, R. O. Jones, M. Kaupp, A. M. Köster, L. Kronik, A. I. Krylov, S. Kvaal, A. Laestadius, M. Levy, M. Lewin, S. Liu, P.-F. Loos, N. T. Maitra, F. Neese, J. P. Perdew, K. Pernal, P. Pernot, P. Piecuch, E. Rebolini, L. Reining, P. Romaniello, A. Ruzsinszky, D. R. Salahub, M. Scheffler, P. Schwerdtfeger, V. N. Staroverov, J. Sun, E. Tellgren, D. J. Tozer, S. B. Trickey, C. A. Ullrich, A. Vela, G. Vignale, T. A. Wesolowski, X. Xu, W. Yang, *Physical Chemistry Chemical Physics* **2022**, *24*, 28700–28781.

[79] P. Hohenberg, W. Kohn, *Physical Review* **1964**, *136*, B864–B871.

[80] W. Kohn, L. J. Sham, *Physical Review* **1965**, *140*, A1133–A1138.

[81] M. Bursch, J.-M. Mewes, A. Hansen, S. Grimme, *Angewandte Chemie International Edition* **2022**, *134*, e202205735.

[82] J. P. Perdew, K. Schmidt in AIP Conference Proceedings, *Vol. 577*, AIP, **2001**, pp. 1–20.

[83] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, K. A. Lyssenko, *Science* **2017**, *355*, 49–52.

[84] S. H. Vosko, L. Wilk, M. Nusair, *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

[85] J. P. Perdew, Y. Wang, *Physical Review B* **1992**, *45*, 13244–13249.

[86] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys Rev Lett* **1996**, *77*, 3865–3868.

[87] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Physical Review Letters* **2003**, *91*, 146401.

[88] A. D. Becke, *The Journal of Chemical Physics* **1993**, *98*, 1372–1377.

[89] J. P. Perdew, J. Tao, V. N. Staroverov, G. E. Scuseria, *The Journal of Chemical Physics* **2004**, *120*, 6898–6911.

[90] F. Jensen, *Introduction to computational chemistry*, John Wiley & Sons, **2007**, p. 599.

[91] A. D. Becke, *The Journal of Chemical Physics* **2014**, *140*, 18A301.

[92] N. Mardirossian, M. Head-Gordon, *Molecular Physics* **2017**, *115*, 2315–2372.

[93] J. Sun, B. Xiao, Y. Fang, R. Haunschild, P. Hao, A. Ruzsinszky, G. I. Csonka, G. E. Scuseria, J. P. Perdew, *Physical Review Letters* **2013**, *111*, 106401.

[94] L. Goerigk, S. Grimme, *Physical Chemistry Chemical Physics* **2011**, *13*, 6670.

[95] Y. Zhao, N. González-García, D. G. Truhlar, *The Journal of Physical Chemistry A* **2005**, *109*, 2012–2018.

[96] D. Jacquemin, V. Wathelet, E. A. Perpète, C. Adamo, *Journal of Chemical Theory and Computation* **2009**, *5*, 2420–2435.

[97] S. Grimme, *WIREs Computational Molecular Science* **2011**, *1*, 211–228.

[98] J. Hermann, R. A. DiStasio, A. Tkatchenko, *Chemical Reviews* **2017**, *117*, 4714–4758.

[99] S. Grimme, *Journal of Computational Chemistry* **2006**, *27*, 1787–1799.

[100] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *The Journal of Chemical Physics* **2010**, *132*.

[101] B. M. Axilrod, E. Teller, *The Journal of Chemical Physics* **1943**, *11*, 299–300.

[102] Y. Muto in Proc. Phys. Math. Soc. Jpn, *Vol. 17*, **1943**, pp. 629–631.

[103] S. Grimme, S. Ehrlich, L. Goerigk, *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

[104] A. D. Becke, E. R. Johnson, *The Journal of Chemical Physics* **2005**, *123*, 154101.

[105] E. R. Johnson, A. D. Becke, *The Journal of Chemical Physics* **2005**, *123*, 024101.

[106] E. Caldeweyher, C. Bannwarth, S. Grimme, *The Journal of Chemical Physics* **2017**, *147*, 034112.

[107] A. Dreuw, M. Head-Gordon, *Chem. Rev.* **2005**, *105*, 4009–4037.

[108] C. Ullrich, *Time-Dependent Density-Functional Theory: Concepts and Applications*, Oxford University Press, **2012**.

[109] E. Runge, E. K. U. Gross, *Physical Review Letters* **1984**, *52*, 997–1000.

[110] M. E. CASIDA in *Recent Advances in Density Functional Methods*, WORLD SCIENTIFIC, **1995**, pp. 155–192.

[111] T.-N. Huynh, N. A. Matsokin, F. Studt, K. Fink, D. I. Sharapa, *The Journal of Physical Chemistry C* **2025**, *129*, 22377–22389.

[112] A. Avdoshin, N. A. Matsokin, T.-N. Huynh, D. I. Sharapa, K. Fink, F. Studt, W. Wenzel, M. Kozlowska, *npj Computational Materials* **2026**, *12*.

[113] A. Forestière, H. Olivier-Bourbigou, L. Saussine, *Oil Gas Science and Technology - Revue de l'IFP* **2009**, *64*, 649–667.

[114] W. Keim, *Angewandte Chemie International Edition* **2013**, *52*, 12492–12496.

[115] A. Alzamly, M. Bakiro, S. Hussein Ahmed, L. A. Siddig, H. L. Nguyen, *Coordination Chemistry Reviews* **2022**, *462*, 214522.

[116] C. P. Nicholas, *Applied Catalysis A: General* **2017**, *543*, 82–97.

[117] M. R. Gogate, *Petroleum Science and Technology* **2019**, *37*, 559–565.

[118] M. Yang, D. Fan, Y. Wei, P. Tian, Z. Liu, *Advanced Materials* **2019**, *31*.

[119] Y. Zhang, E. Xing, W. Han, P. Yang, S. Zhang, S. Liu, D. Cao, M. Li, *Engineering* **2024**, *43*, 99–114.

[120] M. Larsen, J. P. Tilsted, *Energy Research Social Science* **2024**, *118*, 103750.

[121] A. Finiels, F. Fajula, V. Hulea, *Catalysis Science Technology* **2014**, *4*, 2412–2426.

[122] R. Rajapaksha, P. Samanta, E. A. Quadrelli, J. Canivet, *Chemical Society Reviews* **2023**, *52*, 8059–8076.

[123] J. Canivet, S. Aguado, Y. Schuurman, D. Farrusseng, *Journal of the American Chemical Society* **2013**, *135*, 4195–4198.

[124] S. T. Madrahimov, J. R. Gallagher, G. Zhang, Z. Meinhart, S. J. Garibay, M. Delferro, J. T. Miller, O. K. Farha, J. T. Hupp, S. T. Nguyen, *ACS Catalysis* **2015**, *5*, 6713–6718.

[125] E. D. Metzger, C. K. Brozek, R. J. Comito, M. Dincă, *ACS Central Science* **2016**, *2*, 148–153.

[126] J. Lee, O. K. Farha, J. Roberts, K. A. Scheidt, S. T. Nguyen, J. T. Hupp, *Chemical Society Reviews* **2009**, *38*, 1450.

[127] A. Bavykina, N. Kolobov, I. S. Khan, J. A. Bau, A. Ramirez, J. Gascon, *Chemical Reviews* **2020**, *120*, 8468–8535.

[128] E. Pérez-Botella, S. Valencia, F. Rey, *Chemical Reviews* **2022**, *122*, 17647–17695.

[129] L. Zhang, X. Yang, J. Lin, X. Li, X. Liu, B. Qiao, A. Wang, T. Zhang, *Accounts of Chemical Research* **2025**, *58*, 1878–1892.

[130] N.-S. Mussa, K. Toshtay, F. Dumeignil, *Applied Catalysis A: General* **2025**, *708*, 120555.

[131] Z. Li, N. M. Schweitzer, A. B. League, V. Bernales, A. W. Peters, A. B. Getsoian, T. C. Wang, J. T. Miller, A. Vjunov, J. L. Fulton, J. A. Lercher, C. J. Cramer, L. Gagliardi, J. T. Hupp, O. K. Farha, *Journal of the American Chemical Society* **2016**, *138*, 1977–1982.

[132] B. Yeh, S. P. Vicchio, S. Chheda, J. Zheng, J. Schmid, L. Löbbert, R. Bermejo-Deval, O. Y. Gutiérrez, J. A. Lercher, C. C. Lu, M. Neurock, R. B. Getman, L. Gagliardi, A. Bhan, *Journal of the American Chemical Society* **2021**, *143*, 20274–20280.

[133] B. Yeh, S. Chheda, J. Zheng, J. Schmid, L. Löbbert, R. Bermejo-Deval, O. Y. Gutiérrez, J. A. Lercher, L. Gagliardi, A. Bhan, *Catalysis Science Technology* **2023**, *13*, 4213–4222.

[134] A. Avdoshin, W. Wenzel, M. Kozlowska, *Chemical Communications* **2025**.

[135] J. E. Mondloch, W. Bury, D. Fairen-Jimenez, S. Kwon, E. J. DeMarco, M. H. Weston, A. A. Sarjeant, S. T. Nguyen, P. C. Stair, R. Q. Snurr, O. K. Farha, J. T. Hupp, *Journal of the American Chemical Society* **2013**, *135*, 10294–10297.

[136] A. E. Platero-Prats, A. B. League, V. Bernales, J. Ye, L. C. Gallington, A. Vjunov, N. M. Schweitzer, Z. Li, J. Zheng, B. L. Mehdi, A. J. Stevens, A. Dohnalkova, M. Balasubramanian, O. K. Farha, J. T. Hupp, N. D. Browning, J. L. Fulton, D. M. Camaioni, J. A. Lercher, D. G. Truhlar, L. Gagliardi, C. J. Cramer, K. W. Chapman, *Journal of the American Chemical Society* **2017**, *139*, 10410–10418.

[137] T. A. Goetjen, X. Zhang, J. Liu, J. T. Hupp, O. K. Farha, *ACS Sustainable Chemistry & Engineering* **2019**, *7*, 2553–2557.

[138] T. A. Goetjen, M. S. Ferrandon, A. J. Kropf, J. V. Lamb, M. Delferro, J. T. Hupp, O. K. Farha, *The Journal of Physical Chemistry C* **2022**, *126*, 20388–20394.

[139] T. A. Goetjen, J. G. Knapp, Z. H. Syed, R. A. Hackler, X. Zhang, M. Delferro, J. T. Hupp, O. K. Farha, *Catalysis Science & Technology* **2022**, *12*, 1619–1627.

[140] J. Zheng, L. Löbbert, S. Chheda, N. Khetrapal, J. Schmid, C. A. Gaggioli, B. Yeh, R. Bermejo-Deval, R. K. Motkuri, M. Balasubramanian, J. L. Fulton, O. Y. Gutiérrez, J. I. Siepmann, M. Neurock, L. Gagliardi, J. A. Lercher, *Journal of Catalysis* **2022**, *413*, 176–183.

[141] L. Löbbert, S. Chheda, J. Zheng, N. Khetrapal, J. Schmid, R. Zhao, C. A. Gaggioli, D. M. Camaioni, R. Bermejo-Deval, O. Y. Gutiérrez, Y. Liu, J. I. Siepmann, M. Neurock, L. Gagliardi, J. A. Lercher, *Journal of the American Chemical Society* **2023**, *145*, 1407–1422.

[142] A. N. Migues, A. Muskat, S. M. Auerbach, W. Sherman, S. Vaitheeswaran, *ACS Catalysis* **2015**, *5*, 2859–2865.

[143] E. Mansoor, J. Van der Mynsbrugge, M. Head-Gordon, A. T. Bell, *Catalysis Today* **2018**, *312*, 51–65.

[144] S. Svelle, C. Tuma, X. Rozanska, T. Kerber, J. Sauer, *Journal of the American Chemical Society* **2008**, *131*, 816–825.

[145] N. Hansen, T. Kerber, J. Sauer, A. T. Bell, F. J. Keil, *Journal of the American Chemical Society* **2010**, *132*, 11525–11538.

[146] A. S. P. Gomes, C. R. Jacob, *Annual Reports on the Progress of Chemistry Section C: Physical Chemistry* **2012**, *108*, 222–277.

[147] P. N. Plessow, F. Studt, *ACS Catalysis* **2017**, *7*, 7987–7994.

[148] P. Huber, P. N. Plessow, *ChemPhysChem* **2025**, *26*, e202500147.

[149] J. Wellendorff, T. L. Silbaugh, D. Garcia-Pintos, J. K. Nørskov, T. Bligaard, F. Studt, C. T. Campbell, *Surface Science* **2015**, *640*, 36–44.

[150] T. van der Heide, B. Aradi, B. Hourahine, T. Frauenheim, T. A. Niehaus, *Physical Review Materials* **2023**, *7*, 063802.

[151] T. van der Heide, **2024**.

[152] V. Butera, *Physical Chemistry Chemical Physics* **2024**, *26*, 7950–7970.

[153] N. Dohrmann, D. S. King, C. A. Gaggioli, L. Gagliardi, *The Journal of Physical Chemistry C* **2023**, *127*, 16891–16900.

[154] R. J. Bartlett, M. Musiał, *Reviews of Modern Physics* **2007**, *79*, 291–352.

[155] R. J. Bartlett, *Physical Chemistry Chemical Physics* **2024**, *26*, 8013–8037.

[156] K. E. Weflen, M. R. Bentley, J. H. Thorpe, P. R. Franke, J. M. L. Martin, D. A. Matthews, J. F. Stanton, *The Journal of Physical Chemistry Letters* **2025**, *16*, 5121–5127.

[157] J. A. Pople, M. Head-Gordon, K. Raghavachari, *The Journal of Chemical Physics* **1987**, *87*, 5968–5975.

[158] J. F. Stanton, *Chemical Physics Letters* **1997**, *281*, 130–134.

[159] Z. Fang, M. Vasiliu, K. A. Peterson, D. A. Dixon, *Journal of Chemical Theory and Computation* **2017**, *13*, 1057–1066.

[160] B. Chan, J. Ho, *The Journal of Physical Chemistry A* **2023**, *127*, 10026–10031.

[161] L. Gyevi-Nagy, M. Kállay, P. R. Nagy, *Journal of Chemical Theory and Computation* **2021**, *17*, 860–878.

[162] D. G. Liakos, F. Neese, *Journal of Chemical Theory and Computation* **2015**, *11*, 4054–4063.

[163] Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo, F. Neese, *The Journal of Chemical Physics* **2018**, *148*, 011101.

[164] E. Paulechka, A. Kazakov, *The Journal of Physical Chemistry A* **2017**, *121*, 4379–4387.

[165] D. G. Liakos, Y. Guo, F. Neese, *The Journal of Physical Chemistry A* **2019**, *124*, 90–100.

[166] D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin, F. Neese, *Journal of Chemical Theory and Computation* **2015**, *11*, 1525–1539.

[167] S. Mallick, B. Roy, P. Kumar, *Computational and Theoretical Chemistry* **2020**, *1187*, 112934.

[168] M. Besora, F. Maseras, *WIREs Computational Molecular Science* **2018**, *8*, e1372.

[169] A. H. Motagamwala, M. R. Ball, J. A. Dumesic, *Annual Review of Chemical and Biomolecular Engineering* **2018**, *9*, 413–450.

[170] A. H. Motagamwala, J. A. Dumesic, *Chemical Reviews* **2020**, *121*, 1049–1076.

[171] D. R. Glowacki, C.-H. Liang, C. Morley, M. J. Pilling, S. H. Robertson, *The Journal of Physical Chemistry A* **2012**, *116*, 9545–9560.

[172] M. Liu, A. Grinberg Dana, M. S. Johnson, M. J. Goldman, A. Jocher, A. M. Payne, C. A. Grambow, K. Han, N. W. Yee, E. J. Mazeau, K. Blondal, R. H. West, C. F. Goldsmith, W. H. Green, *Journal of Chemical Information and Modeling* **2021**, *61*, 2686–2696.

[173] R. J. Shannon, E. Martínez-Núñez, D. V. Shalashilin, D. R. Glowacki, *Journal of Chemical Theory and Computation* **2021**, *17*, 4901–4912.

[174] M. S. Johnson, X. Dong, A. Grinberg Dana, Y. Chung, D. Farina, R. J. Gillis, M. Liu, N. W. Yee, K. Blondal, E. Mazeau, C. A. Grambow, A. M. Payne, K. A. Spiekermann, H.-W. Pang, C. F. Goldsmith, R. H. West, W. H. Green, *Journal of Chemical Information and Modeling* **2022**, *62*, 4906–4915.

[175] J. Zádor, C. Martí, R. Van de Vijver, S. L. Johansen, Y. Yang, H. A. Michelsen, H. N. Najm, *The Journal of Physical Chemistry A* **2023**, *127*, 565–588.

[176] Y. M. Liu, B. Smit, *ACS Catalysis* **2017**, *7*, 3940–3948.

[177] T.-H. T. Le, D. Ferro-Costas, A. Fernández-Ramos, M. A. Ortuño, *The Journal of Physical Chemistry C* **2024**, *128*, 1049–1057.

[178] M. Andersen, C. Panosetti, K. Reuter, *Frontiers in Chemistry* **2019**, *7*.

[179] N. Planas, J. E. Mondloch, S. Tussupbayev, J. Borycz, L. Gagliardi, J. T. Hupp, O. K. Farha, C. J. Cramer, *The Journal of Physical Chemistry Letters* **2014**, *5*, 3716–3723.

[180] G. Kresse, J. Hafner, *Physical Review B* **1993**, *47*, Publisher: American Physical Society, 558–561.

[181] G. Kresse, J. Hafner, *Physical Review B* **1994**, *49*, Publisher: American Physical Society, 14251–14269.

[182] G. Kresse, J. Furthmüller, *Computational Materials Science* **1996**, *6*, 15–50.

[183] G. Kresse, J. Furthmüller, *Physical Review B* **1996**, *54*, Publisher: American Physical Society, 11169–11186.

[184] V. Bernales, M. A. Ortuño, D. G. Truhlar, C. J. Cramer, L. Gagliardi, *ACS Central Science* **2017**, *4*, 5–19.

[185] S. Grimme, A. Hansen, *Angewandte Chemie International Edition* **2015**, *54*, 12308–12313.

[186] C. A. Bauer, A. Hansen, S. Grimme, *Chemistry – A European Journal* **2017**, *23*, 6150–6164.

[187] W. Jiang, N. J. DeYonker, A. K. Wilson, *Journal of Chemical Theory and Computation* **2012**, *8*, 460–468.

[188] D. A. Wappett, L. Goerigk, *Journal of Chemical Theory and Computation* **2023**, *19*, 8365–8383.

[189] Y. Zhao, D. G. Truhlar, *The Journal of Chemical Physics* **2006**, *125*, 194101.

[190] C. Adamo, V. Barone, *The Journal of Chemical Physics* **1999**, *110*, 6158–6170.

[191] A. D. Becke, *The Journal of Chemical Physics* **1992**, *96*, 2155–2160.

[192] N. Mardirossian, M. Head-Gordon, *The Journal of Chemical Physics* **2018**, *148*, 241736.

[193] Y. Zhao, D. G. Truhlar, *Theoretical Chemistry Accounts* **2007**, *120*, 215–241.

[194] Y. Zhao, F. Chen, *Journal of Multivariate Analysis* **2008**, *99*, 215–231.

[195] H. Eyring, *Chemical Reviews* **1935**, *17*, 65–77.

[196] H.-Z. Ye, T. C. Berkelbach, *Journal of Chemical Theory and Computation* **2024**, *20*, 8948–8959.

[197] Q. Sun, G. K.-L. Chan, *Accounts of Chemical Research* **2016**, *49*, 2705–2712.

[198] A. Shee, F. M. Faulstich, K. B. Whaley, L. Lin, M. Head-Gordon, *The Journal of Chemical Physics* **2024**, *161*, 164107.

[199] F. Maseras, A. Liedós, M. Duran, J. Bertrán, *Journal of the Chemical Society Faraday Transactions* **1992**, *88*, 1111–1117.

[200] H. Neugebauer, P. Pinski, S. Grimme, F. Neese, M. Bursch, *Journal of Chemical Theory and Computation* **2023**, *19*, 7695–7703.

[201] J. Pogrebetsky, A. Siklitskaya, A. Kubas, *Journal of Chemical Theory and Computation* **2023**, *19*, 4023–4032.

[202] K. Yang, J. Zheng, Y. Zhao, D. G. Truhlar, *The Journal of Chemical Physics* **2010**, *132*, 164117.

[203] J. Liu, J. Ye, Z. Li, K.-i. Otake, Y. Liao, A. W. Peters, H. Noh, D. G. Truhlar, L. Gagliardi, C. J. Cramer, O. K. Farha, J. T. Hupp, *Journal of the American Chemical Society* **2018**, *140*, 11174–11178.

[204] V. R. Cooper, C. N. Lam, Y. Wang, B. G. Sumpter in *Non-Covalent Interactions in Quantum Chemistry and Physics*, Elsevier, **2017**, pp. 417–451.

[205] Y. S. Al-Hamdani, A. Tkatchenko, *The Journal of Chemical Physics* **2019**, *150*, 010901.

[206] K. Müller-Dethlefs, P. Hobza, *Chemical Reviews* **1999**, *100*, 143–168.

[207] D. B. Rap, J. G. M. Schrauwen, B. Redlich, S. Brünken, *Journal of the American Chemical Society* **2024**, *146*, 23022–23033.

[208] A. Haque, K. M. Alenezi, M. S. Khan, W.-Y. Wong, P. R. Raithby, *Chemical Society Reviews* **2023**, *52*, 454–472.

[209] E. A. Meyer, R. K. Castellano, F. Diederich, *Angewandte Chemie International Edition* **2003**, *42*, 1210–1250.

[210] S. Jena, J. Dutta, K. D. Tulsiyan, A. K. Sahu, S. S. Choudhury, H. S. Biswal, *Chemical Society Reviews* **2022**, *51*, 4261–4286.

[211] K. Autumn, M. Sitti, Y. A. Liang, A. M. Peattie, W. R. Hansen, S. Sponberg, T. W. Kenny, R. Fearing, J. N. Israelachvili, R. J. Full, *Proceedings of the National Academy of Sciences* **2002**, *99*, 12252–12256.

[212] D. Scheeres, C. Hartzell, P. Sánchez, M. Swift, *Icarus* **2010**, *210*, 968–984.

[213] M. C. T. Fyfe, J. F. Stoddart, *Accounts of Chemical Research* **1997**, *30*, 393–401.

[214] J. Hoja, A. M. Reilly, A. Tkatchenko, *WIREs Computational Molecular Science* **2016**, *7*, e1294.

[215] K. S. Kim, P. Tarakeshwar, J. Y. Lee, *Chemical Reviews* **2000**, *100*, 4145–4186.

[216] *Non-covalent Interactions in the Synthesis and Design of New Compounds*, John Wiley Sons, Ltd, **2016**.

[217] M. Rapacioli, F. Calvo, F. Spiegelman, C. Joblin, D. J. Wales, *The Journal of Physical Chemistry A* **2005**, *109*, 2487–2497.

[218] K. Carter-Fenk, J. M. Herbert, *Chemical Science* **2020**, *11*, 6758–6765.

[219] P. Leyton, S. Sanchez-Cortes, J. V. Garcia-Ramos, C. Domingo, M. Campos-Vallette, C. Saitz, R. E. Clavijo, *The Journal of Physical Chemistry B* **2004**, *108*, 17484–17490.

[220] D. Zhang, T. K. Ronson, R. Lavendomme, J. R. Nitschke, *Journal of the American Chemical Society* **2019**, *141*, 18949–18953.

[221] C. A. Bayse, *Physical Chemistry Chemical Physics* **2023**, *25*, 20451–20461.

[222] W. Wang, Y. Zuo, Z. Xi, H. Yuan, M. He, J. Xie, *Computational and Theoretical Chemistry* **2025**, *1246*, 115127.

[223] G. M. Florio, T. L. Werblowsky, T. Müller, B. J. Berne, G. W. Flynn, *The Journal of Physical Chemistry B* **2005**, *109*, 4520–4532.

[224] L. Piot, A. Marchenko, J. Wu, K. Müllen, D. Fichou, *Journal of the American Chemical Society* **2005**, *127*, 16245–16250.

[225] B. B. Shrestha, S. Higashibayashi, H. Sakurai, *Beilstein Journal of Organic Chemistry* **2014**, *10*, 841–847.

[226] R. Rubert-Albiol, J. Cerdá, D. Aranda, M. Pérez-Escribano, P. Gómez, I. da Silva, S. Georgakopoulos, M. Más-Montoya, D. Curiel, E. Ortí, J. Aragó, *The Journal of Physical Chemistry C* **2025**, *129*, 5960–5972.

[227] M. S. Cubberley, B. L. Iverson, *Journal of the American Chemical Society* **2001**, *123*, 7560–7563.

[228] P. E. Fielding, R. C. Jarnagin, *The Journal of Chemical Physics* **1967**, *47*, 247–252.

[229] S.-H. Lim, T. G. Bjorklund, F. C. Spano, C. J. Bardeen, *Physical Review Letter* **2004**, *92*, 107402.

[230] J. Lee, *Applied Physics Letters* **2013**, *103*, 036101.

[231] X. Feng, X. Wang, C. Redshaw, B. Z. Tang, *Chemical Society Reviews* **2023**, *52*, 6715–6753.

[232] H. Wang, *Proceedings of the Combustion Institute* **2011**, *33*, 41–67.

[233] L. Moreira da Costa, S. R. Stoyanov, S. Gusarov, P. R. Seidl, J. Walkimar de M. Carneiro, A. Kovalenko, *The Journal of Physical Chemistry A* **2014**, *118*, 896–908.

[234] Y. Li, R. G. Clevenger, L. Jin, K. V. Kilway, Z. Peng, *The Journal of Physical Chemistry C* **2016**, *120*, 841–852.

[235] Q. Li, Y. Zhang, Z. Xie, Y. Zhen, W. Hu, H. Dong, *Journal of Materials Chemistry C* **2022**, *10*, 2411–2430.

[236] R. Meena, P. Pandey, C. Zuffa, P. Brázda, E. Samolova, N. McIntosh, M. Volpi, F. Modesti, C. Gatsios, N. Turetta, L. Catalano, W. Choi, S. Seki, J. Cornil, P. Erk, N. Koch, P. Samorì, L. Maini, G. Schweicher, Y. Geerts, *Crystal Growth Design* **2025**, *25*, 3087–3099.

[237] Q. Wang, E. J. Juarez-Perez, S. Jiang, M. Xiao, J. Qian, E.-Y. Shin, Y.-Y. Noh, Y. Qi, Y. Shi, Y. Li, *Physical Review Materials* **2020**, *4*, 044604.

[238] Y. Huang, J. Xing, Q. Gong, L.-C. Chen, G. Liu, C. Yao, Z. Wang, H.-L. Zhang, Z. Chen, Q. Zhang, *Nature Communications* **2019**, *10*.

[239] Monika, A. Verma, M. K. Tiwari, B. Show, S. Saha, *ACS Omega* **2019**, *5*, 448–459.

[240] K. Kanazawa, K. Bulgarevich, K. Kawabata, K. Takimiya, *Crystal Growth & Design* **2023**, *23*, 5941–5949.

[241] Z.-F. Yao, J.-Y. Wang, J. Pei, *Crystal Growth Design* **2017**, *18*, 7–15.

[242] X.-q. Sun, Y. Li, G.-y. Qin, P.-p. Lin, J.-x. Fan, H.-y. Jin, A.-M. Ren, J.-F. Guo, *Dyes and Pigments* **2024**, *228*, 112223.

[243] J. E. Campbell, J. Yang, G. M. Day, *Journal of Materials Chemistry C* **2017**, *5*, 7574–7584.

[244] Y. Tang, Y. Liu, W. Ning, L. Zhan, J. Ding, M. Yu, H. Liu, Y. Gao, G. Xie, C. Yang, *Journal of Materials Chemistry C* **2022**, *10*, 4637–4645.

[245] A. Azhdari Tehrani, H. Ghasempour, A. Morsali, G. Makhloufi, C. Janiak, *Crystal Growth Design* **2015**, *15*, 5543–5547.

[246] X.-L. Yang, X. Chen, G.-H. Hou, R.-F. Guan, R. Shao, M.-H. Xie, *Advanced Functional Materials* **2015**, *26*, 393–398.

[247] S. Mathis, S. Golafale, K. Solntsev, C. Ingram, *Crystals* **2018**, *8*, 53.

[248] F. P. Kinik, A. Ortega-Guerrero, D. Ongari, C. P. Ireland, B. Smit, *Chemical Society Reviews* **2021**, *50*, 3143–3177.

[249] A. M. Kulisiewicz, S. J. Garibay, G. R. Pozza, M. A. Browe, O. Sparr, S. Singh, L. A. Kelly, J. B. DeCoste, *ACS Applied Materials Interfaces* **2023**, *15*, 40727–40734.

[250] R. Haldar, M. Kozlowska, M. Ganschow, S. Ghosh, M. Jakoby, H. Chen, F. Ghalami, W. Xie, S. Heidrich, Y. Tsutsui, J. Freudenberg, S. Seki, I. A. Howard, B. S. Richards, U. H. F. Bunz, M. Elstner, W. Wenzel, C. Wöll, *Chemical Science* **2021**, *12*, 4477–4483.

[251] J. E. Anthony, *Angewandte Chemie International Edition* **2007**, *47*, 452–483.

[252] S. Armaković, S. J. Armaković, V. Holodkov, S. Pelemiš, *Materials Chemistry and Physics* **2016**, *170*, 210–217.

[253] L. A. Martínez-Martínez, M. Du, R. F. Ribeiro, S. Kéna-Cohen, J. Yuen-Zhou, *The Journal of Physical Chemistry Letters* **2018**, *9*, 1951–1957.

[254] J. I. Urgel, S. Mishra, H. Hayashi, J. Wilhelm, C. A. Pignedoli, M. Di Giovannantonio, R. Widmer, M. Yamashita, N. Hieda, P. Ruffieux, H. Yamada, R. Fasel, *Nature Communications* **2019**, *10*.

[255] H. Hayashi, H. Yamada, *Chemical Science* **2025**, *16*, 11204–11231.

[256] N. Zeitter, N. Hippchen, P. Baur, T. V. Unterreiner, F. Rominger, J. Freudenberg, U. H. F. Bunz, *Organic Materials* **2024**, *06*, 12–17.

[257] C. A. Bayse, *Physical Chemistry Chemical Physics* **2025**, *27*, 2958–2967.

[258] K. Alresheedi, A. Alajmi, A. Alrehaili, M. Alzanbaqi, A. Al-Jobory, C. Lambert, A. Ismael, *Carbon Trends* **2026**, *23*, 100630.

[259] M. Watanabe, K.-Y. Chen, Y. J. Chang, T. J. Chow, *Accounts of Chemical Research* **2013**, *46*, 1606–1615.

[260] B. Alam, A. F. Morrison, J. M. Herbert, *The Journal of Physical Chemistry C* **2020**, *124*, 24653–24666.

[261] H. Seiler, M. Krynski, D. Zahn, S. Hammer, Y. W. Windsor, T. Vasileiadis, J. Pflaum, R. Ernstorfer, M. Rossi, H. Schwoerer, *Science Advances* **2021**, *7*.

[262] F. Unger, D. Lepple, M. Asbach, L. Craciunescu, C. Zeiser, A. F. Kandolf, Z. Fišer, J. Hagara, J. Hagenlocher, S. Hiller, S. Haug, M. Deutsch, P. Grüninger, J. Novák, H. F. Bettinger, K. Broch, B. Engels, F. Schreiber, *The Journal of Physical Chemistry A* **2024**, *128*, 747–760.

[263] C. M. Combe, D. T. James, J. Wade, A. J. White, J.-S. Kim, I. McCulloch, *Tetrahedron Letters* **2013**, *54*, 6814–6818.

[264] K. A. McGarry, W. Xie, C. Sutton, C. Risko, Y. Wu, V. G. Young, J.-L. Brédas, C. D. Frisbie, C. J. Douglas, *Chemistry of Materials* **2013**, *25*, 2254–2263.

[265] S. M. Ryno, C. Risko, J.-L. Brédas, *Journal of the American Chemical Society* **2014**, *136*, 6421–6427.

[266] M. Mamada, H. Katagiri, T. Sakanoue, S. Tokito, *Crystal Growth & Design* **2015**, *15*, 442–448.

[267] X. Shi, W. Kueh, B. Zheng, K.-W. Huang, C. Chi, *Angewandte Chemie International Edition* **2015**, *54*, 14412–14416.

[268] C. Sutton, C. Risko, J. L. Brédas, *Chemistry of Materials* **2016**, *28*, 3–16.

[269] K. J. Thorley, T. W. Finn, K. Jarolimek, J. E. Anthony, C. Risko, *Chemistry of Materials* **2017**, *29*, 2502–2512.

[270] M. P. Lijina, A. Benny, R. Ramakrishnan, N. G. Nair, M. Hariharan, *Journal of the American Chemical Society* **2020**, *142*, 17393–17402.

[271] M. Moret, A. Gavezzotti, *New Journal of Chemistry* **2022**, *46*, 7626–7637.

[272] T. F. Headen, P. L. Cullen, R. Patel, A. Taylor, N. T. Skipper, *Physical Chemistry Chemical Physics* **2018**, *20*, 2704–2715.

[273] T. Janowski, P. Pulay, *Journal of the American Chemical Society* **2012**, *134*, 17520–17525.

[274] N. J. Silva, F. B. C. Machado, H. Lischka, A. J. A. Aquino, *Physical Chemistry Chemical Physics* **2016**, *18*, 22300–22310.

[275] V. Fishman, M. Lesiuk, J. M. L. Martin, A. Daniel Boese, *Journal of Chemical Theory and Computation* **2025**, *21*, 2311–2324.

[276] B. Schramm, M. Gray, J. M. Herbert, *Journal of the American Chemical Society* **2025**, *147*, 3243–3260.

[277] V. Fishman, J. M. L. Martin, A. D. Boese, Intermolecular Interactions of Large Systems: Boron Nitrides, Acenes, and Coronenes, **2026**.

[278] S. M. Ryno, C. Risko, J.-L. Brédas, *Chemistry of Materials* **2016**, *28*, 3990–4000.

[279] R. T. Saragi, C. Calabrese, M. Juanes, R. Pinacho, J. E. Rubio, C. Pérez, A. Lesarri, *The Journal of Physical Chemistry Letters* **2022**, *14*, 207–213.

[280] M. Saeki, H. Akagi, M. Fujii, *Journal of Chemical Theory and Computation* **2006**, *2*, 1176–1183.

[281] M. Rubeš, O. Bludský, P. Nachtigall, *ChemPhysChem* **2008**, *9*, 1702–1708.

[282] C. Schouder, A. S. Chatterley, F. Calvo, L. Christiansen, H. Stapelfeldt, *Structural Dynamics* **2019**, *6*, 044301.

[283] A. K. Lemmens, P. Chopra, D. Garg, A. L. Steber, M. Schnell, W. J. Buma, A. M. Rijs, *Molecular Physics* **2020**, *119*, e1811908.

[284] J. Hoche, M. Flock, X. Miao, L. N. Philipp, M. Wenzel, I. Fischer, R. Mitric, *Chemical Science* **2021**, *12*, 11965–11975.

[285] R. Mondal, C. Tönshoff, D. Khon, D. C. Neckers, H. F. Bettinger, *Journal of the American Chemical Society* **2009**, *131*, 14281–14289.

[286] G. Malloci, G. Cappellini, G. Mulas, A. Mattoni, *Chemical Physics* **2011**, *384*, 19–27.

[287] M. M. Muhammed, J. H. Mokkath, *New Journal of Chemistry* **2019**, *43*, 10774–10783.

[288] M. Kołaski, C. R. Arunkumar, K. S. Kim, *Journal of Chemical Theory and Computation* **2012**, *9*, 847–856.

[289] J.-i. Aihara, *The Journal of Physical Chemistry A* **1999**, *103*, 7487–7495.

[290] M. Solà, *Frontiers in Chemistry* **2013**, *1*.

[291] C. Tönshoff, H. F. Bettinger, *Chemistry – A European Journal* **2020**, *27*, 3193–3212.

[292] J. Poater, M. Duran, M. Solà, *Frontiers in Chemistry* **2018**, *6*.

[293] A. Wahab, L. Pfuderer, E. Paenurk, R. Gershoni-Poranne, *Journal of Chemical Information and Modeling* **2022**, *62*, 3704–3713.

[294] I. Sarfraz, A. Roglans, A. Artigas, M. Solà, *European Journal of Organic Chemistry* **2025**, *28*, e202500376.

[295] J. E. Anthony, *Chemical Reviews* **2006**, *106*, 5028–5048.

[296] T. Okamoto, Z. Bao, *Journal of the American Chemical Society* **2007**, *129*, 10308–10309.

[297] W. Fudickar, T. Linker, *Journal of the American Chemical Society* **2012**, *134*, 15071–15082.

[298] C. Chen, Y. Guo, Z. Chang, K. Müllen, X.-Y. Wang, *Nature Communications* **2024**, *15*.

[299] Y. Chen, L. Shen, X. Li, *The Journal of Physical Chemistry A* **2014**, *118*, 5700–5708.

[300] B. Kollmann, Z. Chen, D. Lüftner, O. Siri, P. Puschnig, *The Journal of Physical Chemistry C* **2018**, *122*, 6475–6482.

[301] Y. Appiarius, P. J. Gliese, S. A. W. Segler, P. Rusch, J. Zhang, P. J. Gates, R. Pal, L. A. Malaspina, K. Sugimoto, T. Neudecker, N. C. Bigall, S. Grabowsky, A. A. Bakulin, A. Staubitz, *The Journal of Physical Chemistry C* **2022**, *126*, 4563–4576.

[302] A. Ningthoujam, S. A. Shimray, D. K. S. Khaidem, F. A. Chipem, *Journal of Molecular Structure* **2023**, *1282*, 135198.

[303] S. E. Wheeler, J. W. G. Bloom, *The Journal of Physical Chemistry A* **2014**, *118*, 6133–6147.

[304] D. Bischof, M. W. Tripp, P. E. Hofmann, C.-H. Ip, S. I. Ivlev, M. Gerhard, U. Koert, G. Witte, *Chemistry – A European Journal* **2022**, *28*, e202103653.

[305] R. Amemiya, M. Yamaguchi, *Organic Biomolecular Chemistry* **2008**, *6*, 26–35.

[306] S. R. Domingos, D. S. Tikhonov, A. L. Steber, P. Eschenbach, S. Gruet, H. R. Hrodmarsson, K. Martin, G. A. Garcia, L. Nahon, J. Neugebauer, N. Avarvari, M. Schnell, *Nature Communications* **2024**, *15*.

[307] F. Neese, *WIREs Computational Molecular Science* **2012**, *2*, 73–78.

[308] S. F. Boys, F. Bernardi, *Molecular Physics* **1970**, *19*, 553–566.

[309] L. Turi, J. J. Dannenberg, *The Journal of Physical Chemistry* **1993**, *97*, 2488–2490.

[310] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, T. D. Crawford, *WIREs Computational Molecular Science* **2011**, *2*, 556–565.

[311] D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, C. D. Sherrill, *The Journal of Chemical Physics* **2020**, *152*.

[312] Q. Ma, H.-J. Werner, *Journal of Chemical Theory and Computation* **2018**, *14*, 198–215.

[313] Q. Ma, H.-J. Werner, *Journal of Chemical Theory and Computation* **2020**, *16*, 3135–3151.

[314] Q. Ma, H.-J. Werner, *Journal of Chemical Theory and Computation* **2021**, *17*, 902–926.

[315] H.-J. Werner, A. Hansen, *Journal of Chemical Theory and Computation* **2023**, *19*, 7007–7030.

[316] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *WIREs Computational Molecular Science* **2012**, *2*, 242–253.

[317] H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, I. Miller, Thomas F., A. Mitrushchenkov, K. A. Peterson, I. Polyak, G. Rauhut, M. Sibaev, *The Journal of Chemical Physics* **2020**, *152*, 144107.

[318] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, D. G. Truhlar, *Journal of Chemical Theory and Computation* **2011**, *7*, 3027–3034.

[319] T. Yanai, D. P. Tew, N. C. Handy, *Chemical Physics Letters* **2004**, *393*, 51–57.

[320] N. Mardirossian, M. Head-Gordon, *The Journal of Chemical Physics* **2016**, *144*, 214110.

[321] C. Bannwarth, S. Ehlert, S. Grimme, *Journal of Chemical Theory and Computation* **2019**, *15*, 1652–1671.

[322] F. Neese, *WIREs Computational Molecular Science* **2025**, *15*.

[323] D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csányi, M. Dzamba, P. Eastman, N. C. Frey, X. Fu, V. Gharakhanyan, A. S. Krishnapriyan, J. A. Rackers, S. Raja, A. Rizvi, A. S. Rosen, Z. Ulissi, S. Vargas, C. L. Zitnick, S. M. Blau, B. M. Wood, **2025**.

[324] X. Fu, B. M. Wood, L. Barroso-Luque, D. S. Levine, M. Gao, M. Dzamba, C. L. Zitnick, **2025**.

[325] S. VanZanten, C. Wagen, *ChemRxiv* **2025**.

[326] B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, L. Barroso-Luque, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, K. Michel, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, Z. W. Ulissi, C. L. Zitnick, **2025**.

[327] L. Hedin, *Physical Review* **1965**, *139*, A796–A823.

[328] W. Hanke, L. J. Sham, *Physical Review Letter* **1979**, *43*, 387–390.

[329] G. Strinati, H. J. Mattausch, W. Hanke, *Physical Review B* **1982**, *25*, 2867–2888.

[330] M. Rohlfing, S. G. Louie, *Physical Review Letter* **1998**, *80*, 3320–3323.

[331] S. Albrecht, L. Reining, R. D. Sole, G. Onida, *Physical Review Letter* **1998**, *80*, 4510–4513.

[332] G. Onida, L. Reining, A. Rubio, *Reviews of Modern Physics* **2002**, *74*, 601–659.

[333] A. Dreuw, M. Wormit, *WIREs Computational Molecular Science* **2014**, *5*, 82–95.

[334] O. Christiansen, H. Koch, P. Jørgensen, *Chemical Physics Letters* **1995**, *243*, 409–418.

[335] C. Hättig, F. Weigend, *The Journal of Chemical Physics* **2000**, *113*, 5154–5161.

[336] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński, J. M. Yu, *The Journal of Chemical Physics* **2020**, *152*.

[337] P. Pracht, S. Grimme, C. Bannwarth, F. Bohle, S. Ehlert, G. Feldmann, J. Gorges, M. Müller, T. Neudecker, C. Plett, S. Spicher, P. Steinbach, P. A. Wesołowski, F. Zeller, *The Journal of Chemical Physics* **2024**, *160*, 114110.

[338] K. Carter-Fenk, J. M. Herbert, *Physical Chemistry Chemical Physics* **2020**, *22*, 24870–24886.

[339] J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan, W. Yang, *Journal of Chemical Theory and Computation* **2011**, *7*, 625–632.

[340] D. Chakravarty, M. Guharoy, C. H. Robert, P. Chakrabarti, J. Janin, *Protein Science* **2013**, *22*, 1453–1457.

[341] J. L. Pascual-Ahuir, E. Silla, *Journal of Computational Chemistry* **1990**, *11*, 1047–1060.

[342] E. Silla, I. Tuñón, J. L. Pascual-Ahuir, *Journal of Computational Chemistry* **1991**, *12*, 1077–1088.

[343] J. L. Pascual-ahuir, E. Silla, I. Tuñon, *Journal of Computational Chemistry* **1994**, *15*, 1127–1138.

[344] A. Elmi, K. M. Bąk, S. L. Cockroft, *Angewandte Chemie International Edition* **2024**, *63*, e202412056.

[345] G. Regni, L. Baldinelli, G. Bistoni, *ACS Central Science* **2025**, *11*, 890–898.

[346] A. Li, H. S. Muddana, M. K. Gilson, *Journal of Chemical Theory and Computation* **2014**, *10*, 1563–1575.

[347] L. Goerigk, H. Kruse, S. Grimme, *ChemPhysChem* **2011**, *12*, 3421–3433.

[348] K. M. Herman, E. Aprà, S. S. Xantheas, *Physical Chemistry Chemical Physics* **2023**, *25*, 4824–4838.

[349] O. A. Vydrov, T. Van Voorhis, *The Journal of Chemical Physics* **2010**, *133*, 244103.

[350] X.-X. Fu, J.-F. Li, R.-Q. Zhang, Strong Orbital Interaction in pi-pi Stacking System, **2016**.

[351] P. I. Djurovich, E. I. Mayo, S. R. Forrest, M. E. Thompson, *Organic Electronics* **2009**, *10*, 515–520.

[352] D. Chen, H. Wang, *The Journal of Physical Chemistry C* **2019**, *123*, 27785–27793.

[353] A. Musa, M. Saeed, A. Shaari, R. Sahnoun, M. Lawal, *Molecular Physics* **2015**, *113*, 1347–1358.

[354] Y. Yang, E. R. Davidson, W. Yang, *Proceedings of the National Academy of Sciences* **2016**, *113*, E5098–E5107.

[355] J. Řezáč, A. de la Lande, *ChemPhysChem* **2023**, *24*, e202300329.

[356] X. Blase, I. Duchemin, D. Jacquemin, *Chemical Society Reviews* **2018**, *47*, 1022–1043.

[357] A. Prlj, B. F. E. Curchod, A. Fabrizio, L. Floryan, C. Corminboeuf, *The Journal of Physical Chemistry Letters* **2014**, *6*, 13–21.

[358] E. Brémond, M. Savarese, C. Adamo, D. Jacquemin, *Journal of Chemical Theory and Computation* **2018**, *14*, 3715–3727.

[359] R. M. Richard, J. M. Herbert, *Journal of Chemical Theory and Computation* **2011**, *7*, 1296–1306.

[360] A. D. Laurent, D. Jacquemin, *International Journal of Quantum Chemistry* **2013**, *113*, 2019–2039.

[361] A. G. Martynov, J. Mack, A. K. May, T. Nyokong, Y. G. Gorbunova, A. Y. Tsivadze, *ACS Omega* **2019**, *4*, 7265–7284.

[362] M. Hédouin, E. Luppi, O. Ward, D. Harrowven, C. Fressigné, I. Chataigner, *ChemistrySelect* **2023**, *8*, e202301943.

[363] P. P. Fehér, Á. Madarász, A. Stirling, *Chemistry–Methods* **2023**, *3*, e202200069.

[364] D. Jacquemin, I. Duchemin, X. Blase, *The Journal of Physical Chemistry Letters* **2017**, *8*, 1524–1529.

[365] N. O. C. Winter, N. K. Graf, S. Leutwyler, C. Hättig, *Physical Chemistry Chemical Physics* **2013**, *15*, 6623–6630.

[366] C. Suellen, R. G. Freitas, P.-F. Loos, D. Jacquemin, *Journal of Chemical Theory and Computation* **2019**, *15*, 4581–4590.

[367] R. Huebner, S. Meilczarek, C. Kuyatt, *Chemical Physics Letters* **1972**, *16*, 464–469.

[368] S. Zilberg, Y. Haas, S. Shaik, *The Journal of Physical Chemistry* **1995**, *99*, 16558–16565.

[369] A. Amirav, U. Even, J. Jortner, *The Journal of Chemical Physics* **1981**, *75*, 3770–3793.

[370] A. Klamt, G. Schüürmann, *Journal of the Chemical Society Perkin Transactions 2* **1993**, 799–805.

[371] B. M. Tissue, *Characterization of Materials* **2012**, 1–13.

[372] A. G. L. Olive, A. Guerzo, C. Belin, J. Reichwagen, H. Hopf, J.-P. Desvergne, *Research on Chemical Intermediates* **2008**, *34*, 137–145.

[373] P. B. Coto, S. Sharifzadeh, J. B. Neaton, M. Thoss, *Journal of Chemical Theory and Computation* **2014**, *11*, 147–156.

[374] J. Zhou, S. Jia, X. Xue, A. D. Skitnevskaya, E. Wang, X. Wang, X. Hao, Q. Zeng, A. I. Kuleff, A. Dorn, X. Ren, *The Journal of Physical Chemistry Letters* **2024**, *15*, 1529–1538.

[375] X. López, T. P. Straatsma, A. Sánchez-Mansilla, C. de Graaf, *The Journal of Physical Chemistry C* **2023**, *127*, 16249–16258.

[376] C. Sousa, X. López, X. Dong, R. Broer, T. P. Straatsma, C. de Graaf, *The Journal of Physical Chemistry C* **2025**, *129*, 4290–4302.

[377] K. U. Do, A. V. Conner, S. E. Wheeler, *Journal of the American Chemical Society* **2025**, *147*, 32273–32286.

[378] S. E. Wheeler, *Journal of the American Chemical Society* **2011**, *133*, 10262–10274.

[379] S. E. Wheeler, K. N. Houk, *Journal of the American Chemical Society* **2008**, *130*, 10854–10855.

[380] X. Shao, A. J. A. Aquino, M. Otyepka, D. Nachtigallová, H. Lischka, *Physical Chemistry Chemical Physics* **2020**, *22*, 22003–22015.

[381] E. G. Hohenstein, C. D. Sherrill, *The Journal of Physical Chemistry A* **2009**, *113*, 878–886.

[382] R. G. Huber, M. A. Margreiter, J. E. Fuchs, S. von Grafenstein, C. S. Tautermann, K. R. Liedl, T. Fox, *Journal of Chemical Information and Modeling* **2014**, *54*, 1371–1379.

[383] O. Hübner, J. Thusek, H.-J. Himmel, *Angewandte Chemie International Edition* **2023**, *62*, e202218042.

[384] S. Tsuzuki, M. Mikami, S. Yamada, *Journal of the American Chemical Society* **2007**, *129*, 8656–8662.

[385] Y. Sakamoto, T. Suzuki, M. Kobayashi, Y. Gao, Y. Fukai, Y. Inoue, F. Sato, S. Tokito, *Journal of the American Chemical Society* **2004**, *126*, 8138–8140.

[386] Y. Sakamoto, T. Suzuki, M. Kobayashi, Y. Gao, Y. Inoue, S. Tokito, *Molecular Crystals and Liquid Crystals* **2006**, *444*, 225–232.

[387] D. Bischof, M. Zeplichal, S. Anhäuser, A. Kumar, M. Kind, F. Kramer, M. Bolte, S. I. Ivlev, A. Terfort, G. Witte, *The Journal of Physical Chemistry C* **2021**, *125*, 19000–19012.

[388] Q. Cheng, P. Xing, *Nanoscale* **2025**, *17*, 16056–16070.

[389] J. Burdon, A. C. Childs, I. W. Parsons, J. C. Tatlow, *Journal of the Chemical Society Chemical Communications* **1982**, 534.

[390] Z. Zhang, T. Wächter, M. Kind, S. Schuster, J. W. Bats, A. Nefedov, M. Zharnikov, A. Terfort, *ACS Applied Materials Interfaces* **2016**, *8*, 7308–7319.

[391] A. Friedrich, L. Schraut-May, F. Rauch, P. Durand, J. Krebs, P. N. Ruth, S. Hammer, R. Bertermann, M. Finze, S. J. Clark, J. Pflaum, N. Leclerc, T. B. Marder, *Organic Chemistry Frontiers* **2025**, *12*, 736–753.

[392] X.-H. Ju, X. Liao, *International Journal of Materials Research* **2013**, *104*, 109–113.

[393] Y. Xie, H. F. Schaefer III, F. A. Cotton, *Chemical Communications* **2002**, 102–103.

[394] X. Zhang, Q.-S. Li, Y. Xie, H. F. Schaefer, III, *Molecular Physics* **2007**, *105*, 2743–2752.

[395] X. Feng, Q. Li, J. Gu, F. A. Cotton, Y. Xie, H. F. Schaefer, 3rd, *The Journal of Physical Chemistry A* **2009**, *113*, 887–894.

[396] R. Cardia, G. Malloci, A. Bosin, G. Serra, G. Cappellini, *Chem. Phys.* **2016**, *478*, 8–13.

[397] A. Karton, K. P. Gregory, B. Chan, *Chem. Phys.* **2025**, *595*, 112712.

[398] M. Zhang, Z.-M. Su, Y.-G. Li, H.-L. Xu, *Dyes and Pigments* **2024**, *231*, 112414.

[399] L. Palomino-Ruiz, S. Rodríguez-González, J. G. Fallaque, I. R. Márquez, N. Agraït, C. Díaz, E. Leary, J. M. Cuerva, A. G. Campaña, F. Martín, A. Millán, M. T. González, *Angewandte Chemie International Edition* **2021**, *60*, 6609–6616.

[400] A. N. Lamm, E. B. Garner, D. A. Dixon, S.-Y. Liu, *Angewandte Chemie International Edition* **2011**, *50*, 8157–8160.

[401] T. Ozaki, S. Diamandis, N. Rybansky, X. Yang, F. Bai, B. Li, J. Huang, S.-Y. Liu, *Journal of the American Chemical Society* **2026**, *148*, 3820–3829.

[402] J. Bae, M. Sakai, Y. Tsuchiya, N. Ando, X.-K. Chen, T. B. Nguyen, C.-Y. Chan, Y.-T. Lee, M. Auffray, H. Nakanotani, S. Yamaguchi, C. Adachi, *Frontiers in Chemistry* **2022**, *10*.

[403] C. Yang, X. Luo, *ACS Omega* **2025**, *10*, 4143–4153.

[404] P. Yang, H. Nomura, M. Nagata, N. Takano, E. Watanabe, H. Fukumoto, M. Morita, T. Yasuda, T. Agou, *Asian Journal of Organic Chemistry* **2025**, *14*, e00296.

[405] H. L. van de Wouw, J. Y. Lee, M. A. Siegler, R. S. Klausen, *Organic Biomolecular Chemistry* **2016**, *14*, 3256–3263.

[406] J. S. A. Ishibashi, C. Darrigan, A. Chrostowska, B. Li, S.-Y. Liu, *Dalton Transactions* **2019**, *48*, 2807–2812.

[407] L. Jia, Q. Wu, T. Yang, B. Xie, J. Sheng, W. Xie, J. Shi, *Molecules* **2025**, *30*, 4252.

[408] M. Pinheiro, F. B. C. Machado, F. Plasser, A. J. A. Aquino, H. Lischka, *Journal of Materials Chemistry C* **2020**, *8*, 7793–7804.

[409] J. V. M. Pimentel, J. C. V. Chagas, M. Pinheiro, A. J. A. Aquino, H. Lischka, F. B. C. Machado, *The Journal of Physical Chemistry A* **2025**, *129*, 470–480.

[410] Y. Rahav, O. Gidron, *The Journal of Organic Chemistry* **2025**, *90*, 18148–18154.

[411] M. Pinheiro, L. F. A. Ferrão, F. Bettanin, A. J. A. Aquino, F. B. C. Machado, H. Lischka, *Physical Chemistry Chemical Physics* **2017**, *19*, 19225–19233.

[412] M. Probert, *Contemporary Physics* **2017**, *58*, 365–365.

[413] T. Gould, T. Bučko, *Journal of Chemical Theory and Computation* **2016**, *12*, 3603–3613.

[414] M. Schwilk, P. D. Mezei, D. N. Tahchieva, O. A. von Lilienfeld, *Electronic Structure* **2022**, *4*, 014005.

[415] D.-e. Jiang, S. Dai, *The Journal of Physical Chemistry A* **2007**, *112*, 332–335.

[416] S. Rayne, K. Forest, *Computational and Theoretical Chemistry* **2011**, *976*, 105–112.

[417] B. Hajgató, M. Huzak, M. S. Deleuze, *The Journal of Physical Chemistry A* **2011**, *115*, 9282–9293.

[418] J. R. Carvalho, R. Nieman, M. Kertesz, A. J. A. Aquino, A. Hansen, H. Lischka, *Theoretical Chemistry Accounts* **2024**, *143*.

[419] R. Dorel, A. M. Echavarren, *European Journal of Organic Chemistry* **2016**, *2017*, 14–24.

[420] Z. Ruan, J. Schramm, J. B. Bauer, T. Naumann, L. V. Müller, F. Sättele, H. F. Bettinger, R. Tonner-Zech, J. M. Gottfried, *Journal of the American Chemical Society* **2025**, *147*, 4862–4870.

[421] R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *Journal of Chemical Theory and Computation* **2015**, *11*, 2087–2096.

[422] I. V. Tetko, O. Engkvist, U. Koch, J.-L. Reymond, H. Chen, *Molecular Informatics* **2016**, *35*, 615–621.

[423] C. Draxl, M. Scheffler, *MRS Bulletin* **2018**, *43*, 676–682.

[424] J. F. Rodrigues, Jr, L. Florea, M. C. F. de Oliveira, D. Diamond, O. N. Oliveira, Jr, *Discovery Materials* **2021**, *1*, 12.

[425] S. Tretiakov, A. Nigam, R. Pollice, *Chemical Reviews* **2025**, *125*, 5776–5829.

[426] Y. Zhao, D. G. Truhlar, *Journal of Chemical Theory and Computation* **2005**, *1*, 415–432.

[427] L. F. Molnar, X. He, B. Wang, K. M. Merz, *The Journal of Chemical Physics* **2009**, *131*, 065102.

[428] M. Chourasia, G. M. Sastry, G. N. Sastry, *International Journal of Biological Macromolecules* **2011**, *48*, 540–552.

[429] J. Řezáč, K. E. Riley, P. Hobza, *Journal of Chemical Theory and Computation* **2011**, *7*, 2427–2438.

[430] S. Grimme, *Chemistry – A European Journal* **2012**, *18*, 9955–9964.

[431] R. Sedlak, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay, P. Hobza, *Journal of Chemical Theory and Computation* **2013**, *9*, 3364–3374.

[432] F. Ballesteros, S. Dunivan, K. U. Lao, *The Journal of Chemical Physics* **2021**, *154*, 154104.

[433] J. Řezáč, M. Dubecký, P. Jurečka, P. Hobza, *Physical Chemistry Chemical Physics* **2015**, *17*, 19268–19277.

[434] Z. M. Sparrow, B. G. Ernst, P. T. Joo, K. U. Lao, R. A. DiStasio, *The Journal of Chemical Physics* **2021**, *155*, 184303.

[435] J. Řezáč, *Physical Chemistry Chemical Physics* **2022**, *24*, 14780–14793.

[436] J. D. Herdman, J. H. Miller, *The Journal of Physical Chemistry A* **2008**, *112*, 6249–6256.

[437] R. Ruiterkamp, T. Halasinski, F. Salama, B. H. Foing, L. J. Allamandola, W. Schmidt, P. Ehrenfreund, *Astronomy Astrophysics* **2002**, *390*, 1153–1170.

[438] B. A. McGuire, R. A. Loomis, A. M. Burkhardt, K. L. K. Lee, C. N. Shingledecker, S. B. Charnley, I. R. Cooke, M. A. Cordiner, E. Herbst, S. Kalenskii, M. A. Siebert, E. R. Willis, C. Xue, A. J. Remijan, M. C. McCarthy, *Science* **2021**, *371*, 1265–1269.

[439] V. S. Netchitailo, *Journal of High Energy Physics Gravitation and Cosmology* **2022**, *08*, 243–253.

[440] H. Sabbah, G. Quitté, K. Demyk, C. Joblin, *Natural Sciences* **2024**, *4*, e20240010.

[441] A. Ricca, C. Boersma, A. Maragkoudakis, J. E. Roser, M. J. Shannon, L. J. Allamandola, C. W. Bauschlicher, *The Astrophysical Journal Supplement Series* **2025**, *282*, 7.

[442] J. C. F. de Winter, S. D. Gosling, J. Potter, *Psychological Methods* **2016**, *21*, 273–290.

Permissions to Print

This work contains material adapted from work previously published in scientific journals. Permissions for reuse or reprint of the relevant content has been requested to the copyright owner.

1. Chapter 3, Section 3.2: Partially adapted from Huynh, T.-N.; **Matsokin, N. A.**; Studt, F.; Fink, K.; Sharapa, D. I. Ethylene dimerization on Ni-supported NU-1000: A comprehensive theoretical study. *The Journal of Physical Chemistry C* **2025**, *129*, 22377–22389. American Chemical Society.
2. Chapter 3, Section 3.3: Partially adapted from Avdoshin, A.; **Matsokin, N. A.**; Huynh, T.-N.; Sharapa, D. I.; Fink, K.; Studt, F.; Wenzel, W.; Kozlowska, M. Multiscale kinetic model of ethylene oligomerization in Ni-NU-1000 metal-organic framework. *npj Computational Materials* **2026**. Nature Publishing Group UK.
3. Chapter 4, Section 4.2: Partially adapted from **Matsokin, N. A.**; Kalinin, M.; Buchwald, A.; Werner, H.-J.; Fink, R. F.; Fink, K.; Sharapa, D. I. Revisiting acene dimers: A comprehensive theoretical study of a less explored conformer. *The Journal of Chemical Physics* **2026**, *164*, 13. AIP Publishing.

List of Publications

List of Publications from the Dissertation

1. **Matsokin, N. A.***; Kalinin, M.*; Buchwald, A.; Werner, H.-J.; Fink, R. F.; Fink, K.; Sharapa, D. I. Revisiting acene dimers: A comprehensive theoretical study of a less explored conformer. *The Journal of Chemical Physics* **2026**, *164*, 13. (* indicates equal contribution)
2. Avdoshin, A.; **Matsokin, N. A.**; Huynh, T.-N.; Sharapa, D. I.; Fink, K.; Studt, F.; Wenzel, W.; Kozlowska, M. Multiscale kinetic model of ethylene oligomerization in Ni-NU-1000 metal-organic framework. *npj Computational Materials* **2026**.
3. Huynh, T.-N.*; **Matsokin, N. A.***; Studt, F.; Fink, K.; Sharapa, D. I. Ethylene dimerization on Ni-supported NU-1000: A comprehensive theoretical study. *The Journal of Physical Chemistry C* **2025**, *129*, 22377–22389. (* indicates equal contribution)

Publications Outside the Scope of this Dissertation

1. Vyrko, S. A.; Polynskaya, Y. G.; **Matsokin, N. A.**; Popov, A. M.; Knizhnik, A. A.; Ratkevich, S. V.; Poklonski, N. A. Bistable carbon nanobracelets. *Journal of Molecular Modeling* **2026**, *32*, 43.
2. Nikitin, D. S.; Shanenkov, I. I.; Nassyrbayev, A.; Sivkov, A. A.; Baidyshev, V. S.; Kvashnina, Y. A.; **Matsokin, N. A.**; Pak, A. Ya.; Kvashnin, A. G. High-entropy Ti–Zr–Nb–Hf–Ta carbide and carbonitride coatings fabricated by high-speed arc discharge plasma jet. *International Journal of Refractory Metals and Hard Materials* **2025**, 107537.
3. Nikitin, D. S.; Shanenkov, I. I.; Nassyrbayev, A. R.; Sivkov, A. A.; Baidyshev, V. S.; Kvashnina, Y. A.; **Matsokin, N. A.**; Pak, A. Ya.; Kvashnin, A. G. Synthesis of high-entropy Ti–Zr–Nb–Hf–Ta carbides and carbonitrides in high-

speed arc discharge plasma jet. *Journal of Alloys and Compounds* **2025**, *1010*, 177178.

4. **Matsokin, N. A.***; Eremin, R. A.*; Kuznetsova, A. A.; Humonen, I. S.; Krautsou, A. V.; Lazarev, V. D.; Vassilyeva, Y. Z.; Pak, A. Ya.; Budennyy, S. A.; Kvashnin, A. G.; Osiptsov, A. A. Discovery of chemically modified higher tungsten boride by means of hybrid GNN/DFT approach. *npj Computational Materials* **2025**, *11*, 163. (* indicates equal contribution)
5. Tikhonov, D. S.; Gordiy, I.; Iakovlev, D. A.; Gorislav, A. A.; Kalinin, M. A.; Nikolenko, S. A.; Malaskeevich, K. M.; Yureva, K.; **Matsokin, N. A.**; Schnell, M. Harmonic scale factors of fundamental transitions for dispersion-corrected quantum chemical methods. *ChemPhysChem* **2024**, *25*, e202400547.
6. Sinitsa, A. S.; Polynskaya, Y. G.; **Matsokin, N. A.**; Kedalo, Y. M.; Knizhnik, A. A.; Popov, A. M. Formation of the icosahedral C_{60} fullerene via migration of single sp atoms and annihilation of sp-atom pairs. *Physical Chemistry Chemical Physics* **2024**, *26*, 21905–21911.
7. Kedalo, Y. M.; Polynskaya, Y. G.; **Matsokin, N. A.**; Knizhnik, A. A.; Sinitsa, A. S.; Potapkin, B. V. Theoretical study of the mechanism of hydrogen production by catalytic methane decomposition on the carbon black catalyst. *Journal of Analytical and Applied Pyrolysis* **2024**, *181*, 106611.
8. Radina, A. D.; Baidyshev, V. S.; Chepkasov, I. V.; **Matsokin, N. A.**; Altalhi, T.; Yakobson, B. I.; Kvashnin, A. G. Theoretical study of adsorption properties and CO oxidation reaction on surfaces of higher tungsten boride. *Scientific Reports* **2024**, *14*, 12788.
9. Chepkasov, I. V.; Kvashnin, A. G.; Radina, A. D.; **Matsokin, N. A.**; Jalolov, F. N.; Kvashnin, D. G.; Oganov, A. R.; Dashevsky, Z. Origin of brittle behavior of doped PbTe-based thermoelectric materials. *Applied Physics Letters* **2024**, *124*.
10. Khmelnitsky, R. A.; Martovitsky, V. P.; Bondareva, J. V.; Kolbatova, A. I.; Titova, N. A.; Goltsman, G. N.; Fedorov, F. S.; Egorov, A. V.; **Matsokin, N. A.**; Kvashnin, A. G.; Kvashnin, D. G.; Evlashin, S. A. Synthesis and characterization of niobium carbide thin films on diamond surface for superconductive application. *Journal of Alloys and Compounds* **2024**, *976*, 173266.

11. Vyrko, S. A.; Polynskaya, Y. G.; **Matsokin, N. A.**; Popov, A. M.; Knizhnik, A. A.; Poklonski, N. A.; Lozovik, Y. E. Carbon nanobracelets. *Chemical Physics Letters* **2024**, *835*, 140999.

12. Polynskaya, Y. G.; **Matsokin, N. A.**; Kedalo, Y. M.; Knizhnik, A. A.; Sinitsa, A. S.; Potapkin, B. V. A theoretical study on the mechanism of small carbon clusters growth in low-temperature plasma. *Plasma Chemistry and Plasma Processing* **2024**, *44*, 193–209.

13. **Matsokin, N. A.**; Sinitsa, A. S.; Polynskaya, Y. G.; Lebedeva, I. V.; Knizhnik, A. A.; Popov, A. M. Formation of carbon propeller-like molecules from starphenes under electron irradiation. *Physical Chemistry Chemical Physics* **2023**, *25*, 27027–27033.

14. Adilkhanova, A.; Frolova, V. F.; Yessengazin, A.; Öztopçu, Ö.; Gudun, K. A.; Segizbayev, M.; **Matsokin, N. A.**; Dmitrienko, A.; Pilkington, M.; Khalimon, A. Y. Synthesis and catalytic performance of nickel phosphinite pincer complexes in deoxygenative hydroboration of amides. *Dalton Transactions* **2023**, *52*, 2872–2886.

15. Polynskaya, Y. G.; **Matsokin, N. A.**; Sinitsa, A. S.; Knizhnik, A. A.; Potapkin, B. V. First-principles investigation of interaction between the atomic oxygen species and carbon nanostructures. *Carbon Trends* **2022**, *9*, 100201.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to Professor Karin Fink, who gave me the opportunity to join her group during uncertain times. I am sincerely thankful for her patience, the freedom she granted me, and the trust she placed in me, as well as for the opportunity to pursue genuinely interesting science. Without her, I would never have had the chance to meet so many brilliant and inspiring people as my doctoral studies in her group allowed me to. It was through Professor Fink that I came to understand what it truly means to do science and how deeply one can dig into a problem, how one can look at familiar things from a fresh and unexpected angle. I also wish to extend my special thanks to Professor Felix Studt for broadening my perspective on catalysis and for his insightful comments and advice during our joint discussions.

I would like to give particular recognition to my scientific mentor, Dr. Dmitry Sharapa, who has accompanied me since before the beginning of my doctoral studies and subsequently helped me learn new methods and theories, always sharing his thoughts and guiding me towards the finish line of my PhD. Without Dima, it would be very hard to imagine myself reaching the end of this journey. I am also grateful to my colleagues Dr. Thanh-Nam Huynh and Aleksandr Avdoshin, who, through their greater experience and seniority, helped me find my footing, offered guidance at every step, and made me a more accomplished scientist in many respects.

I would further like to acknowledge my main project, P3 within GRK2450. My sincere thanks go to Dr. Thanh-Nam Huynh, Aleksandr Avdoshin, Prof. Dr. Felix Studt, Dr. Dmitry Sharapa, Prof. Dr. Wolfgang Wenzel, and Dr. Mariana Kozlowska for welcoming my contributions and creating such a comfortable and stimulating scientific environment. Our many meetings together shaped my perspectives, inspired ideas that found their way into our joint work, and ultimately led to shared publications. Financial support from the DFG through GRK2450 is also gratefully acknowledged. I further wish to thank Sabine Holthoff for her invaluable assistance throughout my journey.

I am particularly grateful for the unexpected collaboration on the acene project, which began when I approached Dr. Dmitry Sharapa, who connected me with Mikhail Kalinin. I would like to thank our other collaborators, Prof. Reinhold F. Fink and Andrea Buchwald, with whom I spent a few wonderful days in Tübingen, as well as Prof. Hans-Joachim Werner, whose contributions significantly elevated the quality of our work.

I owe a great deal of gratitude to everyone who contributed to the improvement of this dissertation. Special thanks go to Dr. Christian Pachl, Dr. Dmitry Sharapa, Dr. Denis Tikhonov, Vladimir Fishman, Alexey Gudovannyy, Dr. Siddhi Gojare, Xenia Kraft, Mikhail Kalinin, and Helmy Pacheco Hernandez, who read the manuscript and provided an enormous number of helpful and constructive comments that allowed me to present my results far more clearly. I also thank Dr. Thanh-Nam Huynh, Aleksandr Avdoshin, Dr. Yegor Kedalo, Andrea Buchwald, and Bogdan Orlioglo for commenting on specific parts of the thesis to help improve the quality of the text. A very special thanks goes to my wife, Viktoriia Snyga, who helped me organize all the comments received and made the revision process considerably faster.

I would also like to thank all the members of my group with whom I began and finished this journey: Dr. Christian Pachl, Dr. Siddhi Gojar, Dr. Michael Harding, Dr. Juana Vázquez Quesada, Constantin Weeber, Jonas Weinhold, Marvin Kolter and Xenia Kraft. Thank you for our time together, for the jokes, conversations, and discussions we shared. Thank you for being with me for all these years and for always helping me when I needed it. If I ever fell short in any way, I sincerely apologize.

I would like to give special mention to my officemate Dr. Michael Harding, whose fascinating conversations about science, culture, society, and technology were worth years of study on their own, and who supported me in many of my ideas. I am also deeply grateful to Dr. Juana Vázquez Quesada, who was always ready to help even when I was hesitant to ask, and who consistently offered fresh perspectives on my research. And thank you, Dr. Christian Pachl, for our shared hikes and travels. You are my brother from another mother, and you know it.

Finally, I wish to express my heartfelt gratitude to my entire family: my parents, my sister, my grandmother, my two grandfathers and the rest of my family. These people support me across great distances and keep me from giving up. I will continue to do my best not to disappoint you. I miss you all dearly. I would also like to thank all my friends for their support and warm words, which helped me get through these difficult years.

Last but not least, I want to thank my beloved wife, who endured all these years by my side and supported me in ways no one else could have. I do not know what I would have done without you. You have contributed enormously to who I am and to everything I have achieved. Thank you for everything! For your love and care, which you have given me so generously over all these years.

Declaration

Declaration I, Nikita Matsokin, declare that for the polishing and translation of my thesis, I have utilized the large language models (LLM) Claude Sonnet 4.6 and Opus 4.7, developed by Anthropic. The usage of LLMs wer specifically employed for refining the language as well as for the translation of the abstract into German. I acknowledge that while Claude LLM models has been a valuable aid in the editing and translation processes, the ultimate responsibility for the content and accuracy of my thesis lies with me as the author.