Electronic properties of adsorbate arrays of π-conjugated molecules on coinage metals and hBN surfaces

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Ritu Tomar
aus
Sonipat, Haryana, India

Bonn, 2026

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer:	Prof. Dr. Thomas Bredow
Gutachter:	Prof. Dr. Stefan Grimme
Tag der Promotion:	05.03.2026
Erscheinungsjahr:	2026

Affirmation in Lieu of an Oath

I, Ritu Tomar, hereby declare that I am the sole author of this dissertation.

The doctoral thesis I submitted is my own work and was prepared without unauthorized outside assistance. I have not included text passages, graphics, or other materials from third parties or my own examination papers without identifying them. Only the sources and resources that I have indicated were used. All verbatim and non-verbatim citations from other works are identified in accordance with the citation rules for academic writing. Parts of the thesis that have been published and submitted to journals (Appendices A, B, C and D as indicated in Chapters 3, 4, 5 and 6, respectively) are identified accordingly. All other parts of the thesis have not been published elsewhere. The thesis that I submitted has not yet been submitted in any form as part of an examination/qualification course. I prepared the doctoral thesis that I submitted in accordance with the principles of good research practice. I am aware of the significance and criminal consequences of a false affirmation in lieu of an oath.

My statements are true to the best of my knowledge and belief.

Publications

The following publications serve as part of the thesis.

1. R. Tomar, L. Bernasconi, D. Fazzi, T. Bredow, *Theoretical Study on the Optoelectronic Properties of Merocyanine Dyes*, J.Phys. Chem. A **127** (2023) 9661-9671. DOI:10.1021/acs.jpca.3c04226

2. R. Tomar, A. Kny, M. Sokolowski, T. Bredow, *Modelling the monolayer formation of Merocyanine HB238 on Ag(100) surface*, J. Phys. Chem. C **129** (2025) 7427-7436. DOI:10.1021/acs.jpcc.4c08683

3. R Tomar, T. Bredow, *Machine Learning Accelerated Global Search for Adsorption Geometries of Merocyanine molecule on Hexagonal Boron Nitride*, J. Comput. Chem. **6** (2026), e70332. DOI:10.1002/jcc.70332

4. R Tomar, T. Bredow, *HB238 Adsorption on Coinage Metal Surfaces: Insights from a combined Machine-Learning Interatomic Potential and Bayesian Optimization Framework for Interfacial Electronic Structure*, J. Phys. Chem. C, (2026). DOI: 10.1021/acs.jpcc.5c07480

During the time period of this thesis, the following further studies were carried out.

5. A. Kny, M. Reimer, N. Al-Shamery, R. Tomar, T. Bredow, S. Olthof, D. Hertel, K. Meerholz, M. Sokolowski, *Chiral Self-organized single 2D-layers of tetramers of a functional donor-acceptor molecule by the surface template effect*, Nanoscale **15** (2023) 10319-10329. DOI:10.1039/D3NR00767G

Acknowledgements

I would like to express my deepest gratitude to Prof. Dr. Thomas Bredow for his invaluable guidance, support, and encouragement throughout my PhD journey. His scientific insight, patient supervision, and constant motivation have shaped the way I think and work as a researcher. I am sincerely thankful for the freedom he gave me to explore new ideas, and the constructive feedback that helped me refine my work.

I am deeply grateful to my parents, whose unwavering love, sacrifices, and belief in me have been the foundation of everything I have achieved. Their constant emotional support and encouragement have given me the strength to persevere through every challenge. I also wish to thank my siblings, who have always stood by me with affection and understanding, and my friends, who brought joy, perspective, and laughter during the most intense phases of my PhD.

I would like to thank the Template Designed Organic Electronics (TIDE) Research Training Group for providing an intellectually stimulating environment and a network of inspiring peers. The diverse discussions and exchanges within TIDE enriched my academic experience and broadened my scientific outlook. I am particularly thankful to Stefan, for carefully reading my thesis and for his thoughtful comments and discussions that greatly improved the clarity of my work. I am also very grateful to my workgroup members for the many engaging conversations that made the research process both productive and enjoyable.

This work was made possible through the generous funding and resources provided by the Deutsche Forschungsgemeinschaft (DFG). I gratefully acknowledge the computational resources made available by the University of Bonn, the HoreKa supercomputer at Karlsruhe Institute of Technology (KIT), and the Collaborative Research Center (CRC) at the University of Pittsburgh. These resources were essential for the large-scale simulations and analyses performed in this thesis.

My sincere thanks also go to Jens Mekelberger and Claudia Kronz for their continuous administrative support and assistance throughout my PhD.

Finally, I extend my heartfelt appreciation to everyone who has contributed, directly or indirectly, to this journey.

Abstract

Organic–inorganic hybrid interfaces play a pivotal role in determining the performance of next-generation optoelectronic devices. The structural organization, orientational ordering, and electronic coupling of molecular adsorbates at these interfaces govern fundamental processes such as charge transfer, exciton dissociation, and energy-level alignment. Despite significant advances in experimental surface science, theoretical prediction of adsorption geometries and interfacial electronic properties for large, conformationally flexible organic molecules remains challenging. The associated potential energy surfaces (PES) are inherently high-dimensional and feature a multitude of shallow local minima, rendering first-principles exploration computationally expensive. The present work addresses this challenge by developing a hybrid first-principles and machine-learning strategy to investigate the adsorption and electronic properties of the merocyanine dye HB238 (2-[5-(5-dibutylamino-thiophen-2-ylidenemethyl)-4-tert-butyl-5H-thiazol-2-ylidene]-malononitrile) on metallic and insulating substrates. HB238, a donor–acceptor π-conjugated chromophore, is investigated as a model system to elucidate how molecular polarity, structural flexibility, and substrate reactivity collectively influence self-assembly and interfacial electronic coupling.

Merocyanines represent an important class of polar organic semiconductors, widely used in organic photovoltaics, photodetectors, and nonlinear optical materials. Their optoelectronic responses including charge-transfer character, absorption spectra, and excitonic coupling are highly sensitive to molecular geometry, intermolecular packing, and the surrounding dielectric environment. The HB238 molecule investigated here, exemplifies this family as it exhibits a pronounced dipole moment and a strong tendency for self-organization in thin films. Experimentally, HB238 has been shown to form ordered chiral monolayers on Ag(100) surfaces, providing an ideal benchmark for validating theoretical models of molecular adsorption and self-assembly.

In Chapter 3, various electronic-structure methods were systematically benchmarked for their ability to describe the optoelectronic properties of merocyanine dyes. The study compared time-dependent density functional theory (TD-DFT), the many-body GW–Bethe–Salpeter Equation (GW-BSE) formalism, the coupled-cluster-based STEOM-DLPNO-CCSD approach, and the multireference CASSCF/NEVPT2-FIC framework. It was found that conventional TD-DFT tends to overestimate excitation energies for charge-transfer states, while GW-BSE yields excitation energies in close agreement with experiment when solvent polarization is included. The Time-Dependent Coupled-Perturbed DFT (TDCP-DFT) method provided similar accuracy at substantially reduced computational expense, offering an efficient alternative for modeling merocyanine molecules. This benchmarking established a robust theoretical foundation for subsequent studies of adsorbate systems where optical and charge-transfer properties are critically dependent on the molecular electronic structure.

Chapter 4 investigates the adsorption of HB238 on the Ag(100) surface using dispersion-corrected

density functional theory. The calculations successfully reproduced the experimentally observed chiral tetrameric -phase monolayer, characterized by flat-lying molecular arrangements stabilized through a combination of weak chemisorption to the Ag(100) surface and intermolecular hydrogen-bonding interactions. The study also highlighted the necessity of employing a more robust geometry optimization algorithm, the Preconditioned Limited-memory Broyden–Fletcher–Goldfarb–Shanno (PreconLBFGS) method, which proved essential for achieving reliable and numerically stable convergence during the structural relaxation of the large, flexible HB238 monolayer on the Ag(100) substrate.

To overcome the limitations of conventional first-principles global optimization for complex adsorbates, Chapter 5 introduces a machine-learning–accelerated global search protocol that combines the MACE (Message Passing Atomic Cluster Expansion) interatomic potential with the Bayesian Optimization Structure Search (BOSS) algorithm. This hybrid MACE×BOSS framework enabled the efficient and accurate exploration of high-dimensional PESs by combining data-driven surrogate modeling with probabilistic acquisition strategies. The approach was applied to the adsorption of HB238 on hexagonal boron nitride (hBN), an atomically flat and electronically inert two-dimensional insulator that serves as an archetype for physisorbed interfaces. Within this framework, Bayesian optimization was employed to sample a five-dimensional configurational space (two in-plane translational and three rotational degrees of freedom) for a single molecule, and a ten-dimensional configurational space for the dimer system. This methodology efficiently identified global and near-degenerate local minima with only a few hundred evaluations, representing a substantial computational reduction relative to brute-force DFT sampling.

Chapter 6 extends this machine-learning workflow to the adsorption of HB238 on the series of coinage metal surfaces (Cu, Ag, and Au), which systematically vary in surface reactivity, work function, and electronic screening behavior. Fine-tuned MACE potentials were employed in conjunction with Bayesian optimization to explore the translational and rotational degrees of freedom associated with molecular adsorption, enabling quantitative comparison of adsorption characteristics across these substrates. The analysis provides mechanistic insight into how differences in metal–adsorbate interaction strength and substrate polarizability govern molecular orientation, binding motifs, and the degree of electronic hybridization at donor–acceptor interfaces.

Contents

CHAPTER 1

Introduction

Organic electronics (OE)[1–3] constitute a prominent area of research for flexible optoelectronic devices,[4–6] such as organic solar cells (OSCs),[7–10] organic light-emitting diodes (OLEDs),[11–14] and organic field-effect transistors (OFETs) as shown in Figure 1.1.[15–17] In contrast to traditional silicon-based electronics, these devices rely on π-conjugated molecules or polymers that can be processed into thin films on various substrates.[18, 19] The performance of organic optoelectronic devices is sensitive to the molecular properties and their arrangement in the solid state, particularly at interfaces (e.g., electrode/organic or dielectric/organic interfaces).[20–23] Characteristics such as light absorption, charge-carrier mobility and energy level alignment can be tuned through molecular design and controlled assembly.[24–27] A central paradigm in organic electronic materials is the use of donor–acceptor (D–A) molecular architectures to tailor electronic properties. Early developments in the field demonstrated that incorporating donor and acceptor substituents into π-conjugated frameworks enhance charge-transport and optical response characteristics.[28, 29] Charge-transfer type small molecules and polymers were key to achieving the first efficient organic photodetectors and transistors.[30, 31]

D–A molecules are also used in organic solar cells.[32] By adjusting the donor/acceptor strength, one can control the frontier orbital energies to optimize open-circuit voltage and light-harvesting range.[33, 34] The D–A concept was integral to the development of high-performance solution-processed and vacuum-deposited organic photovoltaics.[33, 35]. Molecular packing in thin films, which is influenced by the presence of dipolar D–A structures, has been recognized as another key determinant of device performance.[29, 36] For example, Koch *et al.* and Duhm *et al.* showed that the orientation of polar organic molecules at electrode interfaces can induce sizeable interface dipoles and shift the energy alignment in organic electronic devices.[37, 38] These insights motivated extensive research into polar donor–acceptor dyes as multifunctional components in optoelectronic applications.

Merocyanines are a prominent class of dipolar D–A dyes that have been studied for decades for their exceptional optical properties. Historically, the term "merocyanine" was introduced in the cyanine dye literature by Brooker and co-workers in 1951.[39] Classic merocyanine dyes consist of an electron-donating group (often an amine or aromatic ring) and an electron-accepting group (such as a cyano, carbonyl, or heterocyclic cationic fragment) connected by a polymethine (polyene) bridge. This conjugated bridge facilitates extensive charge delocalization. A defining feature of merocyanines is the existence of two limiting resonance forms: a neutral form (often quinoid-like) and a zwitterionic form (cyanine-like with formal charges on the end groups).[40, 41] The true ground-state electronic

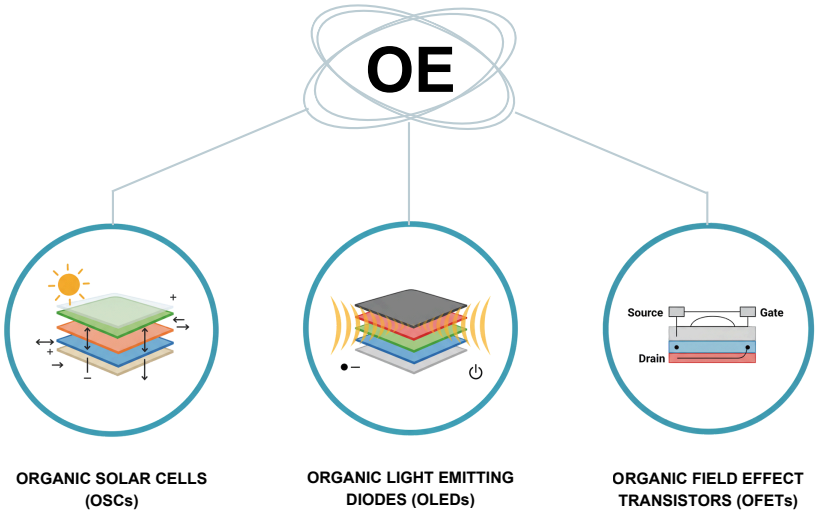


Figure 1.1: Schematic overview showing three categories of organic electronic (OE) devices: organic solar cells (OSCs), organic light-emitting diodes (OLEDs), and organic field-effect transistors (OFETs)

structure is an intermediate hybrid of these forms. The degree of bond-length alternation (BLA) along the polymethine chain serves as a structural indicator of this balance: merocyanines closer to the cyanine limit exhibit very small BLA (nearly equal bond lengths), whereas those with more neutral character show larger BLA values.[42, 43]

One consequence of this D–A resonance is that merocyanines possess large ground-state dipole moments. Dipoles on the order of 5–15 D are common for merocyanine dyes in the solid state.[44, 45] For instance, Arjona-Esteban *et al.* reported that solvent polarity can switch a merocyanine derivative between a predominantly zwitterionic structure and a more neutral cyanine-like state, accompanied by a reversal of the BLA sign.[44] Such solvent-induced conformational switching highlights the polar nature of merocyanines and their sensitivity to the dielectric environment. Stronger donor and acceptor substituents push merocyanines toward the cyanine (zwitterionic) extreme, minimizing BLA, whereas weaker push–pull combinations result in more localized charge and a pronounced polyene character.[43, 45]

In dilute solution, merocyanine molecules typically exist as monomers with well-defined absorption and emission spectra characteristic of their charge-transfer transition.[46–48] In concentrated solutions or solid phases, however, their strong dipole–dipole interactions often drive aggregation into ordered molecular stacks or clusters[49–51] The aggregation behavior of merocyanines and related dyes has been widely studied in the context of J-aggregates and H-aggregates.[49, 50, 52] In J-aggregates (head-to-tail alignment of transition dipoles), merocyanine molecules exhibit a red-shifted, narrow absorption band with enhanced oscillator strength (superradiance) and a small Stokes shift, whereas H-aggregates (face-to-face stacked dipoles) show blue-shifted, broadened absorption with quenched fluorescence.[50, 51, 53] These spectral signatures are manifestations of collective exciton coupling and have been described by the Kasha–Kobayashi exciton model and refined in modern exciton theory including vibrational coupling and charge-transfer effects.[50, 54] Importantly, the propensity of merocyanines to aggregate can be leveraged to tune their material properties like exciton delocalization in J-aggregates can lead to ultrafast energy migration and narrowband optical absorption (useful for

light-harvesting and sensing)[55, 56], while the densely packed H-aggregates can facilitate ultrafast charge separation and photocurrent generation in organic photodiodes.[55, 57] At the same time, uncontrolled aggregation in devices can be detrimental, causing phase separation or spectral shifts; therefore, understanding and controlling the aggregation of such dipolar dyes is a key materials science challenge.[58, 59]

Beyond solution and molecular aggregates, the behavior of merocyanines in bulk solid-state phases and thin films is equally intriguing. In single-crystal form, dipolar merocyanine dyes can achieve remarkably high charge-carrier mobilities if their crystal packing minimizes energetic disorder.[35, 60] A striking example is the merocyanine single-crystal field-effect transistor reported by Liess *et al.*, which attained a hole mobility of ~2.3 $cm^2\ V^{-1}\ s^{-1}$, comparable to the best acene single crystals.[60] This high mobility was enabled by a centrosymmetric (antiparallel) packing of the dipolar molecules in the crystal, effectively cancelling the internal dipole fields and reducing energetic disorder.[33, 35, 60] Such "dipole-elimination" packing motifs were deliberately targeted in molecular design: e.g., Bürckstümmer *et al.* showed that appending compatible substituents can promote antiparallel dimerization of merocyanines in solution-processed films, leading to improved bulk-heterojunction (BHJ) solar cell performance.[33, 47] In general, the solid-state packing of merocyanines strongly influences their optoelectronic properties: head-to-tail columnar stacks can support one-dimensional exciton transport and high polaron mobility along the stack axis, whereas disordered or polar packings yield localized excitations and trap-limited transport.[59] Empirically, thin-film devices based on merocyanines often perform best when some degree of structural order is introduced, for example by thermal annealing, substrate templating, or alignment techniques.[61–63] Thermal annealing of spin-coated merocyanine films was shown to induce the formation of extended J-aggregate domains, concomitantly increasing in-plane charge carrier mobility by orders of magnitude.[55] Similarly, using highly oriented substrates can force dipolar molecules into more ordered orientations, yielding improved charge transport and polarized optical properties.[63–65] These studies highlight that controlling molecular ordering in thin films is critical to harnessing the full potential of donor–acceptor molecules in devices.

Given the importance of molecular organization, there is a strong motivation to investigate how merocyanine donor–acceptor dyes arrange when adsorbed on well-defined surfaces. In device architectures, organic molecules are often in contact with inorganic interfaces (electrodes, transport layers, etc.), where interfacial effects can decisively influence performance.[66–69] Metallic substrates like Cu(111), Ag(111), and Au(111) are widely used as model surfaces in surface science to study adsorption, because they provide atomically flat, crystalline templates with varying chemical activity. Adsorption on a metal can impose specific orientations (e.g. flat-lying vs. upright) and induce interface-specific electronic effects such as charge transfer or dipole screening.[70, 71] For example, the strong image-charge screening in metals tends to reduce the apparent dipole moment of an adsorbed molecule and can pin molecular energy levels relative to the Fermi level.[70, 71] whereas on less-interacting substrates the molecule's gas-phase electronic structure is more preserved. Cu, Ag, and Au represent a spectrum from more reactive (Cu can form chemisorptive bonds or even surface-mediated deprotonation) to more inert (Au often leads to physisorption dominated by van der Waals forces)[72]. By studying a given molecule on Ag vs. Cu vs. Au, one can thus gauge the influence of molecule–substrate interactions on adsorption geometry and molecular properties as shown in Figure 1.2.[72, 73] In addition to metals, insulating or semiconducting 2D materials have gained interest as substrates that can template molecular ordering while electronically decoupling the molecules from a metal support. Hexagonal boron nitride (hBN) is a prime example of such

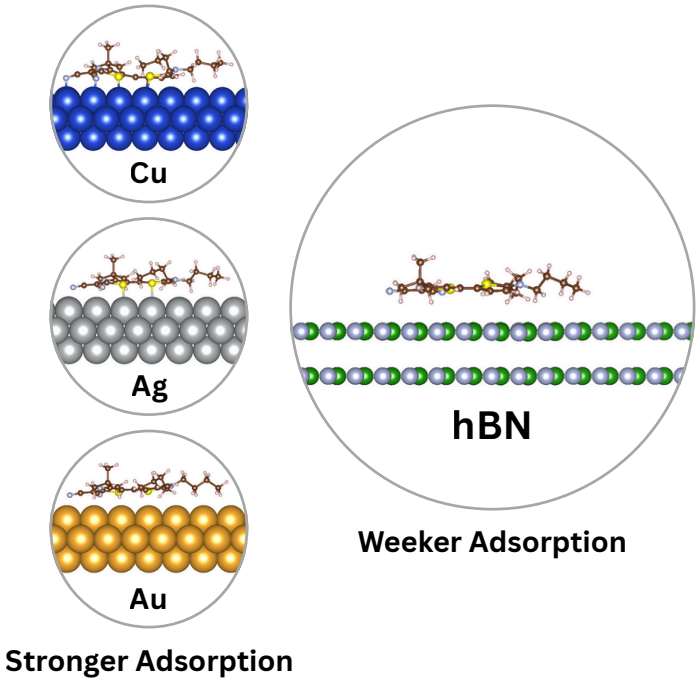


Figure 1.2: Illustative representation of the adsorption of HB238 merocyanine molecule on three metal surfaces (Cu, Ag and Au) and hBN surface.

a 2D insulator: it has an atomically smooth surface and a wide band gap, and when used as a monolayer on metal it presents a chemically inert, weakly interacting interface for adsorbates.[74, 75] Molecules adsorbed on hBN or highly oriented graphite often display more bulk-like aggregation behavior, free from the strong perturbation introduced by metal orbitals.[76–79] In practice, this can lead to long-range ordered molecular islands or monolayers with structures reminiscent of the bulk crystal packing, including the formation of chiral domains or epitaxial registry with the substrate lattice.[27, 78, 79] Understanding how a flexible, dipolar merocyanine arranges on different surfaces (metallic vs. inert) is thus not only academically interesting for surface chemistry, but also directly relevant for improving organic device interfaces (e.g. controlling orientation at electrode contacts to optimize charge injection or exciton quenching)[69]. Furthermore, from a fundamental perspective, the competition between molecule–substrate interactions and molecule–molecule interactions in directing the assembly can be elucidated by comparative studies across substrate types.[72, 80]

In this work, we focus on a particular merocyanine dye, denoted HB238 (Figure 1.3), as a model donor–acceptor molecule for adsorption studies. HB238 (2-[5-(5-dibutylamino-thiophen-2-ylidenemethyl)-4-*tert*-butyl-5*H*-thiazol-2-ylidene]-malononitrile) is a prototypical zwitterionic merocyanine that has been investigated in a variety of contexts, from thin-film photophysics to organic electronics[55, 57, 59]. It features a dialkylamino-substituted thiophene donor and a cyano-substituted thiazole acceptor, yielding a ground-state dipole moment of about 13 D in solution.[59] This large dipole, combined with a nearly planar conjugated core, causes HB238 to strongly aggregate in the solid state: in spin-coated or vapor-deposited films, HB238 readily forms H- and J-aggregate species depending on preparation conditions, as evidenced by pronounced Davydov splitting and distinct dual-band absorption spectra as shown in Figure 1.4.[55] The molecule's tendency to assemble

into ordered domains has been exploited to achieve high-performance organic photodiodes and transistors.[55, 59] These characteristics make HB238 an ideal test case for studying surface-induced ordering. Crucially, HB238 is sufficiently flexible and asymmetric (unlike, say, planar aromatic molecules such as phthalocyanines) that its adsorption configuration is not obvious *a priori*; it has multiple conformational degrees of freedom (rotations of butyl chains, thiophene twist, etc.) and could in principle adsorb in flat-lying or tilted geometries. Experimentally, the adsorption of HB238 on single-crystal surfaces has only very recently begun to be explored. Initial studies using scanning tunneling microscopy (STM) and low-energy electron diffraction have revealed that on Ag(100) HB238 can form an ordered chiral monolayer structure, with the molecules adsorbed roughly flat on the surface in a commensurate arrangement.[81, 82] This finding confirms that even a highly dipolar, bulky merocyanine is capable of two-dimensional self-assembly on a metal, opening the door to systematic investigations and modeling. By choosing HB238 as our model system, we aim to answer how such a flexible donor–acceptor molecule adsorbs on different substrates (Cu, Ag, Au, and hBN) and how the substrate choice influences molecular geometry, orientation, and electronic structure at the interface.

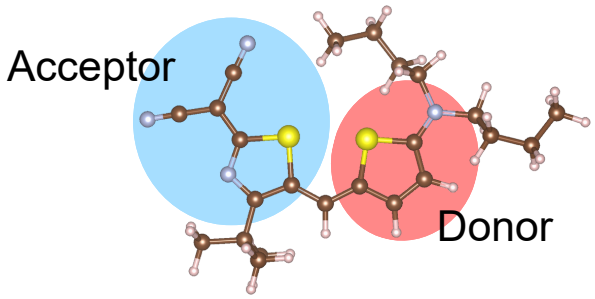


Figure 1.3: Molecular structure of the merocyanine HB238, illustrating the donor (D, red) and acceptor (A, blue) sub-units.

Theoretical studies are essential not only to complement experimental observations but also to provide molecular-level insight into adsorption energetics, interface-induced electronic effects, and structure–property relationships that are often inaccessible experimentally. The studies enable predictive modeling of molecular organization and electronic behavior across different substrates, thereby guiding experimental design and the rational development of advanced optoelectronic materials. From a computational standpoint, predicting the adsorption geometry of a large flexible molecule like HB238 on a surface is a daunting challenge. Standard first-principles approaches (e.g. density functional theory, DFT) can accurately evaluate the total energy of a given adsorbate configuration, but the potential energy surface (PES) governing the molecule–surface interaction is extremely complex and high-dimensional. HB238 has on the order of tens of internal degrees of freedom (bond rotations, conformers) and additionally six rigid-body degrees of freedom relative to the surface (lateral position, height, and orientation angles). A brute-force grid search or manual trial-and-error approach to locate the global minimum adsorption structure would be prohibitively expensive and likely to miss important configurations. Likewise, straightforward molecular dynamics or simulated annealing with first-principles forces is impractical, since the time scales required for a floppy molecule to find its

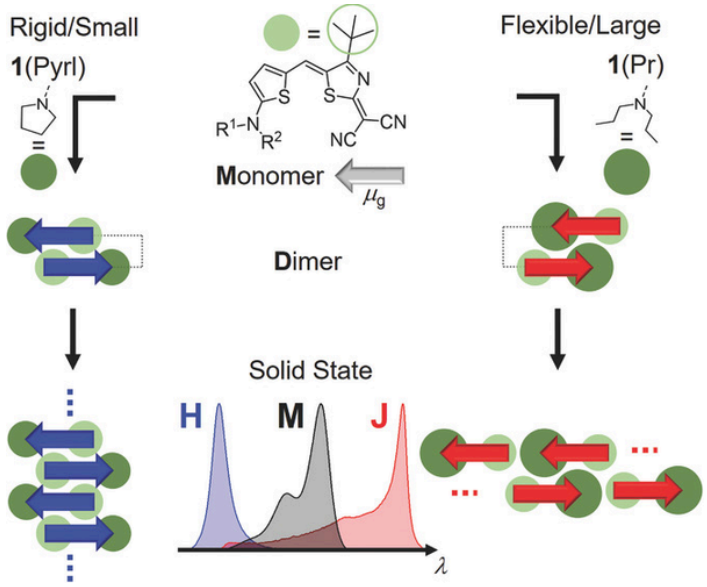


Figure 1.4: Schematic illustration of the crystalline packing variation with molecular flexibility and substituent size: rigid or small donor groups (left), antiparallel alignment of dipolar chromophores, forming card-stack structures with strong H-type coupling, while flexible or bulky donors (right) leading to slipped-stack arrangements with dominant J-type coupling in the solid state. Reprinted with permission from A. Liess et al., Advanced Functional Materials, 2018, 28, 1805058. Copyright 2018 Wiley-VCH..

optimal adsorption site and orientation may be large, and barriers between metastable states can trap the system in local minima[80, 83]. Moreover, periodic DFT calculations for these systems must be performed with large supercells to avoid spurious molecule–molecule interactions between images, and they must include dispersion corrections (or advanced van der Waals-inclusive functionals) to capture the physisorption component accurately[84]. Each single-point DFT evaluation can thus be on the order of tens or hundreds of CPU-hours, making an exhaustive search infeasible. The upshot is that traditional methods would only allow exploration of a very limited set of guessed configurations, running the risk of overlooking the true global minimum structure or other low-energy polymorphs of the adsorbed layer. This is a well-recognized bottleneck in surface science and computational materials design: the curse of dimensionality in configuration space hampers the application of first-principles calculations to complex molecular adsorbates, catalysts, or multi-component interfaces[85–87]. In recent years, several strategies have been developed to tackle this problem, including genetic algorithms and random structure search for surface adsorbates[88, 89], machine-learning accelerated potential energy scanning[90, 91], and Bayesian optimization techniques[92, 93]. In particular, approaches that combine atomistic simulations with machine learning have shown great promise for efficient structure prediction of large adsorbates[80, 94]. By employing an AI surrogate model to guide the search, one can dramatically reduce the number of expensive DFT evaluations needed to find low-energy structures[92, 94]. This thesis adopts such an approach to surmount the computational challenges of the HB238 adsorption problem.

One powerful framework used in this work is the Bayesian Optimization Structure Search (BOSS) method.[92, 95] and the use cases are shown in Figure 1.5.[92] BOSS is a global optimization algorithm that iteratively builds a surrogate model of the PES (using e.g. Gaussian process regression)

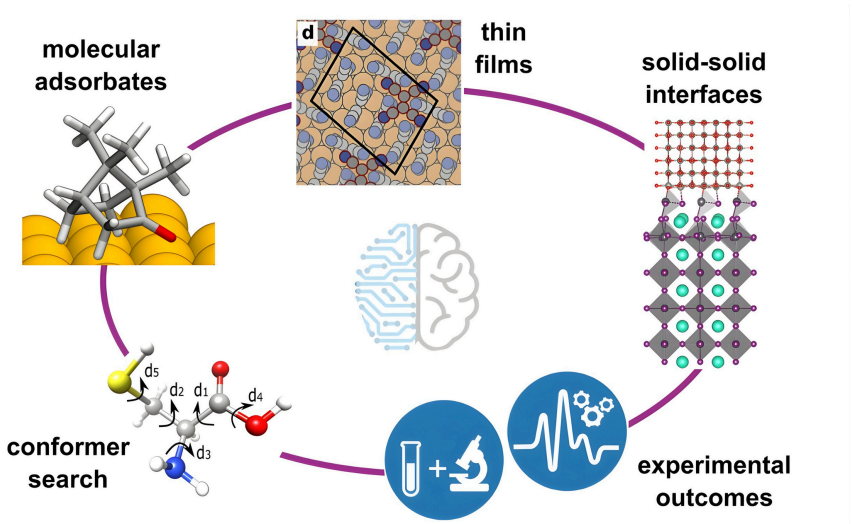


Figure 1.5: Overview diagram of the Bayesian Optimization Structure Search (BOSS) algorithm. Source: BOSS research project website (https://sites.utu.fi/boss/research/, accessed 30 October 2025.[92, 100]

and uses Bayesian optimization to decide where to sample next with DFT in order to locate stable structures efficiently. In essence, BOSS treats each new DFT calculation as an experiment to maximize information gain about the energy landscape. Starting from an initial set of configurations (which may be generated randomly or by coarse heuristics), BOSS fits a probabilistic model (mean and uncertainty) to the DFT energies. It then proposes a new configuration to evaluate by balancing exploitation (sampling where the model predicts a low energy) and exploration (sampling where model uncertainty is high), according to an acquisition function such as expected improvement. The DFT result for that new structure is then used to update the surrogate model, and the cycle repeats. This active-learning loop continues until convergence, gradually honing in on the global minimum and other low-energy minima of the adsorbate PES.[92, 95] The advantage of BOSS is that it can navigate very large configurational spaces with far fewer evaluations than a brute-force or random search. Todorović *et al.*[92] introduced BOSS in 2019 and demonstrated its effectiveness on high-dimensional optimization problems in materials science. Subsequent applications include identifying the stable binding configurations of complex adsorbates on metal surfaces,[79, 95] exploring molecular conformers in gas and solution phases,[96, 97] and even examining spin-crossover and magnetic structure optimizations.[98, 99] For example, Järvi *et al.* applied BOSS to (1*S*)-camphor on Cu(111) – a chiral, flexible adsorbate with a 6D search space – and successfully found multiple unique adsorption geometries, including the global minimum chemisorbed structure, with an order of magnitude fewer DFT calculations than a brute-force approach would require.[95] This and other studies[80, 93] have validated BOSS as an efficient and reliable tool for structure prediction at organic/inorganic interfaces, even when conventional intuition (e.g. assuming flat adsorption) might be misleading. In our work, we employ BOSS to explore the vast phase space of HB238 configurations on various surfaces, aiming to unearth the most stable adsorption arrangements systematically.

Another cutting-edge technique central to our approach is the use of machine-learned interatomic potentials to accelerate energy and force evaluations. Specifically, we utilize the MACE[101] as a highly accurate surrogate for DFT calculations.[101] MACE is part of the latest generation of

machine learning force fields that combine the atomic cluster expansion formalism[102, 103] with graph neural network architectures to achieve fast and precise predictions of atomic energies and forces.[101] This approach allows incorporation of many-body interactions to arbitrary order, and the use of symmetry-preserving neural networks makes it extremely flexible and accurate across a wide range of chemistries.[101] Recent studies have demonstrated that MACE-based potentials can reach near-DFT accuracy for both energies and forces in systems containing up to thousands of atoms, at a computational cost that is orders of magnitude lower than DFT[104, 105]. For example, Yang *et al.* showed that a machine-learned ACE potential could replicate DFT phonon spectra and thermal expansion properties of wurtzite AlN, enabling molecular dynamics simulations over nanosecond timescales that are impossible with direct DFT.[106] In the context of surface chemistry, MACE and related ML potentials have been applied to study catalytic reactions and adsorbate diffusion, providing dynamical insights beyond the reach of static DFT alone.[94, 107] A crucial feature of MACE (and many ML potentials) is the ability to estimate uncertainty or confidence in predictions, typically via an ensemble of models or Bayesian dropout. This is indispensable when integrating an ML potential into an active-learning loop, as it allows the detection of when the ML model is extrapolating beyond its training data.

By combining BOSS with MACE, we establish a powerful workflow for Bayesian optimization guided by a machine-learning surrogate – referred to here as BOSS×MACE. In this approach, the MACE potential serves as the surrogate model within the BOSS loop, providing rapid predictions of adsorption energies and thus enabling an extensive search of the configurational space with minimal direct DFT calls. Initially, a small set of DFT computations (e.g. on a few dozen random structures) is used to train the MACE model. BOSS then uses the MACE model to scan the PES and identify candidate minima, while periodically validating and retraining on new DFT data points to ensure accuracy. This synergy of Bayesian active learning and a high-fidelity ML potential greatly accelerates convergence to the global minimum structure,[108, 109] In essence, BOSS×MACE leverages the strength of ML in interpolating the PES and the strength of DFT in providing absolute accuracy when needed. The development of this combined approach builds upon recent advances by multiple research groups. Todorović and coworkers implemented on-the-fly machine-learning models in their BOSS code to handle molecular adsorbates efficiently[92]. Meanwhile, the ACE/MACE framework was refined by Drautz, Csányi, Ortner and collaborators around 2019–2022[101–103], yielding open-source tools for generating ML potentials with quantified uncertainties. Very recently, Jeschke *et al.* demonstrated a successful example of combining these ideas: they employed a MACE-derived potential within a BOSS workflow to determine the adsorption geometry of furfural on Cu(111), achieving an order-of-magnitude speedup in the global structure search and capturing subtle adsorbate distortions that plain DFT scanning had missed.[108] Their study showcased how an active-learning loop with an ML surrogate can effectively map out a complex PES (in that case involving both flat and tilted adsorption configurations of a flexible molecule) with high confidence. Inspired by these developments, our work applies the BOSS×MACE methodology to the HB238/metal and HB238/hBN systems for the first time. By doing so, we address the twin challenges of configurational complexity and computational cost that have so far limited theoretical studies of large organic adsorbates. We demonstrate that this approach can not only locate the most stable adsorption structures, but also provide insight into the manifold of metastable states, energy barriers for reorientation, and the influence of substrate on molecular conformation.

In this thesis, computational approaches for investigation of merocyanine molecules, specifically their optoelectronic properties and adsorption behaviour on several templating surfaces are presented.

Chapter 1 introduces the motivation, scientific background, and significance of merocyanine dyes as model systems for optoelectronic device applications. Chapter 2 outlines the theoretical details and computational methodologies employed, including the fundamentals of density functional theory (DFT) for periodic systems, dispersion-correction, and the implementation details of BOSS and MACE. Chapter 3 presents a detailed theoretical study of the optoelectronic properties of merocyanine molecules, evaluating the performance of several electronic-structure methods TD-DFT, GW-BSE and STEOM-DLPNO-CCSD in describing excitation energies, solvent effects, and donor–acceptor interactions. Chapter 4 focus on the modeling of HB238 monolayer formation on the Ag(100) surface, analyzing the interplay between molecule–substrate and intermolecular interactions in determining monolayer ordering and adsorption energetics. Chapter 5 introduces a machine-learning-accelerated global search for adsorption geometries of merocyanine molecules on hexagonal boron nitride, establishing the MACE × BOSS framework that couples data-driven interatomic potentials with Bayesian optimization to efficiently explore high-dimensional potential-energy landscapes. Chapter 6 extends this framework to study HB238 adsorption on coinage metal surfaces (Cu, Ag, Au), providing insight into substrate-dependent adsorption behavior and interfacial electronic structure through the combined use of MACE and BOSS. Finally, Chapter 7 summarizes the principal findings, evaluates the applicability and limitations of the MACE × BOSS methodology, and outlines perspectives for future developments in multiscale modeling of organic/inorganic interfaces. Together, these studies demonstrate how the integration of theoretical understanding and machine-learning–based modeling can advance the predictive description of complex molecular materials and their optoelectronic interfaces.

CHAPTER 2

Theoretical Details

2.1 Fundamentals of Density Functional Theory (DFT)

2.1.1 Introduction

Density Functional Theory (DFT) is one of the most widely used methods in quantum chemistry and condensed matter physics due to its balance of accuracy and computational efficiency. Instead of dealing with the complex many-electron wavefunction, DFT reformulates the problem in terms of the electron density, as established by the Hohenberg–Kohn theorems. Practical calculations use the Kohn–Sham formalism, which maps the interacting system onto non-interacting electrons with the same density, while many-body effects are included through an exchange-correlation functional. Approximations such as the Local Density Approximation (LDA), Generalized Gradient Approximation (GGA), meta-GGA, hybrid and double-hybrid functionals have been developed to improve accuracy across diverse systems. Additional techniques, including plane-wave basis sets and dispersion corrections, extend DFT's applicability to periodic systems and van der Waals interactions. This theoretical framework provides the basis for modeling merocyanine molecules in the following chapters. This chapter is based on Refs. [110–114]; for more detailed discussions and comprehensive descriptions, the reader is referred to the original literature.

2.1.2 Many-body Schrödinger Equation and Born–Oppenheimer Approximation

A key goal in molecular quantum mechanics is to solve the time-independent Schrödinger equation to calculate the electronic structures of atoms and molecules. Modern computational methods extend these calculations to increasingly complex molecular, solid and liquid systems [114].

The fundamental equation governing molecular quantum calculations systems is the electronic Schrödinger equation:

$$\hat{H}\Psi(r;R) = E(R)\Psi(r;R) \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian, $\Psi(r;R)$ is the electronic wavefunction that depends on the electronic coordinates r and parametrically on the nuclear positions R and $E(R)$ is the electronic energy, which depends on the nuclear configuration.

The molecular electronic Hamiltonian for a system with N_e electrons and N_n nuclei in atomic units (where $\hbar = m_e = e = 4\pi\varepsilon_0 = 1$) is given by:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N_e}\nabla_i^2 - \sum_{i=1}^{N_e}\sum_{I=1}^{N_n}\frac{Z_I}{r_{iI}} + \frac{1}{2}\sum_{i\neq j}^{N_e}\frac{1}{r_{ij}}. \tag{2.2}$$

Here, Z_I is the charge of the I-th nucleus, r_{iI} is the distance between electron i and nucleus I, and r_{ij} is the distance between electrons i and j. Within the Born–Oppenheimer approximation (BOA), nuclear kinetic energy is neglected and the nucleus–nucleus repulsion term is added separately after the electronic calculation.

Due to the large difference in mass between nuclei and electrons—where nuclei are approximately 10^3 to 10^4 times heavier than electrons, their motions occur on very different timescales. This observation forms the basis of the Born–Oppenheimer approximation, which assumes that the electronic and nuclear motions can be separated. Under this approximation, the total wavefunction can be written as a product of the electronic wavefunction and the nuclear wavefunction:

$$\Psi(r, R) = \psi_e(r; R)\chi_n(R), \tag{2.3}$$

where $\psi_e(r; R)$ represents the electronic wavefunction for a fixed nuclear geometry, and $\chi_n(R)$ is the nuclear wavefunction. Substituting this into the full Schrödinger equation and assuming fixed nuclei yields the electronic Schrödinger equation:

$$\hat{H}_e(R)\psi_e(r; R) = E(R)\psi_e(r; R). \tag{2.4}$$

This equation is solved to obtain the electronic energy $E(R)$, which is then used in subsequent nuclear motion treatments such as vibrational analysis or molecular dynamics.

While the Born–Oppenheimer approximation is widely used due to its ability to significantly reduce computational cost, it may fail in cases where the coupling between electronic and nuclear motion is strong. Examples include systems exhibiting Jahn–Teller distortions, conical intersections, or strong vibronic coupling.[115] Nonetheless, the BOA remains a foundational assumption in most electronic structure calculations, providing a practical and effective framework for exploring the quantum behavior of molecules and materials.

2.1.3 Hohenberg–Kohn Theorems

DFT is based on the Hohenberg–Kohn theorems, which establish the foundation for using the electron density $n(\mathbf{r})$ instead of the wavefunction Ψ to describe electronic systems. The first Hohenberg–Kohn theorem establishes that the ground state of any interacting many-particle system, subject to a fixed inter-particle interaction, is uniquely determined by its electron density $n(\mathbf{r})$.[116] In other words, the ground state wavefunction Ψ_0 can be expressed as a unique functional of the ground state electron density:

$$\Psi_0 = \Psi[n_0] \tag{2.5}$$

This result allows the ground state energy E to also be written as a functional of the electron density:

$$E[\Psi[n_0]] = \langle\Psi[n_0]|\hat{T} + \hat{V} + \hat{U}|\Psi[n_0]\rangle \tag{2.6}$$

$\hat{T}$ is the kinetic energy operator of electrons, $\hat{V}$ is the external potential due to nuclei, and $\hat{U}$ represents electron–electron repulsion. While this theorem guarantees the existence of such an energy functional $E[n_0]$, it does not provide its explicit form.

The second Hohenberg–Kohn theorem states that the true electron density is the one that yields the lowest possible value of the energy functional, corresponding to the solution of the full Schrödinger equation.

If the exact form of this functional were available, different trial electron densities could be evaluated and the one that minimizes the energy would represent the correct density. This concept is known as the variational principle and is widely used in practice, though instead of the exact functional numerous approximations are employed to make the calculations possible. [110]

In principle, knowledge of the exact energy functional would allow determination of the ground-state electron density through minimization. Once obtained, this density would provide access to all ground-state properties of the system. [117]

2.1.4 Wavefunctions and the Kohn–Sham Formalism

To approximate solutions of the many-electron Schrödinger equation, the construction of the electronic wavefunction must be considered. Since electrons are fermions, the total wavefunction must be antisymmetric under exchange, in accordance with the Pauli exclusion principle.[112] This requirement is satisfied by writing the many-electron wavefunction as a *Slater determinant* built from single-particle spin–orbitals:

$$\psi_i(r,\sigma) = \left(\sum_{\mu} c_{i\mu}\,\phi_\mu(r)\right)\chi(\sigma) \tag{2.7}$$

where $\phi_\mu(r)$ are basis functions (atomic orbitals or Gaussians), $c_{i\mu}$ are variational coefficients, and $\chi(\sigma) \in \{\alpha(\sigma), \beta(\sigma)\}$ is the spin function.

The total antisymmetric wavefunction for N electrons is then

$$\Psi(r_1,\sigma_1;\ldots;r_N,\sigma_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(r_1,\sigma_1) & \cdots & \psi_N(r_1,\sigma_1) \\ \vdots & \ddots & \vdots \\ \psi_1(r_N,\sigma_N) & \cdots & \psi_N(r_N,\sigma_N) \end{vmatrix}. \tag{2.8}$$

The Hohenberg–Kohn theorems establish that the ground-state energy is a unique functional of the electron density $n(r)$. Kohn and Sham reformulated this problem by introducing an auxiliary system of non-interacting electrons that reproduces the same $n(r)$. The energy functional is separated as

$$E[n] = E_{\text{known}}[n] + E_{\text{XC}}[n], \tag{2.9}$$

where E_{known} contains terms for non-interacting electrons that are expressible analytically and E_{XC} is the unknown exchange–correlation contribution or here: correction of kinetic energy of the non-interacting particles. In atomic units, the known part is

$$E_{\text{known}}[n] = -\tfrac{1}{2}\sum_i \int \psi_i^*(r)\nabla^2\psi_i(r)\,d^3r - \int V(r)n(r)\,d^3r + \tfrac{1}{2}\iint \frac{n(r)n(r')}{|r-r'|}\,d^3r\,d^3r' + E_{\text{ion}}. \tag{2.10}$$

Minimization of $E[n]$ leads to the Kohn–Sham equations:

$$\left[-\tfrac{1}{2}\nabla^2 - V(r) + V_H(r) + V_{XC}(r)\right]\psi_i(r,\sigma) = \varepsilon_i\psi_i(r,\sigma), \tag{2.11}$$

where $V(r)$ is the external nuclear potential,

$$V_H(r) = \int \frac{n(r')}{|r-r'|}\, d^3r' \tag{2.12}$$

is the Hartree potential, and $V_{XC}(r) = \delta E_{XC}/\delta n(r)$ is the exchange–correlation potential, including the kinetic energy correction with respect to the non-interacting particle approximation, commonly denoted as density functional approximation (DFA).

The electron density is updated self-consistently as

$$n(r) = \sum_i f_i \sum_{\sigma=\alpha,\beta} |\psi_i(r,\sigma)|^2, \tag{2.13}$$

where f_i are orbital occupations. This defines the self-consistent field (SCF) procedure.

2.1.5 Exchange–Correlation Functionals and Approximations

Since the exact form of E_{XC} is unknown, various approximations are used in DFT calculations. Depending on the chosen approximation for the exchange–correlation energy E_{XC}, the corresponding potential $\hat{v}_{\mathrm{XC}}$ is computed differently. In the Kohn–Sham (KS) scheme, [118] it is obtained as the functional derivative of the energy with respect to the electron density:

$$v_{\mathrm{XC}}(r) = \frac{\delta E_{\mathrm{XC}}[n]}{\delta n(r)}.$$

Over the years, numerous approximations for the exchange–correlation functional have been developed. [119, 120] These approximations are commonly organized into a hierarchy known as "Jacob's ladder", ranging from simple to more sophisticated models as shown in Fig 2.1.

A higher rung on Jacob's ladder generally leads to better accuracy but also to higher computational cost because more complex ingredients are involved. The lowest rung, the Local Density Approximation (LDA), depends only on the local electron density and commonly employs the Slater–Dirac exchange or the uniform electron gas model. While LDA often overbinds molecules and is rarely reliable for molecular chemistry, it performs quite well for solid-state systems where the electron density is nearly uniform, which makes the uniform electron gas approximation a reasonable description of bulk materials. The next step is the Generalized Gradient Approximation (GGA), which extends LDA by including the gradient of the electron density ($\nabla\rho$), thereby capturing spatial variations in density and improving the description of bonding and geometries. Building further, meta-GGAs (mGGAs) incorporate higher-order ingredients such as kinetic energy density, providing more flexibility and accuracy across a wide range of systems. An important development in this class is the r^{2}SCAN functional, a numerically stable refinement of SCAN; when combined with Hartree–Fock exchange it forms r^{2}SCANh, a hybrid functional. In this work, r^{2}SCANh was employed for electronic structure calculations because it offers reliable accuracy at a moderate computational cost. The next rung is occupied by hybrid functionals, which replace part of the GGA exchange with Hartree–Fock

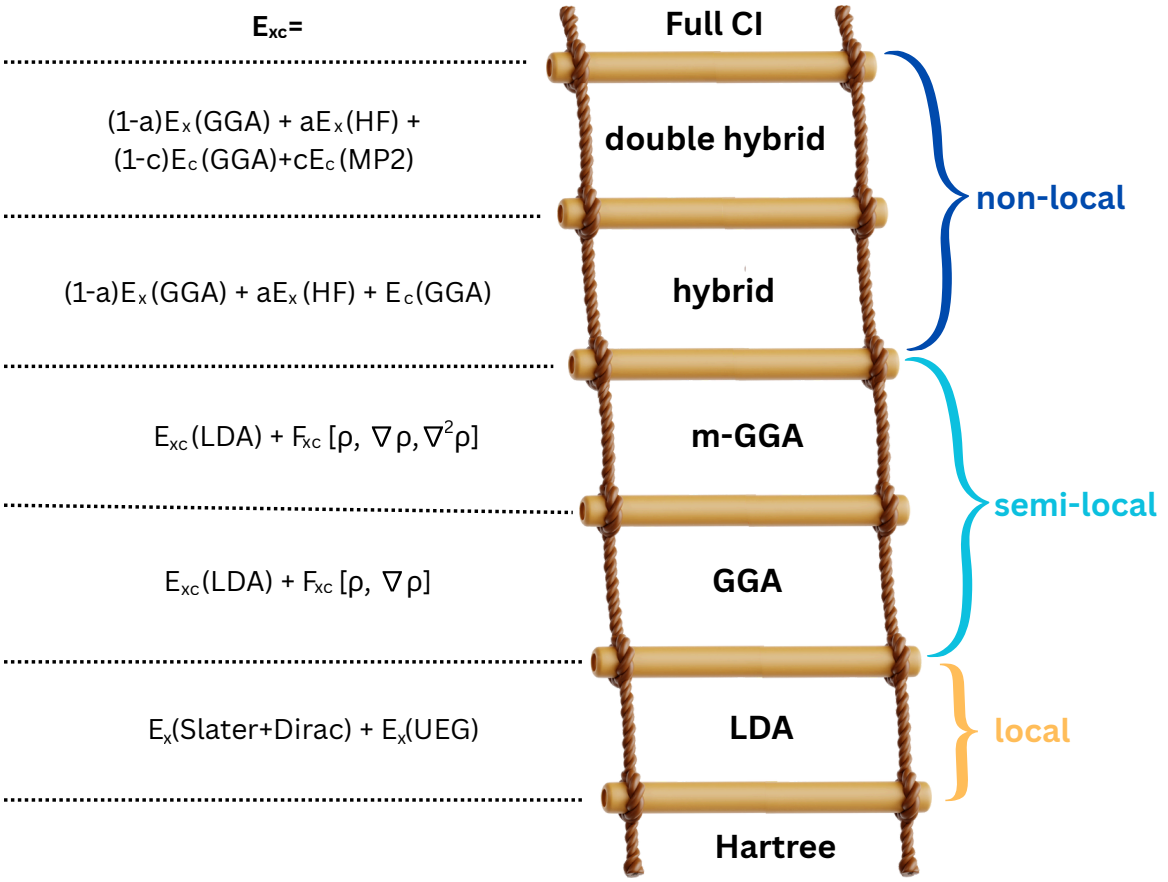


Figure 2.1: Schematic representation of Jacob's ladder of density functional approximations (DFAs).

exchange, introducing non-local effects that reduce self-interaction error and improve accuracy for many chemical and spectroscopic properties. A widely used subclass of hybrids are range-separated hybrids (RSHs), where the fraction of exact exchange depends on the electron–electron distance, thereby improving the treatment of charge-transfer states and noncovalent interactions. Finally, the top rung of practical density functional approximations is represented by double-hybrid functionals, which combine hybrid exchange with perturbative correlation, often from MP2, thus including information from virtual orbitals and achieving very high accuracy for thermochemistry, kinetics, and excitation spectra, although at increased computational cost. Above these, Full Configuration Interaction (Full CI) represents the theoretical limit, providing the exact solution within a given basis set but only feasible for very small systems.

2.1.6 Basis Sets and Plane-Wave Expansion in Periodic Systems

In quantum mechanical calculations, the basis set refers to a set of mathematical functions used to represent the electronic wavefunctions. Since solving the Schrödinger or Kohn–Sham equations analytically for real materials is generally not feasible, wavefunctions are expanded as linear combinations of these basis functions. The choice of basis set influences not only the accuracy of the solution but

also the computational cost and numerical stability of the method.[121]

For isolated systems such as molecules, localized basis functions like Gaussian-type orbitals or Slater-type orbitals are often used. However, for periodic systems which include crystalline solids and extended surfaces, plane waves are the basis functions of choice, owing to their compatibility with the periodic nature of the potential in these systems.[122]

This compatibility stems from Bloch's theorem, a fundamental result in solid-state physics, which states that the eigenfunctions of electrons in a periodic potential can be written as:

$$\psi_{nk}(r) = e^{ik\cdot r} u_{nk}(r) \tag{2.14}$$

where $\psi_{nk}(r)$ is the wavefunction of the n^{th} energy band, k is the wave vector, and $u_{nk}(r)$ is a function that has the same periodicity as the crystal lattice. This form indicates that the wavefunction can be expressed as a sum of plane waves modulated by a latticeperiodic function.[123, 124]

In practice, the electronic wavefunctions are expanded in a finite set of plane waves:

$$\psi_i(r) = \sum_{\mathbf{G}} C_{i\mathbf{G}} e^{i\mathbf{k}+\mathbf{G}} \cdot r \tag{2.15}$$

Here, $\mathbf{G}$ are the reciprocal lattice vectors, $\mathbf{k}$ is a point in the Brillouin zone, and $C_{i\mathbf{G}}$ are the expansion coefficients. Because a full expansion would involve infinitely many plane waves, a kinetic energy cutoff is introduced to truncate the basis set, retaining only those plane waves for which:

$$\frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2 \leq E_{cut} \tag{2.16}$$

The value of E_{cut} must be carefully chosen to balance computational cost with the required accuracy; convergence tests are typically performed to determine a suitable cutoff for each system and pseudopotential.

Moreover, in periodic systems, explicitly treating the core electrons requires a very high plane-wave cutoff due to the rapid oscillations of their wavefunctions near the nuclei. These sharp features, particularly the cusp behavior at the nucleus, require a large number of plane waves to be accurately represented in reciprocal space. To address this, pseudopotentials or the projector augmented-wave (PAW)[125] method are employed to replace the all-electron potential with a smoother, effective potential. This allows for an accurate treatment of valence electrons while significantly reducing the number of plane waves needed.[126]

To evaluate quantities in the reciprocal space, such as integrals over the Brillouin zone, a k-point sampling scheme is used. The Monkhorst–Pack grid is commonly employed to generate a set of discrete k-points that approximate this integral.[127] The density of the k-point mesh is particularly important for systems with partially filled bands, where the electronic structure near the Fermi level is sensitive to the sampling.

Altogether, the plane-wave basis set, combined with Bloch's theorem, appropriate pseudopotentials, and efficient Brillouin zone sampling, forms the computational backbone of periodic DFT calculations. In this work, plane-wave DFT calculations were performed using VASP (Vienna Ab initio Simulation Package)[128, 129], where pseudopotentials and plane-wave expansion are efficiently implemented.

2.1.7 Dispersion Correction

Standard DFT functionals fail to capture van der Waals (vdW) interactions, which are critical for weak intermolecular interactions, physisorption, and long-range correlation effects. This limitation arises because most commonly used exchange-correlation functionals, such as the LDA and GGA, do not include nonlocal correlation effects that are responsible for dispersion forces.[130, 131]

To address this shortcoming, semi-empirical dispersion corrections have been developed and are now widely employed. One of the most successful families of such approaches is the Grimme DFT-D framework. Among these, the D3 method has become especially popular due to its favorable balance between accuracy and computational efficiency.[132] In this scheme, the total energy is corrected by adding an explicit dispersion contribution,

$$E_{\text{DFT-D3}} = E_{\text{DFT}} + E_{\text{disp}}^{\text{D3}}, \tag{2.17}$$

where $E_{\text{disp}}^{\text{D3}}$ accounts for long-range dispersion effects not captured by semilocal functionals.

The D3 correction is formulated as a pairwise sum over atomic pairs:

$$E_{\text{disp}}^{\text{D3}} = -\frac{1}{2}\sum_{i,j}\sum_{n=6,8} s_n \frac{C_n^{ij}}{R_{ij}^n} f_{\text{damp},n}(R_{ij}), \tag{2.18}$$

where R_{ij} is the distance between atoms i and j, s_n are DFA-specific global scaling factors, and $f_{\text{damp},n}(R_{ij})$ is a damping function that smoothly switches off the correction at short range to avoid double-counting correlation effects already described by the DFA.

The dispersion coefficients C_n^{ij} describe the strength of the long-range correlation between two atoms. The leading C_6 term arises from dipole–dipole interactions and dominates at long range, while the C_8 term corresponds to dipole–quadrupole interactions and provides an important correction at shorter distances. These coefficients are not empirical constants; they are derived from first-principles quantities such as atomic polarizabilities and ionization potentials, and in D3 they are further modified to depend on the local environment through coordination numbers.

A widely used damping function in D3 is the Becke–Johnson (BJ) scheme,

$$f_{\text{damp},n}(R_{ij}) = \frac{1}{1 + 6\left(\frac{R_{r,ij}}{R_{ij}}\right)^{\alpha_n}}, \tag{2.19}$$

where $R_{r,ij}$ is a reference distance and α_n is an adjustable parameter. This ensures a smooth transition between the short-range correlation already described by the base functional and the long-range dispersion correction.

The successor method, D4, refines this approach by replacing the coordination-number dependence with density-derived atomic partial charges, providing improved accuracy in charged and polar systems.[133] The total dispersion energy in D4 is given by

$$E_{\text{disp}}^{\text{D4}} = -\frac{1}{2}\sum_{i,j}\sum_{n=6,8} s_n \frac{C_n^{ij}(q)}{R_{ij}^n} f_{\text{damp},n}(R_{ij}) - \frac{1}{6}\sum_{i,j,k} E_{\text{ATM}}^{ijk}, \tag{2.20}$$

where the pairwise contribution is similar to D3 but now depends on atomic partial charges q, and the

second term introduces the Axilrod–Teller–Muto (ATM) three-body correction,

$$E_{\mathrm{ATM}}^{ijk} = \frac{C_9^{ijk}\left(3\cos\theta_i\cos\theta_j\cos\theta_k + 1\right)}{\left(R_{ij}R_{ik}R_{jk}\right)^3}, \tag{2.21}$$

with C_9^{ijk} related to the product of the C_6 coefficients and the angles θ describing the geometry of the atomic triplet.

2.1.8 Extended Tight Binding (xTB)

Extended tight-binding (xTB) methods belong to the family of density functional tight-binding (DFTB) approaches, which approximate the total energy through a Taylor expansion of the density functional around a reference density of neutral non-interacting atoms ρ_0:

$$E[\rho] = E^{(0)}[\rho_0] + E^{(1)}[\rho_0, \delta\rho] + E^{(2)}[\rho_0, (\delta\rho)^2] + \ldots \tag{2.22}$$

Truncation at different orders defines specific models: first order yields a non-self-consistent form, second order corresponds to DFTB2, and third order results in DFTB3. The widely used GFN1-xTB[134] and GFN2-xTB[135] methods are based on third-order expansions and are designed to cover a broad range of molecules and materials.

The total energy can be expressed as a sum of contributions:

$$E_{\mathrm{tot}} = E_{\mathrm{el}} + E_{\mathrm{rep}} + E_{\mathrm{elstat}} + E_{\mathrm{disp}} + E_{\mathrm{xc}} + \ldots \tag{2.23}$$

Electronic interactions are described through effective Hamiltonians with parameters derived from reference quantum chemistry data. The Hamiltonian is constructed analogously to the Roothaan–Hall equation,

$$HC = SC\epsilon, \tag{2.24}$$

with matrix elements fitted to reproduce higher-level results. A self-consistent charge (SCC) scheme accounts for polarization effects.[136]

GFN-type extensions (geometry, frequency, noncovalent) broaden the applicability of xTB across the periodic table. Both GFN1-xTB and GFN2-xTB include empirical dispersion corrections (D3 and D4)[132, 133] to capture long-range correlation. GFN2-xTB further incorporates multipole electrostatics up to quadrupoles, anisotropic exchange contributions, and refined parameterization for noncovalent interactions such as hydrogen bonding and halogen bonding. These refinements improve accuracy for geometries, vibrational frequencies, and binding energies compared to earlier variants.

Despite its efficiency and broad applicability, xTB has known limitations. It inherits the semi-empirical nature of its parametrization and can therefore perform poorly for low-bandgap systems, underestimation of electronic excitation gaps. It has a slight tendency to overbind charge-transfer complexes, and difficulties in describing some transition-metal coordination motifs.

The GFN2-xTB method[135] has been used in this work for conformational sampling using the Conformer Rotamer Ensamble Sampling Tool (CREST).[137] Conformational ensembles were generated with the iMTD-sMTD workflow, which combines parallel metadynamics for initial conformer exploration with multiple parallel static metadynamics (sMTD) simulations for systematic refinement. In this approach, previously identified conformers are used as static references during the

umbrella sampling stage, and several sMTD runs are executed in parallel to enhance coverage of the conformational space.

After each iteration, the ensemble is analyzed through PCA and k-means clustering of dihedral descriptors to identify representative structures and eliminate redundancies. This iterative process continues until convergence of both the number of conformers and the configurational entropy is reached.

2.2 Time-Dependent Density Functional Theory (TDDFT)

Time-dependent density functional theory (TDDFT) extends ground-state DFT to excited states, with its formal foundation given by the Runge–Gross theorem,[138] which establishes a one-to-one correspondence between a system's time-dependent external potential $v_{\text{ext}}(\mathbf{r}, t)$ and its electron density $n(\mathbf{r}, t)$, for a fixed initial state. This is the time-dependent analogue of the Hohenberg–Kohn theorem.

In practice, the interacting system is replaced by an auxiliary Kohn–Sham (KS) system of non-interacting electrons evolving under an effective potential $v_{\text{KS}}(\mathbf{r}, t)$, chosen to reproduce the exact density at all times:

$$i\hbar\frac{\partial}{\partial t}\varphi_i(\mathbf{r}, t) = \left[-\frac{\nabla^2}{2} + v_{\text{KS}}(\mathbf{r}, t)\right]\varphi_i(\mathbf{r}, t), \tag{2.25}$$

with the density reconstructed as

$$n(\mathbf{r}, t) = \sum_{i}^{\text{occ}} |\varphi_i(\mathbf{r}, t)|^2. \tag{2.26}$$

The effective potential has the form

$$v_{\text{KS}}(\mathbf{r}, t) = v_{\text{ext}}(\mathbf{r}, t) + v_{\text{H}}(\mathbf{r}, t) + v_{\text{xc}}(\mathbf{r}, t), \tag{2.27}$$

where v_{xc} is the exchange–correlation (xc) potential. A formal definition of v_{xc} was provided by van Leeuwen via the Keldysh contour formalism,[139] though its exact form remains unknown, making v_{xc} the sole fundamental approximation in TDDFT.

Linear-Response TDDFT

For most applications, the linear-response (LR) formulation of TDDFT is employed.[140] Here, the response of the density $n_1(\mathbf{r}, \omega)$ to a weak perturbing potential $v_1(\mathbf{r}, \omega)$ is expressed via the density–density response function $\chi(\mathbf{r}, \mathbf{r}', \omega)$:

$$n_1(\mathbf{r}, \omega) = \int d^3r'\, \chi(\mathbf{r}, \mathbf{r}', \omega) v_1(\mathbf{r}', \omega). \tag{2.28}$$

The excitation energies correspond to the poles of $\chi(\omega)$. This response is related to the KS response χ_s by

$$\chi = \chi_s + \chi_s(f_H + f_{\text{xc}})\chi, \tag{2.29}$$

where f_H and f_{xc} denote the Hartree and xc kernels, respectively.[141] Reformulated as an eigenvalue problem, this yields the Casida equation, a practical scheme to compute excitation energies from KS orbital differences and kernel contributions.

In practice, the adiabatic approximation is almost universally adopted, assuming a time-local and frequency-independent kernel,

$$f_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') = \frac{\delta v_{\mathrm{xc}}(\mathbf{r})}{\delta n(\mathbf{r}')}, \tag{2.30}$$

which is efficient but limited, notably failing for double excitations and long-range charge transfer.[142]

Despite these shortcomings, LR-TDDFT is widely used for molecular and condensed-phase excited-state calculations due to its balance between accuracy and computational cost.[143, 144] Its efficiency and integration with ground-state DFT codes have established it as a standard tool in electronic structure theory.

2.2.1 Optical Absorption Spectra Calculations

The quantum chemistry code ORCA[145] calculates absorption (UV-Vis) spectra using Time-Dependent Density Functional Theory (TD-DFT) within the linear-response formalism [140]. The process begins with a ground-state geometry optimization, typically performed using DFT methods such as B3LYP [146] with a suitable basis set (e.g., def2-SVP or def2-TZVP [147]). Once the structure is optimized, a single-point DFT calculation is carried out to determine the molecular orbitals, orbital energies, and electron density distribution. These ground-state properties form the basis for the TD-DFT calculations, where ORCA solves the Casida equations, which arise from the linear-response of the electron density to an external oscillating electric field [140]. This yields a set of excited states characterized by excitation energies (ΔE), oscillator strengths (f), and transition dipole moments.

The oscillator strength f_{ij} for a transition from the ground state ψ_i to an excited state ψ_j is given by [148]:

$$f_{ij} = \frac{2}{3}\Delta E_{ij} \sum_{\alpha} \left|\langle \psi_i | \hat{r}_\alpha | \psi_j \rangle\right|^2 \tag{2.31}$$

where $\hat{r}_\alpha$ is the Cartesian component of the electric dipole operator and ΔE_{ij} is the excitation energy in atomic units. These oscillator strengths quantify the probability of light absorption at specific energies and are essential for simulating the optical absorption spectrum. The output of a TD-DFT calculation in ORCA typically includes a list of singlet and/or triplet excited states, their excitation energies, oscillator strengths, and dominant orbital transitions (e.g., HOMO → LUMO).

To generate a continuous absorption spectrum, the discrete excitation lines (line spectrum) are broadened using a Gaussian or Lorentzian convolution:

$$A(E) = \sum_i f_i \cdot \exp\left(-\frac{(E - \Delta E_i)^2}{2\sigma^2}\right) \tag{2.32}$$

Here, $A(E)$ is the absorption intensity at energy E, f_i is the oscillator strength of the i-th transition, and σ is the spectral broadening parameter (typically 0.1–0.3 eV). This post-processing can be performed using ORCA's `orca_mapspc` utility or with external plotting software.

For systems in solution, ORCA enables the inclusion of solvent effects through implicit solvation models such as the Conductor-like Polarizable Continuum Model (CPCM) [149] or the SMD model [150], which modify the dielectric environment and may lead to solvatochromic shifts in the excitation energies.

2.3 GWBSE and TDCP-DFT

Standard TD-DFT often shows systematic errors in excitation energies due to shortcomings of the exchange–correlation functional and the treatment of electron–hole interactions.[151, 152] More reliable predictions can be obtained with the GW approximation combined with the Bethe–Salpeter equation (GW-BSE), which incorporates many-body effects, or with Time-Dependent Coupled Perturbed DFT (TDCP-DFT), which improves the response kernel.[153]

2.3.1 GW and Bethe–Salpeter Equation (BSE) Formalism

The GW-BSE approach systematically goes beyond standard DFT by incorporating dynamic screening and explicit electron–hole interactions, which are essential for an accurate description of excited states.

In the GW formalism, quasiparticle energies are obtained by correcting Kohn-Sham eigenvalues with the electronic self-energy $\Sigma(\omega)$, constructed from the one-particle Green's function G and the screened Coulomb interaction W [154, 155]:

$$\varepsilon_n^{\mathrm{QP}} = \varepsilon_n^{\mathrm{DFT}} + \langle \phi_n | \Sigma(\varepsilon_n^{\mathrm{QP}}) - V_{\mathrm{xc}} | \phi_n \rangle, \tag{2.33}$$

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int d\omega' \, G(\mathbf{r}, \mathbf{r}', \omega + \omega') \, W(\mathbf{r}, \mathbf{r}', \omega'). \tag{2.34}$$

This treatment captures dynamical polarization and correlation effects absent in mean-field approximations, yielding reliable quasiparticle band structures and ionization potentials.

Optical excitations require a treatment of correlated electron–hole pairs, which is achieved within the Bethe–Salpeter Equation (BSE) [156, 157]. The BSE reduces to an effective eigenvalue problem for the excitonic amplitudes A_S^{vc}:

$$\sum_{v'c'} H_{vc,v'c'} A_S^{v'c'} = \Omega_S A_S^{vc}, \tag{2.35}$$

where v and c denote valence and conduction states, and Ω_S is the excitation energy of the system. The effective Hamiltonian matrix is:

$$H_{vc,v'c'} = (\varepsilon_c^{\mathrm{QP}} - \varepsilon_v^{\mathrm{QP}}) \delta_{vv'} \delta_{cc'} + K_{vc,v'c'}, \tag{2.36}$$

with the kernel $K_{vc,v'c'}$ comprising an attractive screened Coulomb term and a repulsive bare exchange contribution term [158].

By solving the BSE, exciton binding energies and oscillator strengths are obtained, enabling the calculation of absorption spectra that explicitly include excitonic effects. The combined GW-BSE approach thus provides a quantitatively accurate description of excited-state properties across molecular systems, semiconductors, and nanostructures, outperforming conventional DFT and TD-DFT particularly in the presence of strong electron correlation, long-range charge transfer, or bound excitons [157, 158].

2.3.2 Time-Dependent Coupled-Perturbed Density Functional Theory (TDCP-DFT)

TDCP-DFT is a linear-response formalism for computing excitation energies of molecules and solids. It is rooted in the same theoretical framework as Time-Dependent Density Functional Theory (TDDFT), but avoids solving the Casida eigenvalue problem. Instead, excitation energies are obtained from the poles of the frequency-dependent polarizability, derived directly from the system's response to an external oscillating electric field [153, 159, 160].

The central quantity is the frequency-dependent polarizability tensor:

$$\alpha_{ij}(\omega) = \sum_{n} \frac{\langle 0|\hat{\mu}_i|n\rangle\langle n|\hat{\mu}_j|0\rangle}{\omega - \omega_n + i\Gamma}, \tag{2.37}$$

where ω is the driving-field frequency, ω_n are the excitation energies, $\hat{\mu}_i$ is the dipole operator, and Γ is a broadening parameter. Resonances in $\alpha_{ij}(\omega)$ occur at $\omega = \omega_n$, yielding the excitation energy.

A key advantage of TDCP-DFT is that it eliminates the explicit construction of unoccupied orbitals. Instead, the first-order density response is obtained by solving the coupled-perturbed Kohn–Sham (CP-KS) equations self-consistently. From this density variation, the polarizability is computed directly. The approach is particularly efficient for extended systems, where the dimension of the unoccupied space is large, and achieves reliable results when used with hybrid functionals.

Excitation energies are extracted by fitting the polarizability in the vicinity of resonances with a rational function,

$$\tilde{\alpha}(\omega) = [A + B(\omega - C)]^{-1}, \tag{2.38}$$

such that the pole position is given by:

$$\omega_I = -\frac{A}{B} + C. \tag{2.39}$$

TDCP-DFT has been successfully applied to both molecular and solid-state systems. For instance, in Cu_2O, the method reproduces absorption peak at 2.63 eV, in excellent agreement with experimental value of 2.61 eV. [161, 162]. In molecular aggregates such as merocyanine dyes, it captures packing-induced redshifts arising from dipole–dipole interactions [163].

2.3.3 Electronic Structure: DOS, Fermi Level, and Band Gap

The density of states (DOS) quantifies the number of electronic states per unit energy and is expressed as

$$g(E) = \sum_{n\mathbf{k}} \delta(E - E_{n\mathbf{k}}), \tag{2.40}$$

where $E_{n\mathbf{k}}$ are the Kohn–Sham eigenvalues for band index n and crystal momentum $\mathbf{k}$. In practice, the delta function is replaced by a Gaussian or Lorentzian broadening to obtain a smooth representation. A more detailed decomposition is provided by the projected density of states (PDOS), obtained by projection of the wavefunctions onto localized atomic orbitals ϕ_α:

$$g_\alpha(E) = \sum_{n\mathbf{k}} \left|\langle \phi_\alpha|\psi_{n\mathbf{k}}\rangle\right|^2 \delta(E - E_{n\mathbf{k}}). \tag{2.41}$$

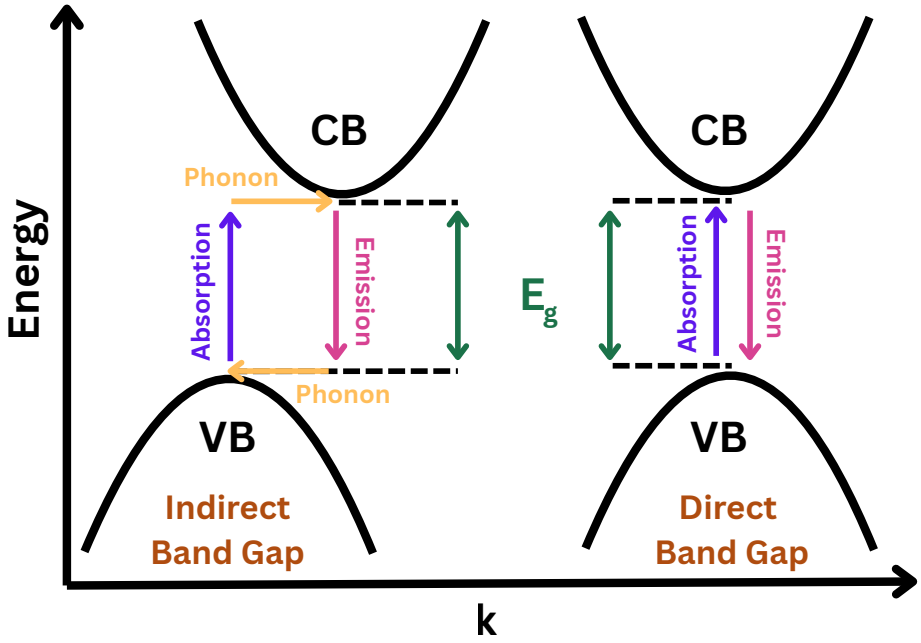


Figure 2.2: Illustration of direct and indirect Band-gap Semiconductor.

The PDOS isolates the contributions of specific orbitals and atoms to the electronic structure.

The Fermi level E_F is defined as the highest occupied electronic energy at $T = 0$ K and is determined by the condition

$$N = \int_{-\infty}^{\infty} f(E)\, g(E)\, dE, \tag{2.42}$$

with N the total number of electrons and $f(E)$ the Fermi–Dirac distribution,

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)}. \tag{2.43}$$

At zero temperature, the distribution reduces to a step function, filling all states below E_F. The electronic band gap is defined as the energy difference between the conduction band minimum (CBM) and valence band maximum (VBM),

$$E_g = E_{\mathrm{CBM}} - E_{\mathrm{VBM}}. \tag{2.44}$$

A direct gap occurs when both extrema are located at the same **k**-point, while an indirect gap arises when they occur at different points in the Brillouin zone as shown in Figure 2.2. Direct gaps permit vertical optical transitions without phonon involvement, whereas indirect transitions require phonon participation to conserve momentum.

2.4 Adsorption and Surface Interactions

The interaction of molecules with solid surfaces underpins processes in heterogeneous catalysis,[164] molecular sensing,[165, 166] organic electronics,[167] and surface functionalization.[168] Despite advances in experimental characterization and electronic-structure methods, predicting adsorption structures and energetics of large, flexible molecules remains challenging. A molecular-level understanding of adsorption geometry and binding interactions is essential for bottom-up nanotechnology approaches where precise control of molecular orientation and stability on surfaces is required.[16, 72]

2.4.1 Geometry Optimization and Global Minimum Search

The adsorption geometry is determined by the strength and character of molecule–surface interactions and is obtained through geometry optimization, where atomic positions are relaxed until forces converge below a threshold. The global minimum represents the most stable adsorption configuration. For flexible molecules, the theoretical search for the global minimum structure is preceded by a gas-phase conformational search to identify the lowest-energy conformer, which is then used as the starting structure for adsorption studies. CREST (Conformer–Rotamer Ensemble Sampling Tool)[137] combined with semi-empirical tight-binding methods such as GFN1-xTB[135] efficiently samples the conformational landscape, ensuring adsorption calculations begin with the most stable molecular geometry.

Once placed on the surface, geometry optimization is performed using standard algorithms such as the Conjugate Gradient (CG)[169] or the preconditioned limited-memory Broyden–Fletcher–Goldfarb–Shanno (PreconLBFGS)[170] methods, which minimize the total energy in systems with many degrees of freedom. Convergence to a stable adsorption geometry is verified by ensuring minimal residual forces and analyzing adsorption height, bonding character, and charge redistribution. In this way however only local minima are obtained depending on the initial structures selected in the optimization process. The search for grobal minima is described in Section 2.5.

Preconditioned LBFGS Optimization

In the present study, the PreconLBFGS algorithm was employed. It is a quasi-Newton method that approximates the inverse Hessian from successive gradients without explicit second derivatives.[171, 172] A limitation of the standard LBFGS scheme is that it typically initializes the approximation with a scaled identity matrix, which can lead to slow convergence in ill-conditioned systems such as flexible molecules adsorbed on rigid surfaces.[173]

The preconditioned LBFGS (precon-LBFGS) method addresses this by replacing the identity initialization with a physically motivated preconditioner that captures approximate curvature.[174] In practice, the preconditioner serves as an approximate inverse Hessian,

$$H_0 \approx M^{-1}, \tag{2.45}$$

where M^{-1} reflects low-cost estimates of the system's stiffness (e.g., neighbor-based force constants or sparse Laplacians). This accelerates convergence by scaling the search directions according to the underlying vibrational modes, without requiring explicit Hessian construction.

Thus, precon-LBFGS preserves the efficiency of LBFGS while greatly improving robustness for large molecule–surface optimizations, where soft and stiff modes coexist and standard optimizers

often stagnate.

2.4.2 Adsorption Energy and Stability Criteria

The stability of adsorption is quantified by the adsorption energy:

$$E_{\text{ads}} = E_{\text{adsorbed system}} - \left(E_{\text{surface}} + E_{\text{gas-phase molecule}}\right), \tag{2.46}$$

where $E_{\text{adsorbed system}}$ is the total energy of the adsorbate–surface complex, E_{surface} the clean substrate energy, and $E_{\text{gas-phase molecule}}$ the optimized isolated molecule. Negative E_{ads} values indicate exothermic, thermodynamically favorable adsorption.

Two regimes of adsorption are typically distinguished. Physisorption is dominated by weak, long-range van der Waals forces or electrostatic interaction and involves minimal charge redistribution; it is characterized by small adsorption energies and larger equilibrium distances. Chemisorption, in contrast, involves the formation of covalent bonds with significant electronic hybridization and charge transfer, yielding larger adsorption energies and shorter molecule–surface separations.

For molecular assemblies, intermolecular stabilization is captured by the interaction energy:

$$E_{\text{int}} = E_{\text{monolayer}} - \sum_i E_{\text{isolated adsorbed molecules}}, \tag{2.47}$$

where $E_{\text{monolayer}}$ is the total energy of the optimized adsorbed layer, and $\sum_i E_{\text{isolated adsorbed molecules}}$ refers to the energies of individual molecules extracted from the monolayer in their fixed monolayer geometries (without further relaxation). A negative E_{int} reflects stabilizing lateral interactions, arising from hydrogen bonding, dipole–dipole coupling, or π–π stacking. Together, E_{ads} and E_{int} provide a comprehensive measure of adsorption stability and collective behavior on surfaces.

2.4.3 Scanning Tunneling Microscopy (STM) Simulations

STM enables atomic-scale imaging of surfaces by exploiting the quantum tunneling of electrons between a conductive tip and the sample under an applied bias. Theoretical STM images can be generated to validate adsorption geometries obtained from DFT and to compare directly with experimental measurements. In this work, STM simulations were performed within the Tersoff–Hamann approximation as implemented in VASP.[129, 175]

The Tersoff–Hamann model assumes that the tip apex can be approximated by a spherical s-type orbital located at the tip position $\mathbf{r}_0$. Within this approximation, the tunneling current is proportional to the integrated local density of states (LDOS) of the sample at the tip position,

$$I(V) \propto \int_{E_F}^{E_F+E} \rho_s(\mathbf{r}_0, E)\, dE, \tag{2.48}$$

where E_F is the Fermi level, V is the applied bias voltage, and $\rho_s(\mathbf{r}_0, E)$ is the LDOS of the sample at energy E and spatial coordinate $\mathbf{r}_0$. The tip position $\mathbf{r}_0$ includes both the lateral coordinates (x, y) and the height z above the surface.

In practice, simulated STM images are obtained by mapping constant-current or constant-height contours of the integrated LDOS, which directly reflects the electronic states accessible for tunneling

in the chosen energy window.

2.4.4 X-ray Absorption Spectroscopy (XAS) Simulations

XAS provides element and site specific information on the unoccupied electronic structure and local chemical environment of adsorbed molecules. To complement experimental spectra, theoretical XAS simulations were performed in this study, using the projector-augmented wave (PAW) formalism as implemented in VASP and GPAW.[176, 177]

Within this framework, core excitations are modeled using the Supercell Core Hole (SCH) approach, where a core electron is removed from the PAW potential of the selected atom, thereby introducing an explicit core hole. To suppress spurious interactions between periodic images of the core-excited site, sufficiently large supercells are employed. The electronic density is then relaxed self-consistently in the presence of the core hole, ensuring that the screening response of the valence electrons is captured.

After convergence, the XAS spectrum is obtained from the imaginary part of the dielectric tensor, computed using Fermi's golden rule:

$$\varepsilon^{(2)}_{\alpha\alpha}(\omega) = \frac{4\pi^2 e^2 \hbar^2}{\Omega\, \omega^2 m_e^2} \sum_{\text{core},c,\mathbf{k}} 2w_{\mathbf{k}} \left|\left\langle \psi_{c\mathbf{k}} \middle| i\nabla_\alpha - \mathbf{k}_\alpha \middle| \psi_{\text{core}} \right\rangle\right|^2 \delta\left(\varepsilon_{c\mathbf{k}} - \varepsilon_{\text{core}} - \omega\right), \tag{2.49}$$

where Ω is the supercell volume, $w_{\mathbf{k}}$ are the Brillouin-zone sampling weights, $\psi_{c\mathbf{k}}$ are the conduction-band wavefunctions, and ψ_{core} is the core state wavefunction. The Dirac delta function δ enforces energy conservation between the absorbed photon energy ω and the transition from the core level $\varepsilon_{\text{core}}$ to the conduction state $\varepsilon_{c\mathbf{k}}$.

This methodology enables quantitative analysis of local electronic structure, coordination environment, and oxidation state, providing a direct link between first-principles simulations and experimental spectra.

2.5 Machine Learning Potentials and Global Optimization Algorithms

Machine Learning Interatomic Potentials (MLIPs) provide a data-driven framework for predicting potential energy surfaces (PES) with near–DFT accuracy at a fraction of the computational cost.[178] Instead of solving the electronic structure problem explicitly, MLIPs learn a mapping between atomic configurations and corresponding energies and forces, trained on reference quantum-mechanical data. By construction, MLIPs must respect fundamental physical symmetries, including translation, rotation, and permutation invariance of identical atoms.

A rigorous foundation for modern MLIPs is the Atomic Cluster Expansion (ACE), which systematically expands atomic energies in terms of many-body basis functions of atomic neighbor densities.[102] In practice, the ACE framework captures the local atomic environment through systematically increasing body-order terms—pair, triple, quadruple, and higher-order correlations between neighboring atoms. The pair (two-body) terms describe radial interactions between an atom and each of its neighbors, similar to classical pair potentials. The triple (three-body) terms introduce angular dependencies by considering atom triplets, thereby encoding how bond angles influence local energetics. Higher-order terms capture more complex geometric correlations and many-body

effects. Each contribution is expressed in a complete and orthogonal basis constructed from products of radial basis functions and spherical harmonics, ensuring a compact yet systematically improvable description of the atomic neighbor density. This hierarchical expansion allows ACE to bridge smoothly from simple empirical pair potentials to highly expressive many-body models approaching quantum mechanical accuracy.

The ACE representation of the energy of atom i can be expressed as:

$$E_i = \sum_{\nu} c_{\nu} B_{\nu}(\{\mathbf{r}_{ji}\}), \tag{2.50}$$

where c_{ν} are expansion coefficients, and B_{ν} are basis functions constructed from products of radial basis functions and spherical harmonics that encode the geometry of the local atomic environment $\{\mathbf{r}_{ji}\}$.

The **MACE** (Message Passing Atomic Cluster Expansion) framework extends ACE by integrating it with equivariant message-passing neural networks (MPNNs).[101] In MACE, atomic environments are represented as geometric tensors derived from spherical harmonics and radial functions, and atomic interactions are updated through tensor products and Clebsch–Gordan contractions. Crucially, the architecture enforces *rotation equivariance* (outputs transform covariantly under SO(3) rotations) and *reflection invariance* (energies do not change under inversion). This guarantees that the predicted energies and forces are physically consistent with the symmetries of the underlying potential energy surface.

The total energy is written as a sum of atomic contributions learned by the model:

$$E_{\text{total}} = \sum_{i} E_i(\{\mathbf{r}_{ji}\}), \tag{2.51}$$

with E_i obtained via graph-based message passing. The model separates body-order (expressivity of geometric correlations) from the number of message-passing layers (spatial range). This decoupling enables high accuracy with shallow networks: in practice, MACE achieves near–DFT fidelity with only 1–2 message-passing layers of the neural network and body order up to 4, substantially lowering training and inference costs relative to earlier equivariant architectures.[179]

Pretrained MACE-MP-0 Model:

In this work, the pretrained `mace_mp` model (MACE-MP-0)[101] was employed, which was trained on a dataset of more than 1.5 million DFT-calculated molecular dynamics configurations (MPtrj) derived from Materials Project structures, covering 89 elements. Reference energies and forces were computed using the PBE(+U) functional.[180] Key training parameters include:

- `num_interactions`: 2 layers
- `hidden_irreps`: $256 \times (0e + 1o + 2e + 3o)$
- Radial cutoff: $r_{\text{max}} = 5$ Å
- Radial basis: Bessel functions with polynomial envelope
- Optimizer: AMSGrad, initial learning rate 10^{-3}
- Loss weights: $\lambda_E = 1$, $\lambda_F = 1000$

To specialize the model for adsorption studies, the pretrained MACE-MP-0 was fine-tuned on custom datasets comprising adsorption geometries and energies of the HB238 molecule on various metal surfaces. This transfer-learning step increased sensitivity to system-specific interactions, such as charge redistribution, molecule–surface bonding, and dispersion forces. The resulting fine-tuned potential retained the efficiency of MACE while achieving accuracy comparable to DFT for adsorption energetics. It was subsequently employed for large-scale geometry optimization and potential energy surface exploration within the Bayesian Optimization Structure Search (BOSS) framework.[92]

2.5.1 Bayesian Optimization Structure Search

BOSS is a data-efficient global optimization framework developed to identify low-energy configurations of complex atomic and molecular systems. It is particularly well-suited for structure prediction tasks involving hybrid organic–inorganic interfaces, such as molecular adsorption on surfaces [92]. It couples first-principles evaluations with a Bayesian optimization strategy, constructing a probabilistic surrogate model of the PES using a Gaussian Process (GP). The GP posterior provides both a predictive mean $\mu(\mathbf{x})$ and an uncertainty estimate $\sigma(\mathbf{x})$, which are combined in an acquisition function to guide sampling.

For training data $\mathcal{D} = \{\mathbf{x}_i, E_i\}$, the GP posterior mean and variance are given by:

$$\mu(\mathbf{x}) = k(\mathbf{x}, X)\,[K + \delta^2 I]^{-1}\mathbf{E}, \tag{2.52}$$

$$\sigma^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - k(\mathbf{x}, X)\,[K + \delta^2 I]^{-1}k(X, \mathbf{x}), \tag{2.53}$$

where $k(\mathbf{x}, \mathbf{x}')$ is the kernel function, ρ the signal variance, ℓ the length scale, and δ the regularization parameter accounting for numerical noise. A Gaussian-type kernel is typically used:

$$k(\mathbf{x}, \mathbf{x}') = \rho^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell^2}\right). \tag{2.54}$$

New candidate configurations are chosen via the Lower Confidence Bound (LCB) acquisition function:

$$\mathrm{LCB}(\mathbf{x}) = \mu(\mathbf{x}) - \kappa\,\sigma(\mathbf{x}), \tag{2.55}$$

where the exploration–exploitation parameter κ controls the balance between sampling uncertain regions (σ) and refining low-energy predictions (μ). A representative illustration of this Bayesian optimization process is shown in Figure 2.3, where the surrogate Gaussian Process, its associated uncertainty, and the acquisition function are depicted in a 1D schematic.

At each optimization step, the GP posterior is updated with newly evaluated configurations, refining both $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$. The acquisition function—here, the Lower Confidence Bound—then determines the next configuration to sample by minimizing the trade-off between exploration of uncertain regions (large σ) and exploitation of promising low-energy regions (small μ). This iterative process progressively improves the surrogate model while efficiently converging toward the global minimum of the potential energy surface.

Initial sampling is performed using Sobol sequences, which are quasi-random low-discrepancy sequences that uniformly explore the high-dimensional search space. This ensures broad initial coverage before refinement by Bayesian optimization.

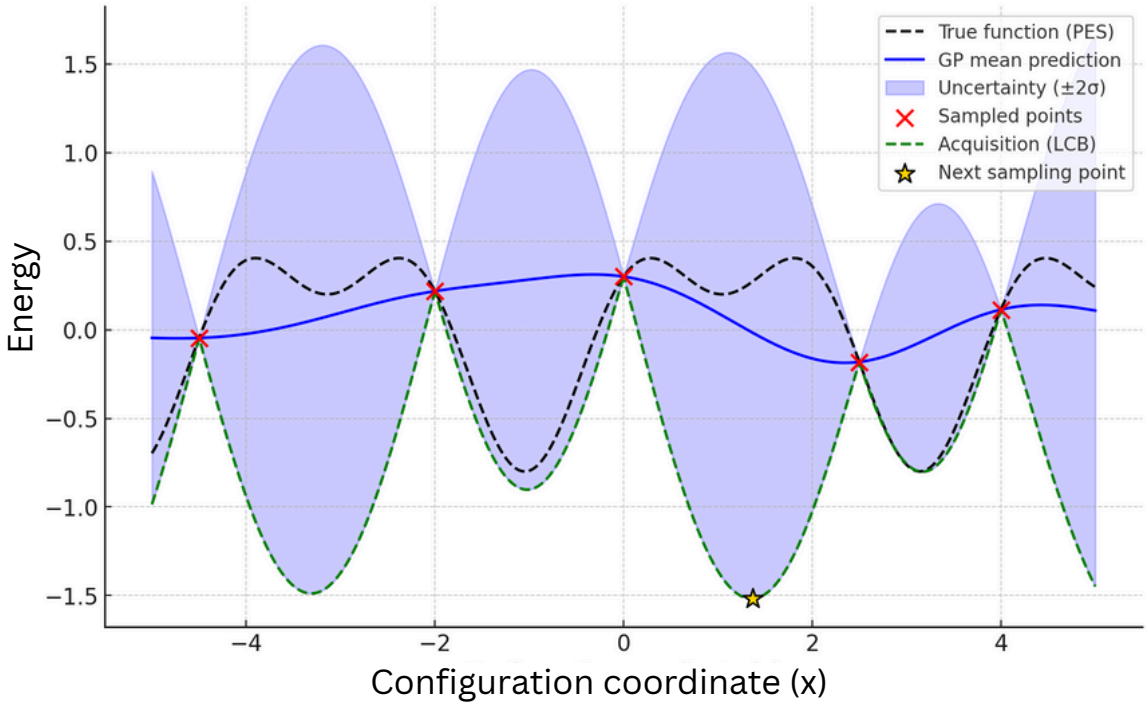


Figure 2.3: Representative schematic of Bayesian Optimization in 1D. The true potential energy surface (black dashed) is approximated by a Gaussian Process surrogate (blue line) with uncertainty bands (shaded). Sampled configurations (red points) update the surrogate, while the acquisition function (green dashed) guides exploration. The next sampling point, chosen as the minimum of the acquisition function, is indicated by the gold star.

The configurational degrees of freedom for molecular adsorption typically include translational (x, y, z) and rotational (α, β, γ) coordinates, with the adsorption energy expressed as:

$$E_{\text{ads}} = E(x, y, z, \alpha, \beta, \gamma). \tag{2.56}$$

Rigid-body approximations are employed to constrain intramolecular degrees of freedom, which is appropriate for physisorption and moderate chemisorption scenarios.

BOSS has demonstrated outstanding efficiency in molecular adsorption studies. Todorović *et al.*[92] reconstructed a 5D PES using fewer than 700 DFT evaluations with sub-angstrom accuracy. Järvi *et al.*[181] identified eight distinct adsorption geometries of camphor on Cu(111) from 1380 BOSS-sampled points, achieving structural deviations < 0.2 Å and energy deviations < 0.05 eV compared to relaxed DFT results. Extensions of BOSS to semiconducting surfaces further confirmed its ability to resolve adsorption geometries, configurational energy barriers, and charge-transfer trends essential for surface engineering.[182]

CHAPTER 3

Theoretical Study on the Optoelectronic Properties of Merocyanine Dyes

Ritu Tomar[†], Leonardo Bernasconi[‡], Daniele Fazzi[§], Thomas Bredow[†]

Received: 23 June 2023
Published online: 14 November 2023

Reprinted in Appendix A (adapted) with permission[¶] from
R. Tomar, L. Bernasconi, D. Fazzi, and T. Bredow, *Theoretical Study on the Optoelectronic Properties of Merocyanine Dyes*, *J. Phys. Chem. A* **127** (2023) 9661-9671.
DOI:10.1021/acs.jpca.3c04226

Own Contributions

- performing of all calculations except coupled cluster and CAS-SCF
- interpretation of the results
- writing of the manuscript

§

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[‡] Center for Research Computing and Department of Chemistry, University of Pittsburgh, PA 15260, USA
[§] Dipartimento di Chimica ”Giacomo Ciamician", Universitá di Bologna, 40126 Bologna, Italy

Merocyanines (MCs) are highly polar π-conjugated donor–acceptor molecules widely studied for their self-assembly and optoelectronic behavior, with applications in nonlinear optics, photorefractivity, and organic solar cells. Each molecule contains electron-donating and electron-accepting groups bridged by an ethylene or polyethylene chain. A key structural parameter is the bond-length alternation (BLA), defined as the difference between the average lengths of single and double C–C bonds along the π-conjugated bridge. The magnitude and sign of BLA distinguish neutral from zwitterionic resonance forms: small or vanishing BLA corresponds to the cyanine-like state, while large positive or negative BLAs mark increasing contributions of neutral or zwitterionic character, respectively. Standard generalized-gradient DFT functionals often fail to reproduce the experimentally observed BLA and excitation energies accurately.

In this study, several theoretical approaches like TD-DFT, GW-BSE, STEOM-DLPNO-CCSD, CASSCF/NEVPT2-FIC, and the TDCP-DFT were benchmarked for the computation of optoelectronic properties with respect to experimental results on four representative MCs (D1A1, D1A2, D1A3, D2A4) composed of donor units such as 5-ethenyl N,N-disubstituted thiophenamine (D1) and a Fischer-base derivative (D2) paired with electron-withdrawing acceptors A1–A4. Initial conformer searches were carried out with CREST using GFN1-xTB. The geometry of the most stable conformers for each respresentative molecule were optimizated using PBE, PBE0, B3LYP, CAM-B3LYP, B97X, and B2PLYP functionals with def2-TZVP and Grimme-D3 dispersion corrections. Solvent effects were included via the CPCM model.[183] Dispersion-corrected double-hybrid B2PLYP-D3 and hybrid PBE0-D3 gave the best match to X-ray BLAs, though both still underestimated and occasionally produced the wrong sign for D1A2 and D1A3. Reoptimizing structures in polar solvents such as acetone or dichloromethane changed the BLA sign—indicating a neutral–zwitterionic shift and lowered mean absolute deviations relative to X-ray data, demonstrating the importance of solvation effects.

Optical properties were then evaluated. HOMO energies and vertical $S_0 \rightarrow S_1$ excitation energies were calculated using TD-DFT with various DFAs. Hybrid functionals such as PBE0 and B3LYP reproduced experimental HOMO energies within ~0.1 eV, whereas CAM-B3LYP underestimated the ionization energy. All TD-DFT methods overestimated S_1 by 0.7–0.9 eV relative to experiment, and increasing the basis set from def2-DZVP to def2-QZVPP did not correct this. Monte Carlo sampling of conformers showed only minor broadening of the absorption band (on the order of 0.1 eV), indicating that conformational variability does not explain the discrepancy. It was therefore concluded that TDDFT is not suitable for accurate calculation of the excitation energies of MCs.

Benchmarking excitation energy across correlated methods revealed mean absolute deviations of 0.34 eV for GW-BSE, 0.11 eV for STEOM-DLPNO-CCSD, and 0.14 eV for CASSCF/NEVPT2-FIC and TDCP-DFT. Applying solvent shifts (δS) computed with TD-B2PLYP further reduced GW-BSE errors to $\approx$ 0.05–0.07 eV, yielding near-quantitative agreement with experiment. Finally, TDCP-DFT calculations for the crystalline MC HB238 showed S_1 values 0.1–0.3 eV lower than for the gas-phase molecule due to weak exciton coupling. The accurate modeling of merocyanine dyes requires double-hybrid DFT geometries with dispersion, explicit solvent treatment, and correlated excited-state methods—GW-BSE being the most reliable, while TDCP-DFT offers a computationally efficient alternative for both molecular and crystalline phases.

CHAPTER 4

Modelling the monolayer formation of Merocyanine HB238 on Ag(100) surface

Ritu Tomar[†], Anna J. Kny[‡], Moritz Sokolowski[‡], Thomas Bredow[†]

Received: 24 December 2024
Published online: 8 April 2025

Reprinted in Appendix B with permission[¶] from
R. Tomar,Anna J. Kny, M. Sokolowski, and T. Bredow, *Modelling the monolayer formation of Merocyanine HB238 on Ag(100) surface*, *J. Phys. Chem. C* **129** (2025) 7427-7436.
DOI:10.1021/acs.jpcc.4c08683

Own Contributions

- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[‡] Clausius Institute of Physical and Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to the American Chemical Society

The adsorption behavior of the merocyanine HB238 (2-[5-(5-dibutylamino-thiophene-2-yl-methylene)-4-tert-butyl-5H-thiazol-2-ylidene]-malononitrile) on the Ag(100) surface has been characterized experimentally by STM, SPA-LEED, XPS, and UPS by Kny et al.[82] These studies revealed that HB238 forms a chiral tetrameric monolayer structure, referred to as the α-phase, which is structurally distinct from its bulk crystal forms. Such symmetry breaking and chiral ordering highlight the role of molecule–substrate interactions in directing merocyanine self-assembly.

In the present work, quantum-chemical calculations were performed to complement these experimental results and provide a microscopic description of adsorption geometries, energetics, and the resulting electronic structure. A full conformer search of HB238 in the gas phase was conducted using the Conformer–Rotamer Ensemble Sampling Tool (CREST), which identified 356 conformers within 6 kJ mol^{-1}. Geometry optimizations of these conformers at the PBE/D3 level yielded the most stable gas-phase structure, which served as the starting point for adsorption modeling.

The Ag(100) surface was represented by a three-layer slab, with the bottom two layers fixed and the topmost layer relaxed. The experimentally determined α-phase unit cell parameters (a = 23.30 Å, b = 24.69 Å, γ = 103.43°) were employed to construct monolayer models containing four HB238 molecules per unit cell. Multiple initial orientations were generated by systematically varying the azimuthal rotation of the molecules while maintaining fourfold symmetry. Subsequent optimization and adsorption energy evaluation identified the most stable configuration (model A), which closely matches the experimental α-phase arrangement.

The optimized structure exhibits a face-on adsorption geometry, in which the π-conjugated backbone of HB238 lies nearly parallel to the Ag(100) surface. The calculated adsorption energy of −3.51 eV per molecule indicates strong stabilization through a combination of dispersion and electrostatic interactions, complemented by intermolecular C–H$\cdots$N hydrogen bonds. The monolayer structure was stabilized through a balance of weak chemisorption to the Ag substrate and intermolecular interactions. Charge analysis using DDEC6 population analysis confirmed minor charge redistribution between HB238 and the Ag substrate, with a small depletion at the sulfur sites and accumulation on the tert-butyl carbons. Simulated STM images, obtained within the Tersoff–Hamann approximation,[175] reproduce the experimentally observed tetrameric contrast, confirming the assignment of the α-phase. Adsorption height analysis shows that the donor sulfur atom lies at 3.09 Å and the acceptor sulfur at 2.88 Å above the surface, consistent with partial covalent character in the S–Ag interaction.

Electronic structure calculations using hybrid functionals (r^{2}SCANh, HSE06, and B3LYP) revealed that the meta-GGA hybrid functional r^{2}SCANh best reproduces the S bonding as observed in experimental UPS spectra, with calculated HOMO and HOMO−1 energies of −1.80 eV and −2.90 eV, respectively with respect to the Ag Fermi energy. The projected density of states (PDOS) and charge-density-difference maps indicate weak hybridization between molecular frontier orbitals and Ag states, accompanied by slight charge transfer from the surface to the molecule, which stabilizes the adsorbed layer. Calculated x-ray absorption spectra (XAS) further demonstrate that the donor and acceptor sulfur atoms in the monolayer exhibit nearly degenerate S 2p features, consistent with experimental observations and reflecting surface-mediated charge redistribution.

The excellent agreement between calculated and experimental STM, PDOS, and XAS data confirms that the proposed model accurately captures the structure and bonding of the α-phase. The results establish HB238 as a prototypical donor–acceptor merocyanine for studying charge-transfer processes at organic–metal interfaces.

CHAPTER 5

Machine Learning Accelerated Global Search for Adsorption Geometries of Merocyanine molecule on Hexagonal Boron Nitride

Ritu Tomar[†], Thomas Bredow[†]

Recieved: 31 October 2025
Published online: 17 February 2026

Reprinted in Appendix C (adapted) with permission [¶] from
R. Tomar and T. Bredow, *Machine Learning Accelerated Global Search for Adsorption Geometries of Merocyanine molecule on Hexagonal Boron Nitride*, *J. Comput. Chem.*, **6** (2026), e70332
DOI:10.1002/jcc.70332

Own Contributions

- coding and implementation of MACE×BOSS
- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to Wiley

Understanding how large organic molecules adsorb on solid surfaces is crucial for designing catalysts, sensors, and optoelectronic materials.[184–187] Hexagonal boron nitride (hBN) is considered a templating surface candidate for self-assembled monolayer growth because it is atomically flat, chemically inert, and supports ordered molecular organization.[78, 188–190] However, the adsorption of merocyanine dyes like HB238 on hBN has not been explored. Since HB238 is a dipolar donor–acceptor chromophore with tunable optical properties, studying its interaction with hBN can reveal how such molecules form ordered, weakly bound monolayers—an essential step toward developing hybrid organic–2D optoelectronic system.[81, 82] Traditional DFT-based global optimization methods are computationally expensive for large, flexible molecules like HB238. This study aims to bridge that gap by combining machine learning (MACE) potentials with Bayesian optimization (BOSS) to efficiently and accurately explore the adsorption landscape of HB238 on hBN, providing fundamental insights into dipolar dye–surface interactions.

Initial conformational sampling of gas-phase HB238 was performed using CREST as described in the previous chapter. The most stable conformer was relaxed using an untrained MACE potential and positioned on a periodic hBN(001) slab. A training set of adsorption geometries for single and dimer configurations was generated and optimised using PBE/D3 density functional theory (DFT) in VASP. Three machine-learning potentials were subsequently fine-tuned: a model (MACE-1M) trained on 1191 single-molecule structures, a second model (MACE-1M2M) expanded to 2354 structures including dimers, and a third model (MACE-1M2M+C) incorporating 356 gas-phase conformers and total 2660 structures. Training employed weighted energy and force targets producing energy mean absolute errors below 2.5 meV per atom and force errors under 5 meV/Å^{-1}. These potentials were integrated into a Bayesian optimisation framework; each search started from 20 random samples followed 980 optimisation cycles, with translational bounds chosen to avoid unphysical dimer overlaps.

The fine-tuned MACE potentials accurately ranked low-energy adsorption configurations, unlike the untuned foundation model, which failed to distinguish stable structures. In the five-dimensional (5D) single-molecule search—exploring two in-plane translations (x, y) and three rotational degrees of freedom (α, β, γ)—the top 700 configurations were 0.2 eV above the global minimum energy, featuring a flat-lying HB238 molecule with sulfur atoms over hollow sites and high surface mobility. For dimers, a ten-dimensional (10D) configurational search was performed, combining translations and rotations of two molecules. The lowest-energy dimer showed two flat, parallel HB238 molecules stabilized by lateral hydrogen bonding. Overall, 175 dimer configurations had energies 0.1 eV above the minimum energy, with adsorption energies of –1.69 eV (monomer) and –1.95 eV (dimer), reflecting a flat potential energy landscape dominated by weak intermolecular interactions.

Electronic-structure analysis corroborates the physisorptive nature of adsorption. Projected density of states calculation indicate that the frontier orbitals remained sharp and localized. Charge density difference maps show only minor redistribution around the cyano groups and little change on the hBN substrate, implying weak electronic coupling and the absence of covalent bonding. This study demonstrates that machine-learning-accelerated global search can efficiently map the adsorption landscape of large organic molecules on 2D materials. The lack of strong site specificity and the flat potential surface imply that HB238 molecules can diffuse on hBN, potentially hindering ordered monolayer formation, while the approach itself offers a transferable framework for other organic adsorbates.

CHAPTER 6

HB238 Adsorption on Coinage Metal Surfaces: Insights from a combined Machine-Learning Interatomic Potential and Bayesian Optimization Framework for Interfacial Electronic Structure

Ritu Tomar[†], Thomas Bredow[†]

Received: 31 October 2025
Published online: 20 April 2026

Reprinted in Appendix D (adapted) with permission [¶] from
R. Tomar and T. Bredow, *HB238 Adsorption on Coinage Metal Surfaces: Insights from a combined Machine-Learning Interatomic Potential and Bayesian Optimization Framework for Interfacial Electronic Structure*, *J. Phys. Chem. C*, (2026)
DOI: 10.1021/acs.jpcc.5c07480

Own Contributions

- finetuning of MACE potentials
- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to the American Chemical Society

The adsorption behavior of HB238 monolayers on Ag(100) surfaces was investigated in Chapter 4. In the present study, we focus on the adsorption of a single HB238 molecule on Cu(100), Ag(100), and Au(100) surfaces using a hybrid machine learning–based framework that integrates MACE interatomic potentials with Bayesian optimization (MACE×BOSS), as described in Chapter 5. While the PreconLBFGS optimizer employed in Chapter 4 offers improved convergence compared to conventional conjugate gradient methods, identifying the true global minimum on the complex potential energy surface remains a significant challenge.

The choice of three different metal substrates—Cu, Ag, and Au allows systematic comparison of molecule–surface interactions across metals with varying chemical reactivity and work functions. A machine-learned interatomic potential MACE based on the MACE-MP-0 model was fine-tuned for each metal substrate separately. The finetuned potential is used as an optimizer with Bayesian Optimization Structure Search (BOSS) to efficiently sample 5 configurational degrees of freedom (lateral translations along the surface lattice vectors and 3 rotational angles). Each run starts with 20 random configurations followed by 980 adaptive evaluations, enabling exploration of several candidate geometries at negligible cost. Promising structures were subsequently relaxed with dispersion-corrected DFT (PBE/ D3, 400 eV cutoff) using slab models that allowed top-layer relaxation. For the electronic analysis, the r^2SCANh hybrid functional was employed for projected density of states (PDOS) calculations, while sulfur 2p X-ray absorption spectra (XAS) were simulated using GPAW's real-space projector augmented wave method with ΔSCF core-hole treatment using PBE.

Bayesian exploration revealed that flat-lying adsorption is energetically favoured on all three coinage metals. The global minima correspond to HB238 oriented face-on with its π-system parallel to the surface, maximising dispersion and S–metal interactions. Side-on and edge-on structures were 1.5–3.5 eV less stable, due to loss of molecule-metal contact and weaker S/N–metal bonding. The MACE×BOSS workflow efficiently ranked structures and reproduced DFT energy ordering. Adsorption energies followed the trend Cu >Au >Ag. The donor (S1) and acceptor (S2) sulfur atoms prefer atop sites on Cu and Au, while on Ag the larger lattice constant forces S1 off-atop with only S2 near atop. Vertical distances range from 2.31–2.41 Å on Cu and 2.72–3.19 Å on Au/Ag, reflecting stronger Cu–S bonding. Nitrogen atoms interact weakly and exhibit larger lateral deviations; cyano nitrogens (N3/N4) lie closer to Cu (2.09 Å) than to Ag or Au.

Simulated sulfur L-edge XAS spectra show that the donor–acceptor splitting of S2p excitations is 0.44 eV in the gas phase and reduces to 0.29 eV on Cu(100), 0.26 eV on Ag(100) and 0.20 eV on Au(100). The reduction reflects substrate-mediated charge redistribution that partially neutralize the internal dipole. In monolayer simulations, intermolecular interactions further collapse the splitting, explaining experimental XPS spectra that show no resolvable donor–acceptor separation. PDOS revealed strong hybridisation between HB238 frontier orbitals and metal states. In the gas phase, HB238 has a HOMO at –3.41 eV and LUMO at –2.05 eV. Upon adsorption, these levels renormalize relative to the Fermi level (E_F). On Cu(100), the HOMO lies 1.39 eV below E_F and the LUMO is 1.98 eV above E_F, yielding a 3.37 eV gap. Au and Ag shift the HOMO closer to E_F and the LUMO further away (gap 3.97eV on Au and 4.13 eV on Ag).

The study demonstrates that flat-lying adsorption is preferred on all 3 metal surfaces and Cu provides the strongest binding and greatest spectral renormalisation, while Ag and Au offer weaker coupling. Surface screening drastically reduces the donor–acceptor splitting in sulfur XAS. Further studies on the systemetic generation of monolayers employing the MACE×BOSS approach can provide insights directly for tuning interfacial dipoles, work functions and charge transfer barriers in organic optoelectronic devices.

CHAPTER 7

Conclusions and Outlook

Organic optoelectronics[1–3] has advanced from a niche research area to a central component of modern materials science,[191–193] enabling innovations in flexible displays,[4] solar cells,[194] and molecular-scale information processing.[195, 196] The success of these technologies relies on the ability to tailor the electronic and optical responses of organic semiconductors through molecular design. Among these, merocyanines are donor–acceptor conjugated molecules characterized by strong intramolecular charge transfer and have emerged as exemplary materials because of their large dipole moments, strong nonlinear optical responses, and environment-dependent absorption spectra[48, 49, 197]. Their structural tunability and pronounced polar character make them attractive candidates for next-generation optoelectronic devices, including organic photovoltaics, nonlinear optical switches, photodetectors, and field-effect transistors[34, 55, 59]. Despite extensive experimental studies, predicting how these flexible dipolar molecules behave at solid interfaces remains a significant theoretical challenge. Molecular ordering, adsorption geometry, and interfacial charge redistribution strongly affect device performance, yet their accurate description requires methods that can bridge quantum-level accuracy with mesoscale structural complexity. This thesis addressed these challenges by integrating first-principles simulations and machine-learning–accelerated structure search to investigate the optoelectronic and adsorption behavior of the prototypical merocyanine HB238 across a hierarchy of systems—from isolated molecules to ordered monolayers and complex organic/inorganic interfaces.

In Chapter 3, a comprehensive theoretical investigation was conducted on the optoelectronic properties of merocyanine dyes in gas phase and in solution, focusing on benchmarking computational methods for accurately predicting excitation energies. The study compared TD-DFT, GW-BSE, STEOM-DLPNO-CCSD and CASSCF/NEVPT2. The results demonstrated that while TD-DFT tends to overestimate excitation energies, the GW-BSE approach provides excellent agreement with experimental spectra. Solvent induced polarization and structural relaxation were shown to significantly influence optical transitions, highlighting the importance of environmental effects. The TDCP-B3LYP approach,[159] adapted here for donor–acceptor systems, reproduced key spectral features with near-GW accuracy at much lower computational cost. In summary, these results established a reliable computational framework for studying excitation energy of merocyanines and provided the foundation for subsequent surface and interface investigations.[163]

The focus of Chapter 4 lies on the theoretical framework to model the self-assembly and adsorption of HB238 on the Ag(100) surface using periodic density functional theory. The global minimum

configuration revealed a flat-lying molecular orientation, forming a chiral tetramer arrangement in agreement with scanning tunneling microscopy observations.[82, 198] The monolayer structure was stabilized through a balance of weak chemisorption to the Ag substrate and intermolecular interaction. Electronic structure analysis indicated partial charge transfer from the substrate to the molecule, resulting in a filled LUMO just below the Fermi level—consistent with UPS data. The DFT-optimized geometry reproduced the experimental adsorption height and electronic level alignment. This chapter demonstrated the use of DFT-based modeling can accurately describe the adsorption geometry and interfacial electronic coupling in complex organic monolayers.

In Chapter 5, the adsorption of HB238 on hBN was studied using a machine-learning–accelerated global optimization workflow that combined BOSS[92] with MACE[101] potential denoted as MACE×BOSS . This framework enabled efficient exploration of the high-dimensional configurational space associated with flexible molecular adsorption. The global search revealed that HB238 adsorbs on hBN in a flat configuration with multiple close-lying local minima, reflecting the weak and non-site-specific nature of molecule–surface interactions. These findings show that while metallic substrates have strong templating effects, hBN provides a largely inert platform where intermolecular interactions dominate over substrate interactions. The success of the MACE×BOSS approach demonstrated that machine learning reduces the computational cost of global structure prediction while maintaining DFT-level accuracy, offering a scalable strategy for studying large organic adsorbates on weakly interacting surfaces.

In Chapter 6, the MACE×BOSS methodology was applied to study adsorption of an isolated HB238 molecule on the (100) surfaces of coinage metals (Cu, Ag, and Au) to elucidate how substrate geometry and electronic structure influences binding geometry. On Cu(100), the adsorption was predominantly chemisorptive, characterized by strong Cu–S bonding. On Ag(100) and Au(100), the interaction was primarily physisorptive, leading to weaker adsorption energies and greater orientational flexibility. Analysis of projected density of states revealed systematic variations energy level alignment across the three metals, highlighting how substrate reactivity governs the balance between molecule–substrate and intermolecular interactions.

Building on the above findings, several future developments are envisioned to extend the theoretical and computational framework for organic–inorganic interfaces. The MACE × BOSS methodology can be generalized to other complex adsorbate–surface systems. Future work should explore its applicability to diverse molecule/surface combinations, including different classes of organic semiconductors and 2D materials, to test the limits of the surrogate model's transferability. One promising direction is to adopt an active learning loop where the ML potential identifies new, informative configurations (for example, unusual binding modes or transition states) that can be confirmed with targeted DFT calculations and then added to the training data. Such iterative refinement would improve robustness and reliability, ensuring the surrogate does not extrapolate beyond its validated domain.

In order to achieve multiscale modeling, the approach should be extended to capture finite-temperature effects and dynamical processes. While the current work focused on 0 K equilibrium structures (global minima and near-minima), real interfaces often involve entropy, kinetics, and metastable states. Integrating the ML potential into molecular dynamics (MD) or Monte Carlo simulations could enable the study of adsorbate diffusion and rotation on surfaces over longer time scales. ML-driven MD can be used to simulate migration pathways of HB238 molecules on hBN at experimental deposition temperatures and provide insight into nucleation mechanisms that may lead to ordered monolayers. The fast speed of ML potentials makes it feasible to simulate larger systems and longer time spans than with ab-initio MD. Additionally, using techniques like enhanced sampling

or metadynamics with the ML potential can be used to study transition pathways between different adsorption configurations and quantify energy barriers for surface diffusion or assembly.

Future theoretical studies should aim to link the atomistic modeling with macroscopic observables relevant to organic electronics. For example, having identified the stable interface structures, their electronic and optical properties using high-level methods (as benchmarked in Chapter 3) can be calculated to predict how the substrate influences excitation energies, charge transfer states, or charge transport at the interface. Combining the ground-state structural models with excited-state calculations (TD-DFT or GW-BSE on the adsorbed molecules) would yield predictions of interfacial spectral shifts or quenching/enhancement of optical absorption – directly testable against experiments. Likewise, evaluating charge injection barriers or band alignment for the computed interface geometries would inform device modeling (e.g. estimating how an hBN layer modifies the energy level alignment between a merocyanine film and an electrode).

Future improvements to ML interatomic potentials could focus on more accurate treatment of dispersion and long-range induction effects. Incorporating classical electrostatic descriptors or many-body dispersion schemes directly into the machine learning framework would significantly improve the description of weak intermolecular and molecule–surface interactions, thereby enhancing accuracy for physisorbed systems. On the Bayesian optimization side, scaling up to even higher-dimensional searches (e.g. considering flexible substrate lattices or multi-component co-adsorption) is a challenge. Future research could incorporate more advanced global optimization algorithms or coarse-grained models for initial screening. Cross-validation with experimental data will remain crucial, as new interface systems are studied, theoretical predictions should be continuously benchmarked against microscopic structural probes (STM, X-ray diffraction) and spectroscopy (UPS, XPS, etc.) to ensure the models capture reality. As the models mature, an important future prospect lies in predicting and designing ordered molecular monolayers directly from first principles—capturing the balance between molecule–substrate interactions and self-assembly forces that govern large-scale pattern formation.

APPENDIX A

Theoretical Study on the Optoelectronic Properties of Merocyanine Dyes

Ritu Tomar[†], Leonardo Bernasconi[‡], Daniele Fazzi[§], Thomas Bredow[†]

Received: 23 June 2023
Published online: 14 November 2023

Adapted with permission[¶] from
R. Tomar, L. Bernasconi, D. Fazzi, and T. Bredow, *Theoretical Study on the Optoelectronic Properties of Merocyanine Dyes*, *J. Phys. Chem. A* **127** (2023) 9661-9671.
DOI:10.1021/acs.jpca.3c04226

Own Contributions

- performing of all calculations except coupled cluster and CASSCF
- interpretation of the results
- writing the manuscript

[§]

[¶] Permission requests to reuse material from this chapter should be directed to the American Chemical Society
[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[‡] Center for Research Computing and Department of Chemistry, University of Pittsburgh, PA 15260, USA
[§] Dipartimento di Chimica "Giacomo Ciamician", Universitá di Bologna, 40126 Bologna, Italy

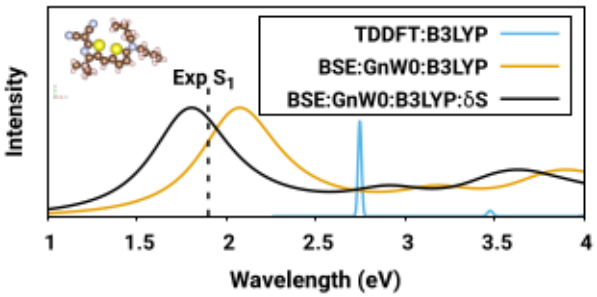


Figure A.1: Table of content graphic (ToC)

Abstract Merocyanines, as prototypes of highly polar π-conjugated molecules, have been intensively investigated for their self-assembly and optoelectronic properties, both experimentally and theoretically. However, an accurate description of their structural and electronic properties remains challenging for quantum-chemical methods. We assessed several theoretical approaches, TD-DFT, GW-BSE, STEOM-DLPNO-CCSD, and CASSCF/NEVPT2-FIC for their reliability to reproduce optoelectronic properties of a series of donor/acceptor (D/A) merocyanines, focusing on the first excitation energy. Additionally, we tested an all-electron perturbative method based on time-dependent coupled-perturbed density functional theory, denoted as TDCP-DFT. Particular focus was set on direct and indirect solvent effects which affect excited-state energies by electrostatic interaction and molecular geometry. The molecular configuration space was sampled at the semiempirical tight-binding level. Our results corroborate previous investigations, showing that the S_0-S_1 excitation energy strongly depends on the merocyanine molecular structure and the dielectric constant of the solvent. We found significant effects of the polar solution environment on the geometry of the merocyanines, which strongly affects the calculated excitation energies. Taking these effects into account best agreement between calculated and measured excitation energies was obtained with TDCP-DFT and GW-BSE. We also calculated excitation energies of molecular crystals at TDCP-DFT level, and compared the results to the corresponding monomers.

A.1 Introduction

Merocyanines (MCs) are excellent candidates for optoelectronic applications because their optical properties can be easily tailored using chemical and physical methods. MCs have been intensively investigated, both experimentally and theoretically, for over four decades and have numerous applications ranging from textile colorants to high-tech opto-electronic devices. [42, 199–201] Pioneering work by Marder[202], Blanchard-Desce[203], Meerholz[204], and Würthner et al.[47] paved the way for their application in nonlinear optics[200, 205], photorefractivity[201, 204, 206], and solar cells (OSC)[35]. The donor and acceptor groups present in MCs are linked by ethylene or polyethylene bridges. As a result, a crucial structural characteristic of MCs is the so-called bond length alternation (BLA) defined as a measure of the difference between single and double carbon-carbon bonds of the polyethylene bridge. Small (large) BLA indicate increased (decreased) electron delocalization and strong (weak) donor-acceptor coupling, resulting in a red (blue) shift in the spectra.[207] The BLA parameter measures the electron delocalization within a molecule, with smaller BLA values generally

indicating lower excitation energies. Conversely, larger BLA values signify reduced delocalization and higher excitation energies, resulting in absorption at higher energy levels. A small or vanishing BLA denotes an electronic structure that is close to the cyanine form, whereas a large positive or negative BLA denotes an increase in the weight of the neutral or the zwitterionic form, respectively.[208] From the theoretical point of view, MC BLA is not accurately described using standard DFT based on the generalized gradient approximation. [209–211] In this work, we report the importance of optimizing the molecular structure of MCs with double-hybrid density functionals in a polar environment to accurately capture the measured BLA pattern of merocyanines.

The quantitative prediction and comprehension of low-lying excitations in MC molecules are both fundamentally interesting and technologically important. For instance, the calculation of excitation energies provides valuable insights into the photophysical properties of merocyanines[34, 35, 212], including absorption spectra[207, 213], fluorescence emission [214], and photochemical behavior[211, 215]. A standard approach to calculate molecular excitation spectra is time-dependent density-functional theory (TD-DFT) based on various density-functional approximations (DFA). In previous studies it was found that the first excitation energy is overestimated by about 0.3 eV with TD-DFT, even with fourth-rung DFA such as CAM-B3LYP[203]. Here, we assess the performance of TD-DFT based on DFAs up to the fifth rung to reproduce the experimental first singlet excitation energies of a selection of MCs with different donor/acceptor functional groups, denoted as **D1A1**, **D1A2**, **D1A3**, **D2A4** (see Figure A.2). Solvent effects were taken into account using the conductor-like screening polarizable continuum model (CPCM) [183]. We compared the TD-DFT results with higher-level methods such as the Green function-based Bethe-Salpeter Equation (GW-BSE) [216], the complete active space self-consistent field (CASSCF)[217] and the similarity transformed equation of motion coupled cluster theory (STEOM-DLPNO-CCSD) [218]. Additionally, the time-dependent coupled-perturbed DFT (TDCP-DFT) method [153] was applied. This method has been demonstrated to provide highly accurate excitation energies for inorganic solids [160, 161, 219], and in this work it is applied to molecular systems for the first time. We analyze the electrostatic interaction between merocyanine molecules in the crystalline state by aligning the transition dipole vectors. This method yields deeper insights into the aggregation behavior that influences the optical absorption properties of the merocyanines. Additionally, we analyze distances between the barycenters of adjacent molecules to infer intermolecular coupling, providing a new perspective on the molecular interactions within these systems.

A.1.1 Methods

All molecular structure optimizations and TD-DFT calculations were performed with ORCA version 5.0 [145]. The geometry of the MC molecules was optimized with the DFAs mentioned below, def2-TZVP basis sets [147] and Grimme's D3 dispersion correction [132]. An initial conformer search was performed with CREST [137] combined with the semiempirical GFN1-xTB method [134]. The most stable conformers were then optimized in gas-phase and in solution applying the CPCM [183] solvation model in ORCA.[145] All TDDFT calculations are performed with non-equilibrium and linear response PCM regime. The solvent shift of the first excitation energy $\delta S = S_1^{sol} - S_1^{gas}$ is calculated at TD-DFT level and applied to the higher-level methods where the CPCM cannot be applied.

We employed the following density functional approximations (DFA): PBE[220], PBE0[221], B3LYP[146, 222–224], CAM-B3LYP[225],wB97X[226], B2PLYP [227], and basis sets def2-TZVP,

def2-TZVPP and def2-QVZPP [228] for the TD-DFT calculations. In order to investigate the influence of electron correlation, STEOM-DLPNO-CCSD[218] calculations were performed on DFT-optimized geometries using ORCA version 4.2.0.[229] We utilized two different basis sets (namely def2-SVP and def2-TZVP), and considered up to five singlet states in our calculations, employing standard threshold parameters, including DTol, STol, OThresh, and VThresh, as specified in the ORCA manual. GW-BSE calculations of the optimized MC structures were performed with the MOLGW code [230] applying the G_nW_0 approximation based on B3LYP [222, 223] wavefunctions obtained with Dunning correlation-consistent polarized triple-zeta (cc-pVTZ) basis sets [231]. We observed that quasiparticle energies converged within a few meV within six iterations. The GW approach takes into account electron correlation and polarization effects around holes and electrons, and provides physically meaningful one-electron energy levels. By solving the Bethe-Salpeter equations (BSE) electron-hole interaction is included for excited states. This approach has been proven to provide reliable results for excitation energies of organic molecules.[232]

TDCP-DFT [138, 157], also based on B3LYP wavefunctions, was applied as implemented in a locally modified version of CRYSTAL14 [153, 159, 233]. We used optimized BSSE-corrected triple-zeta plus polarization basis sets (pob-TZVP-rev2). [234] The poles of the dynamical polarizability tensor are calculated to approximate the optical transition energies. The TDCP-DFT method only provides dipole allowed transitions and does not account for quadrupole or dipole magnetic effects. In addition, because we used a purely adiabatic approximation for the non-local hybrid kernel, this approach can only determine the fundamental mode of the exciton series. Further, we calculated the dielectric constant of the molecular crystals using the CPHF method in CRYSTAL.[235, 236]

Fractional occupation number weighted electron density (FOD) calculations were conducted using ORCA to assess the partial occupation of frontier orbitals and determine the FOD number (N_FOD).[237] The FOD number serves as a molecular parameter indicating the presence of "hot" electrons, specifically correlated electrons within the system. Default Density Functional Approximations (DFA) as implemented in ORCA version 4.2.0 were employed for these FOD calculations.

Furthermore, FOD calculations were utilized as an initial screening tool to determine the active space for further Complete Active Space Self-Consistent Field (CASSCF) calculations.[238] For the MCs under investigation, various active spaces were considered, spanning from (4,4) to (12,12). Unless otherwise stated, the chosen active space is (12,12). To refine the energy determination of singlet excited states, second-order N-electron valence perturbation theory (NEVPT2) corrections with fully internally contracted (FIC) wavefunctions were incorporated using the CASSCF/NEVPT2-FIC method, which is available in ORCA version 4.2.0.[229]

Unless stated otherwise, the B2PLYP optimized geometries were used for all calculations to achieve consistency. Only the first singlet excited state S_1 was considered for each chromophore, as the higher excited states are outside the visible region of the absorption spectrum (TDDFT artifact), which were not the focus of this work. We compare the computed excited state S_1 of the investigated MCs to experimental absorption maxima obtained from the UV-Vis adsorption measurements.[199, 239–241]

A.2 Results and Discussion

The structures of the donor (D) and acceptor (A) units investigated in this study are shown in Figure A.2. The four reference systems studied here are classified according to the D/A combin-

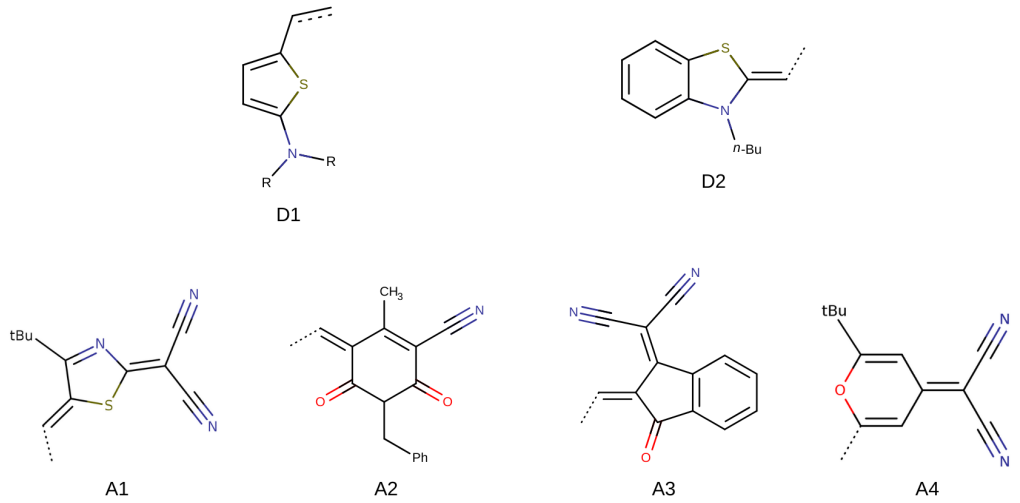


Figure A.2: Molecular structures of the donor (D) and acceptor (A) units constituting the merocyanines investigated. R can be either n-butyl(n-Bu) or ethyl(Et) group.

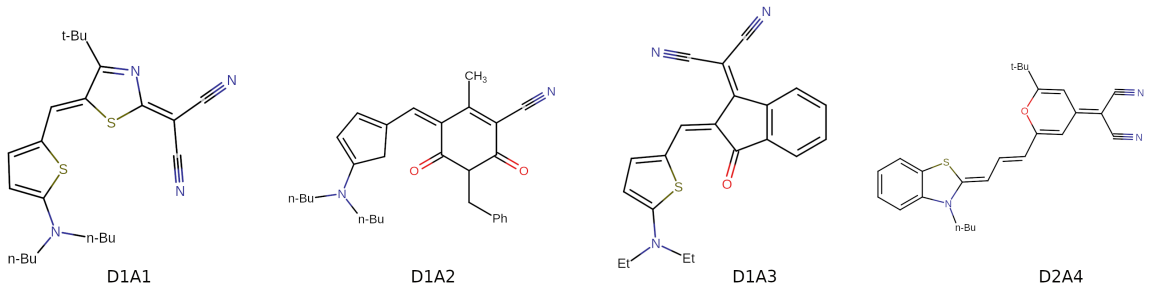


Figure A.3: Molecular structures of the investigated merocyanines.

ations that define the π-conjugated backbone: **D1A1**, **D1A2**, **D1A3**, and **D2A4** (See Figure A.3). The following donor and acceptor groups were selected as representatives of a wide range of MCs, and since experimental data are available for structure and excitation energies: **D1**=5-ethenyl N,N-disubstituted thiophenamine, **D2**=(2Z)-butyl-ethylidene 2,3-dihydro-1,3-benzothiazole ('Fischer base'), **A1**=2-[(5Z)-4-tert-butyl-5-ethylidene-1,3-thiazol-2-ylidene]propanedinitrile, **A2**=(3Z)-5-benzyl-3-ethylidene-2-methyl-4,6-dioxocyclohex-1-ene-1-carbonitrile, **A3**=2-[(2Z)-2-ethylidene-3-oxo-2,3-dihydro-1H-inden-1-yl]propanedinitrile, **A4**=2-(2-tert-butyl-6-methyl-4H-pyran-4-ylidene) propanedinitrile.

First we discuss the influence of the DFA and the solvent effect on the structures of **D1A1** and **D1A2** with an emphasis on the BLA.

Geometry optimizations

The accurate description of the ground-state geometry of MCs is challenging for standard DFAs [242]. This is due to the presence of two resonance structures, representing a neutral and a zwitterionic form, which require a careful description of static and dynamic correlation effects. Since geometry optimization with more sophisticated methods is cumbersome for screening purposes, standard DFAs have been nonetheless applied to calculate MC geometries. The BLA, as a simple measure of the relative weights of the neutral and zwitterionic forms[208], has been used to assess the reliability of the theoretical approaches. d_{BLA} is defined as the difference between the average of single and double bond lengths R $\left(d_{\mathrm{BLA}} = \sum_i^N (R^i_{\mathrm{single}})/N - \sum_j^M (R^j_{\mathrm{double}})/M\right)$ where N and M are the numbers of single and double bonds, respectively. In general, the bonds connecting the donor and acceptor units are used for BLA calculation, however, we consider the alternating single and double bonds across the π-conjugated backbone of the molecules as suggested by Gildemeister et al.[242] (see also Figure S1 in the Supporting Information). For the sample chromophores of our selection, we compared the BLA as calculated with various DFAs in the gas phase with experimental data obtained from X-ray diffraction (XRD) given in Table A.1.[239–242]

Table A.1: Mean absolute deviation (MAD) with respect to experimental structures of d_{BLA} (Å) calculated in the gas phase with various DFAs including PBE, ωB97X, CAM-B3LYP, B3LYP, and B2PLYP with respect to experimental XRD data. All calculations were performed with Grimme's D3 dispersion correction and def2-TZVP basis set.

Merocyanine	PBE0	ωB97X	CAM-B3LYP	B3LYP	B2PLYP	XRD
D1A1	−0.024	−0.049	−0.039	−0.024	−0.024	−0.001[a]
D1A2	−0.007	−0.024	−0.017	−0.008	−0.008	0.016[b]
D1A3	−0.012	−0.026	−0.019	−0.027	−0.012	0.011[c]
D2A4	0.047	0.071	0.062	0.048	0.048	0.047[d]
MAD	0.018	0.037	0.029	0.022	0.018	

[a]:Reference[199]; [b]:Reference [239]; [c]:Reference[240]; [d]:Reference [241];

Among the considered DFAs, the dispersion-corrected double-hybrid functional B2PLYP-D3 and PBE0-D3 provide best agreement with the experiment, also the global hybrid functional B3LYP-D3 combined with def2-TZVP basis sets provide results reasonably close to the aforementioned functionals. Nonetheless, both double hybrid and global hybrid functionals significantly underestimate the experimental BLA and provide the wrong sign for **D1A2** and **D1A3**. Possible reasons will be discussed below.

Since UV-Vis absorption spectra are measured in solution, we focus on the effects of solvation on the MC structure in the next section. Previously it was shown that these effects can be effectively addressed using a polarizable continuum method (PCM), by considering highly polar solvents, such as acetone and DMSO [212], or by using a constrained-DFT approach (CDFT) [242]. For this purpose, we re-optimized the structures by applying the CPCM with the solvent acetone (in order to show the solvent effect of a polar solvent) and the respective solvents as used in the UV-Vis experiments, dichloromethane (DCM) for D1A1 and D1A3, 1,4-dioxane for D1A2, and dimethylsulfoxide (DMSO)

for D2A4. The results are shown in Table A.2 and Figure A.4.

Table A.2: Mean absolute deviation (MAD) of $d_{BLA}()$ calculated with B2PLYP-D3 and B3LYP-D3 for the MC geometries as optimized in respective solvents.

Merocyanine	Solvent	B2PLYP-D3	B3LYP-D3
D1A1	Acetone	0.013	0.014
	DCM	0.008	0.009
D1A2	Acetone	0.019	0.023
	1,4-dioxane	0.002	0.003
D1A3	Acetone	0.010	0.013
	DCM	0.008	0.010
D2A4	Acetone	0.025	0.022
	DMSO	0.023	0.021
MAD	Acetone	0.010	0.012

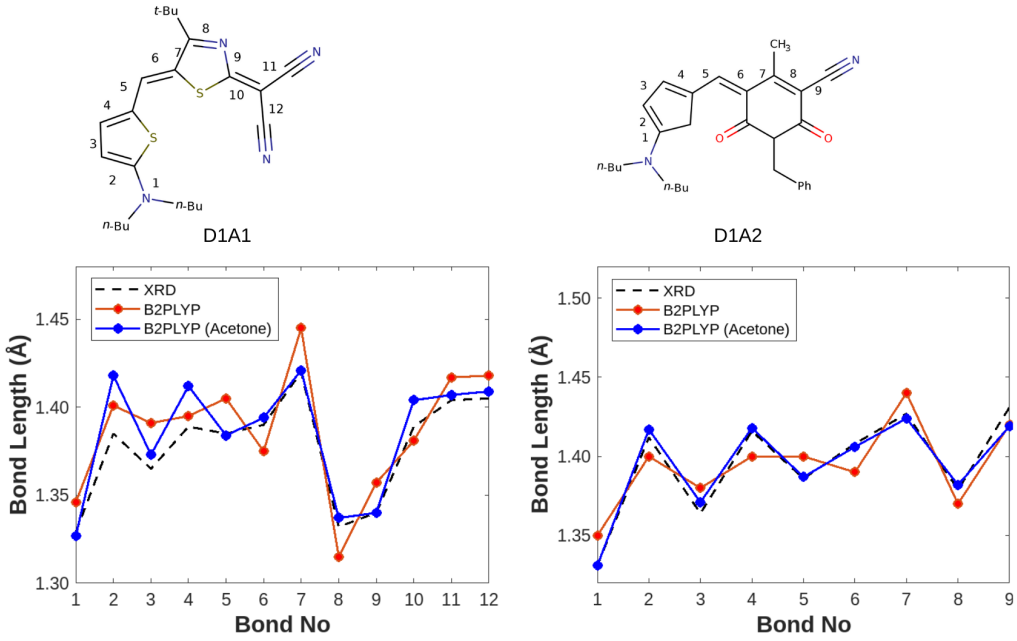


Figure A.4: Comparison of calculated and measured BLA of **D1A1** and **D1A2**; black dotted lines: XRD [199], red solid lines: B2PLYP (gas phase) and blue solid lines: B2PLYP/CPCM (acetone).

A.2.1 Solvent effects on Merocyanine Geometry

The solvatochromic shifts of merocyanine (MC) molecules have been extensively investigated, and we here report that the shifts are dependent on the geometry, namely the BLA of the molecule. The

solvent polarity changes the equilibrium between the neutral and zwitterionic states, shifting the structure towards one of the two forms.

In a previous study, CDFT has been used to distinguish between the two electronic situations.[242] Here we show that the effects of solvent polarity on the geometry of MCs are crucial. Solvent effects steer the molecule's behavior towards one of two extreme states of cyanine compounds: neutral or zwitterionic. When optimizing the geometry of a merocyanine molecule in a polar solvent like acetone, the BLA parameter's sign changes. This shift reflects a transformation in the molecule's arrangement from a zwitterionic state ($d_{BLA} < 0$) to neutral form ($d_{BLA} > 0$) for MC's D1A1, D1A2 and D1A3. For the structures optimized in acetone, the MAD of the calculated BLA is significantly reduced (from 0.018 to 0.010 for B2PLYP and 0.022 to 0.012 for B3LYP, Table A.2). MAD refers to the average of the absolute deviations of the (d_{BLA}) calculated using various density functionals from the experimental XRD data. This indicates that the solvent approximates the electrostatic intermolecular interactions in the bulk which affects the molecular structure. Agreement with the experimental structures is slightly better with B2PLYP-D3 than with B3LYP-D3. Therefore, all subsequent electronic and optical properties calculations were performed on the B2PLYP-D3 optimized structures.

The solvent effects on the excitation energy of MC **D2A4** have been reported experimentally.[241] Therefore, we calculated the geometry dependence of **D2A4** on the polarity of the solvent (Figure S3). In agreement with previous theoretical studies [203, 241, 242] there is a strong dependence of the calculated BLA on the dielectric constant of the solvent. For comparison, we also computed the first excitation energy of molecular crystals of D2A4. The calculation of the dielectric constant of the merocyanine crystals involves computing the components of the dielectric tensor ϵ_{tu}, which can be obtained from the polarizability tensor α_{tu}components as per the equation:

$$\epsilon_{tu} = \delta_{tu} + \frac{\alpha_{tu}}{\varepsilon_0 V}$$

where δ_{tu} is Kronecker's delta, ε_0 is the permittivity of vacuum, and V is the unit cell volume. The term $\alpha_{tu}/\varepsilon_0 V$ represents the first dielectric susceptibility $\chi^{(1)}_{tu}$, which is proportional to the linear coefficient of the derivative of the induced polarization vector $P_t(\mathcal{E})$, as given by the equation:

$$P_t(\mathcal{E}) = \frac{1}{V}[\mu_t(\mathcal{E}) - \mu_t(0)]$$

The dielectric constant is then obtained from the dielectric tensor by taking its trace and computing it in the limit as the frequency (ω) approaches zero, as per the Time-Dependent Coupled Perturbed (TD-CP) equations. This approach is implemented in the CRYSTAL software, which we used to compute the dielectric constant of the merocyanine crystals.[233, 243]. The best agreement between theoretical and experimental structure is obtained when the dielectric constant of the solvent is close to that of the merocyanine crystal. We calculated the dielectric constant of the merocyanine crystals using TDCP-B3LYP method and pob-TZVP-rev2 basis-set. This finding highlights the importance of taking into account solvent effects not only in the calculation of the excited-state wavefunction, but also on the MC geometry when studying the optoelectronic properties.

A.2.2 Method dependence of optoelectronic properties of D1A1

In the following, we assess the accuracy of various DFAs (PBE, PBE0, B3LYP, CAM-B3LYP, and B2PLYP) for the calculation of electronic and optical properties of the MC **D1A1**. The highest

occupied molecular orbital (HOMO) energy and the first singlet excitation energy S_1 were considered. While the HOMO energy can be accurately measured experimentally and is directly related to the ionization potential, the experimental determination of the LUMO energy is more challenging and often involves approximations.[244, 245] Therefore, we focused our analysis on the HOMO energy to provide a more reliable assessment of the accuracy of the calculated electronic and optical properties.

Table A.3: Optoelectronic properties of **D1A1**: HOMO energy and singlet $S_0 - S_1$ vertical energy transition in the gas phase in eV. All DFT calculations were performed using the def2/TZVP basis set, G_6W_0 calculations were performed on B3LYP wavefunctions obtained with cc-pVTZ basis sets. All TDDFT calculations were performed on the geometries optimized with B2PLYP and B3LYP geometry was used for G_6W_0 calculations.

Method	PBE	PBE0	B3LYP	CAM-B3LYP	B2PLYP	G_6W_0	Exp
HOMO (eV)	−4.88	−5.65	−5.40	−6.64	−6.05	−6.59	−5.52[a]
$S_0 - S_1$ (gas phase)	2.75	2.87	2.84	2.92	2.46	2.05	—

[a]:Reference [199]

The rung four global hybrid functionals PBE0 and B3LYP give a good account of the measured HOMO energy with deviations of ∼0.1 eV (Table A.3). As expected, PBE provides a too-high value due to the self-interaction error, while CAM-B3LYP strongly underestimates the experimental reference value. We performed G_nW_n calculations at different levels including G_0W_0, G_6W_0 and G_6W_6, cc-pVDZ and cc-pVTZ basis sets, and found that the present results are converged in terms of GW self-consistency, but not with the basis set size (Table S2 in SI).Due to the limitations of our computer resources, it was not possible to employ cc-pVQZ basis sets. Surprisingly, the G_6W_0 HOMO energies, which should be the best representation of ionization potentials, are more than 1 eV too low. At the moment we have no explanation for this discrepancy.

All TD-DFT methods strongly overestimate the vertical excitation energy, by 0.7-0.9 eV with respect to the experimental λ_{max} energy (1.90 eV as measured in DCM).[199] We verified by a systematic increase of the basis set size from def2-TZVP to def2-TZVPP and def2-QVZPP (see Table S1 in the Supporting Information), that this is not a basis set incompleteness error. Such an overestimation has been observed previously for similar systems.[246, 247] As we compare the computed vertical excitation energy S_1 of the investigated MCs to experimental absorption maxima, it can lead to an overestimation of computed energies. In order to simulate the band shape and the origin of the spectra accurately nuclear effects (e.g., vibronic effects and Wigner ground state sampling) can be included.[248, 249] It is important to note that the neglect of vibronic effects in our calculations can lead to a systematic error, i.e. ∼ 0.1eV in the predicted excitation energies as suggested by Fang et al. [250] and Santoro et al.[248] MCs have been shown to exhibit large solvatochromic effects, with both hypsochromic and bathochromic shifts measured.[241] The effect of structural changes due to solvent effects will be discussed later. First, we consider the possibility that the discrepancies between experiment and theory are due to the conformational flexibility of MCs, which requires sampling of spectra for many configurations. Previous studies by Tirri et al., Georg et al. and Coutinho et al. report improvements in the description of excitation spectra of merocyanines using Monte Carlo sampling of molecular configurations.[207, 251, 252] We tested the conformational sampling tool CREST and performed sampling of MC geometries in gas phase as well as solvent phase using iMTD-sMTD

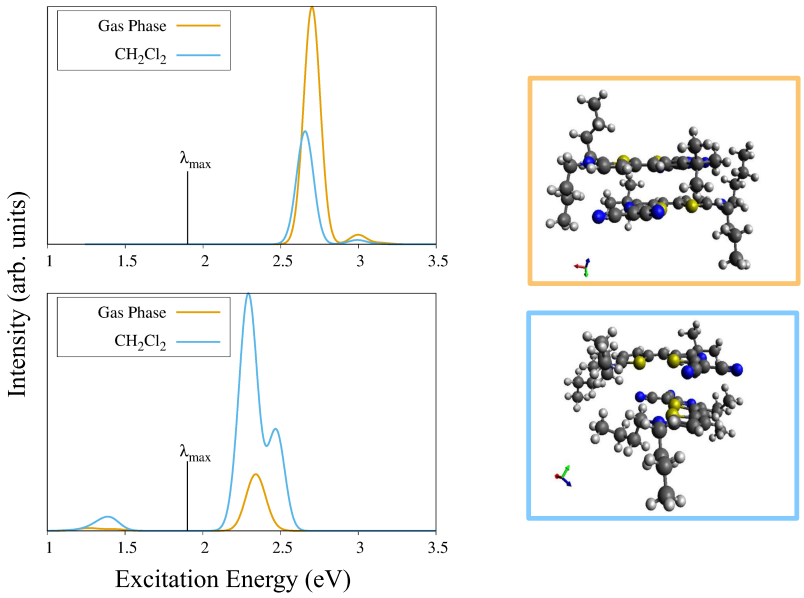


Figure A.5: Boltzmann weighted absorption spectra of 356 most stable **D1A1** monomer conformations(top left) and 197 most stable dimer configurations (bottom left) as obtained via CREST, calculated using TD-DFT at B3LYP/def2-TZVP level. The experimental λ_{max} value is shown by the black vertical line at 1.9 eV.[199]. The most stable conformers of D1A1 dimers in gas phase (top right) and DCM solvent (bottom right) are reported.

algorithm and GFN1-xTB Hamiltonian. A conformer study of **D1A1** is reported in the next section.

Conformers of D1A1

In the neutral ground state, **D1A1** has a planar π-conjugated backbone. However, the t-butyl groups are rather flexible and free to rotate. Since the relative orientation of these groups may affect the optical excitation spectrum in terms of band broadening and band shift,we calculated (at TD-DFT level) the excitation energy of 356 most stable conformers of **D1A1** obtained by CREST[137] (Figure A.5). However, the broadening effect due to conformation sampling is quite small (~0.04 eV in gas phase and 0.10 eV in DCM solution). The TD-DFT S_1 excitation energy ranges from 2.68 to 2.72 eV in the gas phase and from 2.59 to 2.69 eV in the solvent (CPCM/DCM). These values suggest a systematic error of ~ 0.1eV in the computed excitation energies using TDDFT methods. However, these excitation energies are well above 1.90 eV (λ_{max}) obtained from UV-Vis experiment.[199]

In the present study, we performed a comprehensive analysis by investigating both the monomeric and dimeric forms of merocyanines and comparing them with the available experimental spectra. This approach is crucial because it helps to understand the aggregation behavior of merocyanines in solution. We also explored the possibility of red-shift in the excitation energy going from the MC monomers to the dimers, all due to mutual polarization as suggested by Tirri et al.[207] We used the iMTD-sMTD method as implemented in CREST to perform conformational sampling on two **D1A1** molecules positioned side-by-side. In the gas phase, the **D1A1** molecules form stacked antiparallel dimers, whereas in dichloromethane they are oriented orthogonal to each other (Figure A.5). This

shows the significant influence of the solvent on the structural characteristics of MCs. The excitation energies of the 197 lowest-energy dimer configurations in gas phase and dichloromethane solvent are shown in Figure A.5. In the dimer configuration, S_1 energy varies based on the specific geometry and electronic structure of the dimers. This is evident from the transitions observed around 1.5 eV, indicating that S_1 can be active or dark depending on the dimer's geometry. In the case of the antiparallel configuration (H-like dimer stacking), the higher energy states (S_2 and S_3) gain oscillator strength, while S_1 becomes dark or weakly allowed. In monomer spectra, the weak peaks observed at 2.9-3.0 eV represent transitions to a higher excited state (S_2) of the molecule (see Figure A.5).

A.2.3 Excitation energies in gas-phase and in solution

We examined different levels of theory, including TD-DFT, GW-BSE, STEOM-DLPNO-CCSD, CAS-SCF, and TDCP-DFT, in search of an efficient and accurate computational approach for characterizing the first excitation energy of MCs. It has been shown before, that TD-DFT significantly overestimates the excitation energies of merocyanines [203, 213], which was attributed to the lack of differential correlation energy in DFT. Double-hybrid functionals (e.g. B2PLYP) do not fully alleviate this problem, although they improve the description of excited states in general. [227, 253] The treatment of electron correlation effects using coupled cluster (CC) theory,[254] or the Bethe-Salpeter approach based on Green function with screened Coulomb interaction (GW-BSE) [157, 255, 256] approaches provided accurate excitation energies, however at high computational cost.

In the case of the MC's examined in this study, N_FOD values ranged from 0.5 to 0.7, indicative of a single-reference character for describing the electronic structure of the ground state. We further used the N_FOD analysis for active space determination in CASSCF/NEVPT2-FIC calculations. Active spaces (12,12) are used for MC D1A1 and D1A3 and (4,4) for D1A2.

TD-DFT is computationally efficient and allows to take solvent effects into account via continuum models such as CPCM. This is presently not possible with the more sophisticated methods. We distinguished between two kinds of solvent effects: (I) excitation energy shifts for a given structure, and (II) structural changes. Both solvent shifts (δS) were calculated with B2PLYP/def2-TZVP at TDDFT level(Table A.4) and added to the GW-BSE, STEOM-DLPNO-CCSD, CASSCF/NEVPT2, and TDCP-DFT gas-phase results (Table A.5) in Table A.6.

In the first series of calculations, we used the same geometries of the MCs in gas-phase and solvent (effect I). Table A.5 shows the gas-phase results, in Table A.6 the S_1 energies obtained for the B2PLYP gas-phase structures are shifted by δS calculated with TD-B2PLYP.

Table A.4: S_0-S_1 energies and solvent shift $\delta S = E_{ex}^{sol} - E_{ex}^{gas}$ (eV) for the selected MCs calculated using B2PLYP/def2-TZVP and CPCM in ORCA. Fixed gas-phase geometries obtained with B2PLYP/def2-TZVP are used.

Merocyanine	Solvent	gas phase	solution	δS
D1A1	DCM	2.46	2.19	−0.27
D1A2	1,4-dioxane	2.93	2.66	−0.27
D1A3	DCM	2.59	2.27	−0.32
D2A4	DMSO	2.82	2.37	−0.45

By adding the solvent shifts, GW-BSE S_1 results are considerably improved, the MAD decreases from 0.34 eV to 0.09 eV. A similar trend is observed for TD-B2PLYP, however, the MAD is still much larger, 0.25 eV. STEOM-DLPNO-CCSD and CASSCF/NEVPT2 (MAD=0.36 and 0.18 eV) reproduce the experimental S_1 energies more accurately without solvent shift (MAD=0.11 and 0.14 eV, respectively). We noted that the STEOM-DLPNO-CCSD and CASSCF/NEVPT2 calculations give comparable results if a polar solvent-optimized geometry is considered for $S_0 - S_1$ computation. The excitation energy so obtained are 1.73 and 2.04 eV respectively for MC D1A1 with the geometry optimized using CPCM/Acetone. For the gas-phase optimized geometry of MCs, experimental S_1 energies are efficiently and closely described with TDCP-DFT. The MAD is not significantly changed after adding the solvent shifts (MAD=0.14 and 0.19 eV, respectively). The higher accuracy of TDCP-DFT S_1 energies compared to TD-DFT can be attributed to the better screening of electrons in the hole for the extended-π delocalization of MCs. In the present case, TDCP-DFT is a more efficient alternative to STEOM-DLPNO-CCSD and CASSCF/NEVPT2. Another advantage is that this method can also be applied to periodic calculations of molecular crystals (see next section), which is at present not possible with CCSD and CASSCF. GW-BSE calculations with VASP [**vasp3**, 128, 129] were not successful, neither for the isolated molecules nor for molecular crystals, due to limitations of our computational resources.

In the second step, we optimized the MC structures in the solution with B2PLYP (effect II). Since it is known that the molecular geometries are significantly affected by solvent effects,[212] we optimized the MC structures under conditions that are as close as possible to the experiment. For this purpose, we employed toluene ($\varepsilon = 2.4$) for **D1A1**, **D1A2**, **D1A3** and chloroform ($\varepsilon = 4.9$) for **D2A4**, respectively. The dielectric constants of these solvents are close to those of the molecular crystals, obtained with TDCP-DFT (see Table A.9). Indeed we obtained improved BLA compared to other solvents (see Table A.2) in this way. BLA indirectly reflects the degree of π-electron delocalization in a conjugated molecular system. Greater electron delocalization results in a more uniform molecular structure, such as equalized bond lengths, leading to smaller BLA values. A more delocalized electronic structure, however, is connected with a low energy gap between the occupied and unoccupied molecular orbitals (typically the HOMO and LUMO), thus leading to low values of the excitation energies (see Fig A.6). In this context, BLA can be associated with the electronic absorption spectrum of individual MC molecules. MCs showing large values of BLA result in single-molecule absorption spectra blue-shifted with respect to MCs with vanishing BLA. This confirms the high importance of the solvent effects on MC structure, which in turn affects optoelectronic properties reported in Table A.7.

For a second series of excited-state calculations, we therefore, used the MC geometries obtained by B2PLYP in the respective solvents of the solution experiments, DCM for **D1A1** and **D1A3**, 1,4-dioxane for **D1A2**, and DMSO for **D2A4**. After taking into account both the geometry change and the solvent shift in solution, Tables A.7 and A.8, GW-BSE again outperforms all other methods (MAD = 0.05 and 0.07 eV, respectively). Also, TD-DFT results are further improved while TDCP-DFT excitation energies are again not affected much.

A.2.4 Excitation energies in molecular crystals

We applied the TDCP-DFT method to calculate the excitation energy of merocyanines. For the calculation of the MC crystal's excitation energy and dielectric constant, the positions of hydrogen atoms were optimized starting from the XRD data for the complete unit cell, and the optimized unit cell structure was used for the TDCP-DFT calculations. For the gas phase calculations, the

Table A.5: S_0-S_1 energy (eV) for selected MCs calculated using TD-DFT (B2PLYP-D3), GW-BSE (B3LYP), STEOM-DLPNO-CCSD, CASSCF/NEVPT2, TDCP-DFT. MAD (eV) for individual methods compared to the experimental excitation energy (B2PLYP geometry and excitation energy calculations performed in gas phase.)

Merocyanine	TD-DFT	GW-BSE	STEOM-DLPNO-CCSD	CASSCF/NEVPT2-FIC	TDCP-DFT	Exp.
D1A1	2.46	2.05	1.87	2.03	1.90	1.90[a]
D1A2	2.93	2.62	2.47	2.26	2.37	2.31[b]
D1A3	2.59	2.37	2.01	2.38	2.31	2.15[c]
D2A4	2.82	2.70	—	—	2.48	2.15[d]
MAD	0.57	0.34	0.11	0.14	0.14	

[a]: Reference [199], [b]: Reference [239], [c]: Reference [240], [d]: Reference [241]

Table A.6: S_0-S_1 energies (eV) after addition of the TD-DFT solvent shift δS for the same MCs and methods as in Table A.5; MAD for individual methods with respect to the experimental excitation energy..

Merocyanine	TD-DFT	GW-BSE	STEOM-DLPNO-CCSD	CASSCF/NEVPT2-FIC	TDCP-DFT	Exp.
D1A1	2.19	1.79	1.60	1.76	1.63	1.90
D1A2	2.66	2.35	2.00	1.99	2.10	2.31
D1A3	2.27	2.05	1.69	2.06	1.99	2.15
D2A4	2.37	2.25	—	—	2.03	2.15
MAD	0.25	0.09	0.36	0.18	0.19	

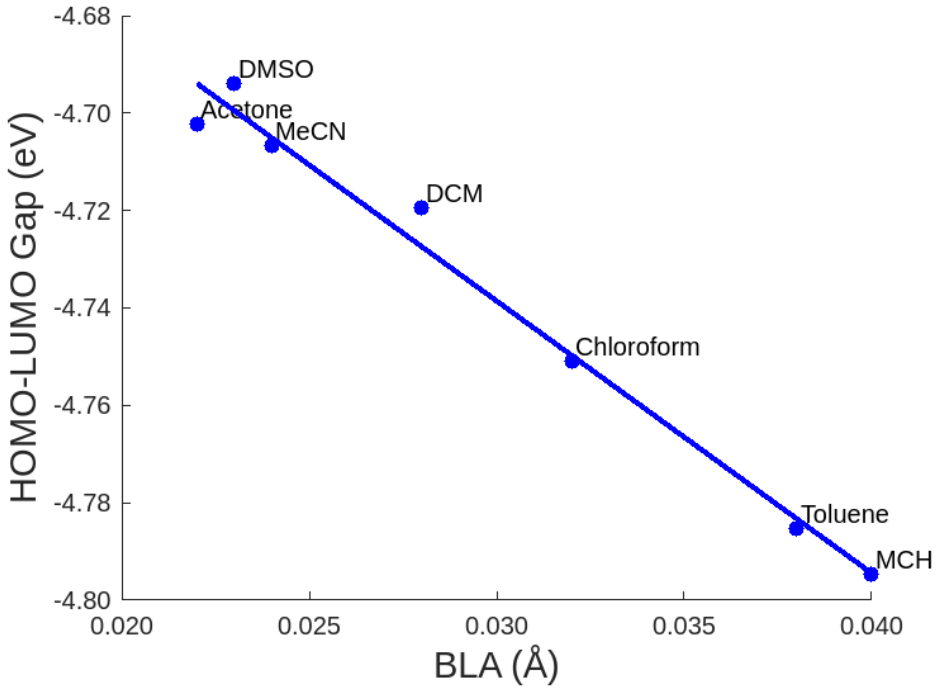


Figure A.6: The BLA dependence of MC D2A4 on the HOMO-LUMO gap of the solution calculated at B2PLYP/D3 level with def2-TZVP basis set. Actual values are represented by circles, solid line shows a linear fit of the data points.

Table A.7: S_0-S_1 energies (eV) obtained for MC structures optimized with the solvents that have similar dielectric constant as the respective molecular crystal. The S_1 energies are shifted by δS of Table A.4 and MAD is calculated with respect to the experimental excitation energy.

Merocyanine	Solvent[a]	Solvent[b]	TD-DFT	GW-BSE	TDCP-DFT	Exp.
D1A1	Toluene	DCM	2.16	1.83	1.52	1.90
D1A2	Toluene	1,4-dioxane	2.64	2.32	2.10	2.31
D1A3	Toluene	DCM	2.29	2.05	2.01	2.15
D2A4	Chloroform	DMSO	2.34	2.18	1.99	2.15
MAD			0.23	0.05	0.22	

[a]: Solvent used for geometry optimization, [b]: Solvent used in TDDFT for excitation energy calculation.

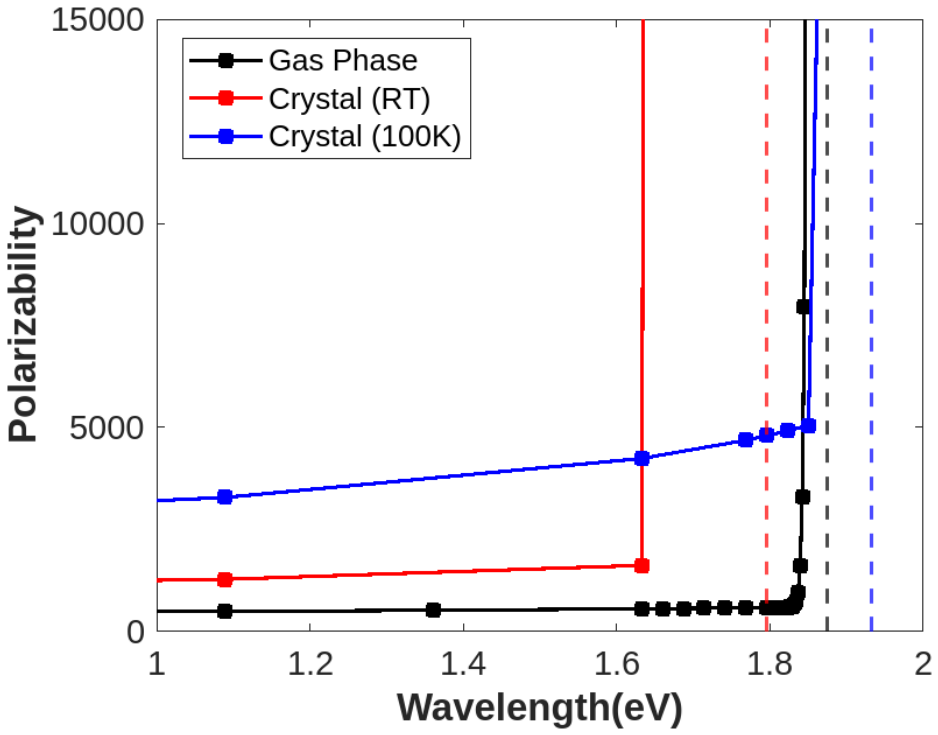
Table A.8: S_0-S_1 energies (eV) obtained for MC structures optimized with B2PLYP in the respective solvents shifted by δS of Table A.4. MAD with respect to the experimental excitation energy.

Merocyanine	Solvent	TD-DFT	GW-BSE	TDCP-DFT	Exp.
D1A1	DCM	2.10	1.75	1.43	1.90
D1A2	1,4-dioxane	2.64	2.32	2.10	2.31
D1A3	DCM	2.30	2.04	1.99	2.15
D2A4	DMSO	2.31	2.15	1.94	2.15
MAD		0.21	0.07	0.26	

B2PLYP geometry of the molecules was used. The excitation energy of the MC is 0.1-0.3 eV lower in the molecular crystals than in gas phase except for P1-D1A1. (Table A.9). The corresponding crystal structures and crystallographic data are shown in Figure S5 and Table S3. Figure S4 shows the calculated mean dynamical polarizability (MDP) of the two **D1A1** molecular crystals along the three crystallographic directions. The first and only observed absorption peak was detected in all three crystallographic orientations at 1.89 eV. In Table A.9 we also report the difference in excitation energy of two D1A1 polymorphs denoted as P1 (CCDC 2073437) and P2 (CCDC 2073438).[242] The excitation energy of P1-D1A1 and P2-D1A1 differ by 0.04 and −0.09 eV, respectively, from the excitation energy of gas phase D1A1 (dashed lines in Figure A.7). Since it is known that such shifts are caused by intermolecular coupling of transition dipoles (the so-called J- and H-couplings) [52, 55], we inserted the transition dipole vectors (calculated for gas-phase D1A1 at PBE/D3/def2-TZVPP level) in the unit cells. This representation of P1-D1A1 and P2-D1A1 suggests weak couplings in the two polymorphs (see Figure S7 in the SI). Consequently, the shifts of the S_1 peaks with respect to isolated molecules are relatively small compared to ordered aggregates on surfaces. [55]

Table A.9: Dielectric constants and excitation energies (eV) of the molecular crystals calculated using TDCP-DFT at B3LYP/pob-TZVP level

2*Merocyanine	2*Dielectric constant	Excitation energy	
		Crystal	Gas phase
P1-D1A1	2.69	1.93	1.89
P2-D1A1	2.69	1.80	1.89
D1A2	2.66	2.29	2.37
D1A3	3.12	2.04	2.31
D2A4	5.15	2.18	2.48



H

Figure A.7: Energy dependence of the many-body polarizability tensor for two **D1A1** polymorphs measured at 100K and room temperature (RT) computed using TDCP-B3LYP. The dynamical polarizabilities are represented by diamonds, solid lines serve as guide to the eye. Red and blue dashed lines represent the optical gap of the crystal structures at the respective temperatures.

A.3 Conclusions

We calculated the structural and optoelectronic properties of selected MCs with different donor (D) and acceptor (A) chromophores for which experimental absorption spectra have been reported earlier. Excitation energies were calculated with the TD-DFT, GW-BSE, STEOM-DLPNO-CCSD, CASSCF/NEVPT2, and TDCP-DFT methods. Using the hybrid DFT functionals PBE0 and B3LYP, the experimental ionization energy is best reproduced. Surprisingly, the G_6W_0 method gave too low HOMO energies. However, for the calculation of the first excitation energy, GW-BSE was by far the most accurate approach. The accuracy was further improved by taking into account the geometry changes and solvent shifts in solution. TD-DFT methods overestimate the first excitation energy, which was also observed in previous studies. The more sophisticated STEOM-DLPNO-CCSD and CASSCF/NEVPT2 methods also provide accurate excitation energies but at much higher computational cost. Furthermore, no systematic improvement was possible by adding the solvent shifts obtained with TD-DFT. The TDCP-B3LYP method, which has been previously applied only for crystalline semiconductors and insulating materials, was extended to organic molecules and molecular crystals. In the gas phase, its accuracy is comparable to GW-BSE, STEOM-DLPNO-CCSD, CASSCF/NEVPT2, but it can also not be systematically improved by TD-DFT solvent shifts. We observed a lowering of the first excitation energy in the solids compared to the gas-phase molecule due to exciton coupling. The excitation energy of MCs exhibits sensitivity to the crystal structure. However, in the solid state,

the intermolecular coupling of transition dipoles is relatively weak, resulting in modest variations in excitation energy.

A.4 Supporting Information

Supporting information includes BLA pathways of MC's and comparative line graphs illustrating bond length variations of different MCs using diverse functionals; Basis set dependence of $S_0 - S_1$ and HOMO energy for D1A1; BLA vs ϵ graph for D2A4; Polarizability tensors of MC D1A1 along the three crystallographic axis are provided along with crystallographic Parameters and molecular crystal structures of all MCs; Transitional dipole moment vectors of MC's and representative image of the crystal structures using dipole moment vectors.

A.5 Acknowledgement

The authors gratefully acknowledge funding from the DFG (RTG-2591 *TIDE Template-designed Organic Electronics*), and the computer resources from the HPC Centre at the University of Bonn. This research was supported in part by the University of Pittsburgh Center for Research Computing through the resources provided. Specifically, this work used the H2P cluster, which is supported by NSF award number OAC-2117681. D.F. gratefully acknowledges N. Gildemeister and S. Lekat for their initial explorative works on the excited states of merocyanines and the computer center "Regionales Rechenzentrum der Universität zu Köln" (RRZK).

APPENDIX B

Modelling the monolayer formation of Merocyanine HB238 on Ag(100) surface

Ritu Tomar[†], Anna J. Kny[‡], Moritz Sokolowski[‡], Thomas Bredow[†]

Received: 24 December 2024
Published online: 8 April 2025

Adapted with permission[¶] from
R. Tomar, Anna J. Kny, M. Sokolowski, and T. Bredow, *Modelling the monolayer formation of Merocyanine HB238 on Ag(100) surface*, *J. Phys. Chem. C* **129** (2025) 7427-7436.
DOI:10.1021/acs.jpcc.4c08683

Own Contributions

- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[‡] Clausius Institute of Physical and Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to the American Chemical Society

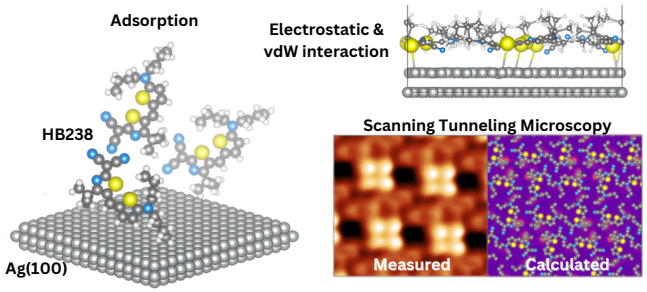


Figure B.1: Table of content graphic (ToC).

Abstract We recently published a study on the adsorption of the merocyanine (MC) HB238 (2-[5-(5-dibutyl-amino-thiophene-2-yl-methylene)-4-tert-butyl-5H-thiazol-2-ylidene]-malononitrile) on the Ag(100) surface using scanning tunneling microscopy (STM), spot profile analyzing low electron diffraction (SPA-LEED), x-ray photoelectron spectroscopy (XPS) and ultraviolet photoelectron spectroscopy (UPS) techniques[Kny et al., Nanoscale, 2023, 15, 10319]. We observed that HB238 self-organizes as a chiral tetramer upon adsorption. The adsorbate structure, denoted as α-phase, is completely different from the known bulk structures of HB238. Whereas, the previous study was mainly experimental, in the present study we focus on the results of quantum chemical calculations to corroborate the previous experimental findings and interpret measured spectra in order to deepen the understanding of the adsorption process. For this purpose, we first screened the full conformer space of the HB238 molecule, particularly with respect to its flexible alkyl side chains which obtain several degrees of freedom due to possible rotations along their single bonds. The most stable structures were then used to determine the optimal orientations of HB238 on the Ag(100) surface and to compare to the experimental α-phase structure.

We systematically generated various initial configurations of the HB238 monolayer on a three-layer Ag slab model. For the most stable adsorption model, STM images, partial density of states (PDOS), x-ray absorption spectra (XAS) and adsorption height profiles were calculated and compared with the previous experimental results. Our observations reveal a significant alteration in the HB238 conformation upon adsorption, with monolayers forming through stabilization mainly by electrostatic and dispersion interactions with the Ag(100) surface.

B.1 Introduction

Organic electronics have emerged as a promising research topic, presenting new opportunities for advancements in technologies such as organic photodetectors and organic field-effect transistors (OFETs).[30, 31, 44] The study of the electronic and optical properties of organic molecules adsorbed on metal surfaces is of high importance in the field of molecular electronics and optoelectronics.[31, 82] Merocyanines (MC) are interesting molecules for optoelectronic applications as they feature electron donor (D) and acceptor (A) groups interconnected by a conjugated methine bridge. The D-A charge transfer leads to large dipole moments and high oscillator strengths.[45, 206] These properties contribute to high absorption coefficients and adjustable optical and electronic traits, making this class of molecules potential candidates for optoelectronic applications.[49] MCs have a non-planar structure, low symmetry, and possess side-groups that are non-rigid and structurally flexible. This

makes it initially counterintuitive to consider MCs as candidates for forming ordered structures on surfaces. Yet, scanning tunneling microscopy (STM) studies on MC layers deposited in vacuum have uncovered intriguing characteristics, such as the formation of dimers or tetramers.[82, 257, 258] The Ag(100) surface was chosen here due to its high symmetry and moderate chemical reactivity.

The adsorption of the MC dye on the surface is expected to lead to changes in the electronic and optical properties of both molecule and surface. In addition, the formation of ordered structures can significantly impact the overall properties of the adsorbate.[55]

Despite its importance, theoretical investigations of MC monolayers at substrate interfaces remain scarce.[81, 259] Owing to their complex, flexible structure, and low symmetry, MCs present challenges for the theoretical modeling of ordered structures on surfaces. Recently, an experimental study was conducted by Kny et al. [82], focusing on the adsorption and ordering of the MC HB238 (2-[5-(5-dibutyl-amino-thiophene-2-yl-methylene)-4-tert-butyl-5H-thiazol-2-ylidene]-malononitrile) on the Ag(100) surface using STM, SPA-LEED, UPS and XPS. To complement these experimental findings, our present study focuses on the theoretical understanding of the HB238 interaction and the Ag surface. Specifically, the present investigation aimed to clarify the chemical bonding of the HB238 molecule to the Ag(100) substrate, to deduce the changes in the electronic states due to interfacial bonding and to explain the monolayer formation based on interfacial and intermolecular bonding.

A significant aspect of our study is the application of density functional theory (DFT) calculations to explore the electronic structure, intermolecular interactions, and adsorption interactions of HB238 on the Ag surface. This approach contrasts with previous studies that have primarily employed molecular mechanics (MM) calculations for surface interaction with the NaCl surface.[260] Although MM provides a suitable approximation scheme for a larger application field with less computational demand, it does not replace the rigor of ab initio quantum mechanical approaches such as DFT which are required for adsorption on metal surfaces.

We compare calculated adsorption structures, STM images, PDOS, and XAS with the experimental data to aid in the interpretation of the measured spectra, thereby providing a deeper analysis of the interaction between HB238 and the Ag(100) surface.

B.2 Computational methods

Prior to the investigation of adsorption structures of HB238 on Ag(100), we performed a conformer search of the molecule in gas phase using the conformer rotamer ensemble sampling tool (CREST) as shown in Figure B.2. CREST identified 356 stable conformers, and the geometries of these conformers were optimized at the PBE/D3 level using ORCA.[137, 229] A selection of stable conformers was used for subsequent adsorption calculations.

For the single-molecule adsorption on Ag(100), we initially performed a semi-empirical pre-optimization using the GFN1-xTB method.[134] All semi-empirical calculations were carried out using DFTB+.[261] The final structure optimizations were then conducted using the VASP (version 6.3.2) plane-wave DFT code, employing the PBE/D3 functional.[128, 129, 132, 220] For optimization, we employed the standard PAW sets provided by the potcar-PBE.5.4 suite.[125, 262] The key computational parameters included an energy cutoff of 400 eV, the VeryFast electronic optimization algorithm, and a Gaussian smearing (ISMEAR=2) with σ =0.2 eV. A k-point grid of (1x1x1) was used with the Γ-point approximation, and automated real-space projection settings (LREAL=Auto) were applied to balance computational efficiency with the precision required for studying complex surface

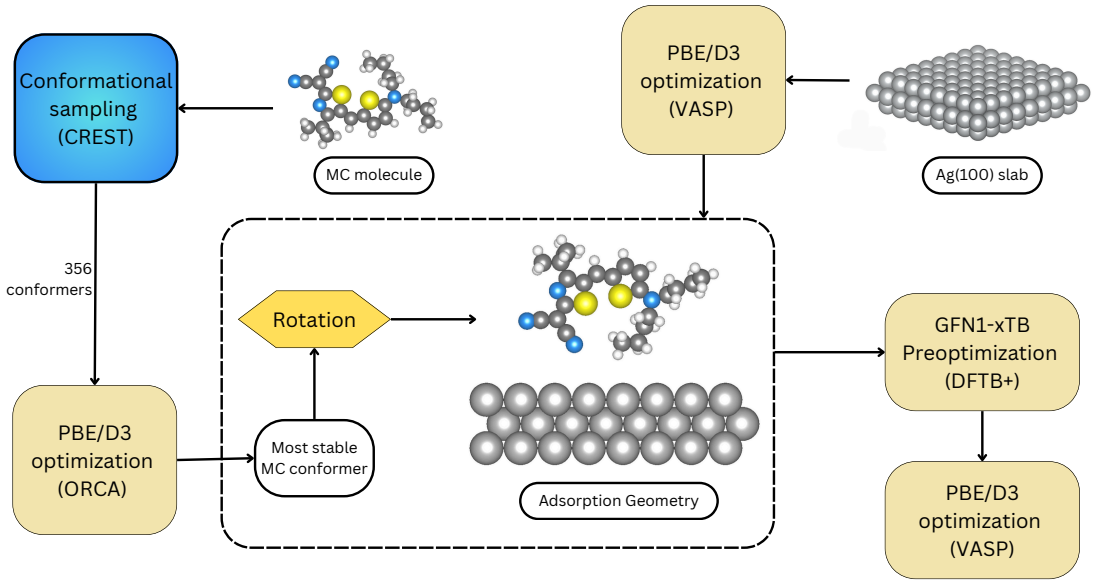


Figure B.2: A workflow of the simulation process for the adsorption of a single HB238 molecule adsorbed on the Ag(100) surface.

phenomena. The surface model was based on a three-layer Ag(100) slab, where the bottom two layers were fixed, and the topmost layer was allowed to relax. The convergence test calculations are given in Table S1 and S2 in SI. The supercell was separated by a vacuum distance of approximately 20 .

Starting from the optimized gas-phase conformers, monolayer models of HB238 were constructed. Experimentally, the -phase of HB238 is known to form a commensurate structure with four symmetry-equivalent domains. The experimental unit cell parameters were a = 23.30Å, b = 24.69Å, and γ = 103.43° in the P211 slab group.[82] The PBE/D3-optimized unit cell parameters were used to model the adsorption of HB238 on the Ag(100) surface. The transformation matrix for the primitive Ag(100) unit cell is given by

$$(\vec{a}, \vec{b}) = \begin{pmatrix} 1 & -8 \\ 8 & 3 \end{pmatrix} \begin{pmatrix} \vec{a_1} \\ \vec{b_1} \end{pmatrix} \tag{B.1}$$

where $\vec{a_1}$ and $\vec{b_1}$ correspond to Ag(100) lattice vectors of 2.89 (experimental) and 2.88 (calculated). The monolayer models were initially screened using GFN1-xTB optimization, and the resulting stable structures were further optimized with VASP (PBE/D3) using the PreconLBFGS optimizer. The azimuthal orientation of the four HB238 molecules in the unit cell was systematically varied while maintaining fourfold symmetry. The individual HB238 molecules were placed at (0,0,h), (0,-0.5b,h), (-0.5a,0,h), and (-0.5a,-0.5b,h) positions in the monoclinic unit cell, where h is the initial height above the surface, here 2.5 . Due to the high flexibility of the molecules, the PreconLBFGS optimizer significantly reduced the number of optimization steps required to reach convergence and, in all investigated cases, led to more stable local minima (see Figure S2). The present example thus displays a case where the use of a highly efficient optimizer is mandatory, differently to earlier studies on

smaller unit cells of π-conjugated molecules on surfaces.[263]

The adsorption energies E_{ads} per molecule are calculated as

$$E_{ads} = \left(E^{opt}_{\text{Ag:HB238tetramer}} - E^{opt}_{\text{Ag}} - 4E^{opt}_{\text{HB238}}\right)/4 \tag{B.2}$$

where $E^{opt}_{\text{Ag:HB238tetramer}}$ represents the energy of the HB238 tetramer adsorbed on Ag(100), E^{opt}_{Ag} and E^{opt}_{HB238} are the energies of the pristine Ag(100) slab and isolated HB238 molecule(most stable conformer) optimized at PBE/D3 level, respectively.

The intermolecular interaction energy E_{int} is calculated as

$$E_{int} = \left(E^{sp}_{\text{HB238tetramer}} - \sum E^{sp}_{\text{HB238}}\right)/4 \tag{B.3}$$

where $E^{sp}_{\text{HB238tetramer}}$ and E^{sp}_{HB238} represent the single-point energies of the HB238 monolayer in adsorption model and the HB238 molecule in gas phase with adsorbate structure, respectively, calculated independently at the PBE/D3 level. The reorganization energy E_{reorg} is calculated as the difference between the molecular energies in the adsorption structure and in optimized gas-phase structure. It is a measure of the destabilization of molecule due to geometry changes upon adsorption.

B.3 Results and Discussion

B.3.1 Single molecule adsorption

The adsorption geometry of a single HB238 molecule on the Ag(100) surface was investigated using a stepwise approach. Initially, conformer sampling of HB238 molecules in the gas phase was performed using CREST identifying the most stable conformer (see Figure B.3). In this structure, the CN groups and the methylene group are part of the π-conjugated backbone and lie within the same plane, the n-butyl groups form a zigzag arrangement, with the connecting methylene (CH_2) group lying in the same plane as the π-conjugated system. One CH_3 group of the tert-butyl group is in the molecular plane and the other two are positioned symmetrically above and below following the tetrahedral bonding arrangement. This conformer was placed 2 Å above the top Ag layer on the Ag(100) supercell (eqn. 1), in various orientations including face-on, side-on and edge-on as shown in Figure B.4. Geometry optimization of each initial structure was performed using the GFN1-xTB [134] Hamiltonian in DFTB+[261]. Further refinement was performed with PBE-D3 as described above. The resulting minimum energy configuration corresponds to a face-on adsorption of HB238 on the Ag(100) surface which also gives the largest footprint of the molecule. This adsorption geometry indicates a strong interaction between the HB238 molecule and Ag surface, with the molecular plane parallel to the surface, as observed experimentally for the α-phase [82].

The single-molecule adsorption study was expanded by considering a series of different conformers to investigate the influence of the adsorbate's initial geometry on the adsorption process and its interactions. This investigation involved the utilization of 50 distinct and most stable conformers from the 356 structures obtained with CREST, each placed in random orientations upon Ag the surface. Each configuration was within an energy range of 0.4 eV compared to the most stable structure. Subsequent optimization of these systems was conducted using DFTB+ at the GFN1-xTB level of theory. We observed that all selected conformers achieved a face-on adsorption orientation on the Ag

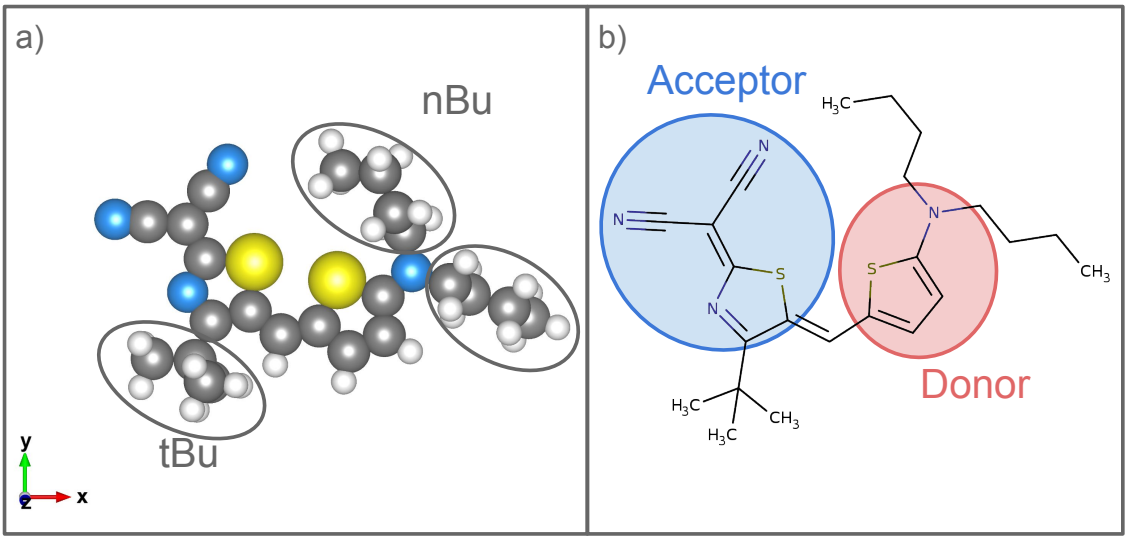


Figure B.3: a) PBE/D3 optimized structure of the most stable HB238 conformer in gas phase obtained using CREST. The C, H, S and N atoms are represented using grey, white, yellow and blue spheres, respectively; the nBu and tBu groups are marked with grey circles. b) The valence structure of HB238. The red and blue colors indicate the donor and acceptor group, respectively.

substrate as a minimum structure. The face-on configurations were characterized by the π-conjugated backbone of the HB238 molecule being almost parallel to the surface. The minor variations in the adsorption energy ($\approx$ 0.01 eV) with respect to the most stable conformer of these systems can largely be attributed to the influence of the flexible n-butyl groups on the surface (See Figure S1).

Further, we focused on differentiating the energy and structural variations in the adsorption modes of the HB238 molecule on the surface. This involved reorienting the HB238 molecule in different adsorption modes as depicted in Figure B.4. The most stable gas phase structure of HB238, determined through CREST and subsequent DFT calculations (PBE/D3), was positioned in various orientations on the surface: face-on, side-on, and edge-on. Additionally, we introduced variations by rotating the HB238 molecule by 180° in each of these configurations. This approach was designed to provide an understanding of how different orientations of the HB238 molecule affect its adsorption characteristics and energetics on the surface. The side-on and edge-on orientations are less stable than face-on orientation by 1.63 $\pm$ 0.02 eV and 2.06 $\pm$ 0.41 eV, as defined in Figure 2a) respectively.

B.3.2 Adsorption models for the monolayer

In the first set of structure optimizations, the monolayer adsorption models maintained a consistent initial orientation of HB238 molecules, guided by a structure model of HB238 tetramers derived from experimental data by Kny et al. [82]. This predicted α-phase structure places four tert-butyl groups in close proximity, resulting in the distinct STM contrast that is explained in the subsequent section. We introduce variations in the HB238 conformations within this model, obtained through CREST

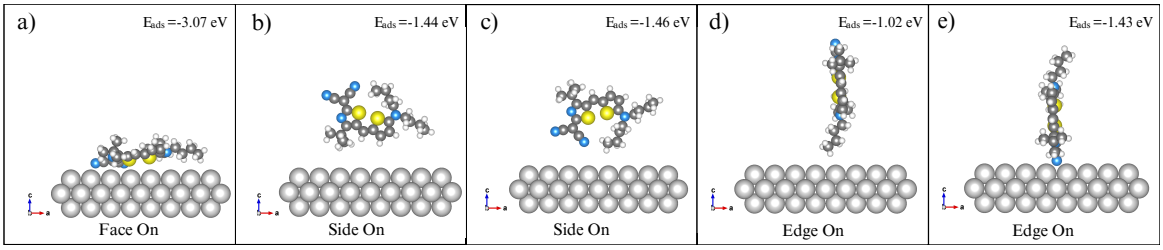


Figure B.4: A single HB238 molecule adsorbed on the Ag(100) surface in various orientations (side view). a) Face-on orientation, where the molecular plane lies parallel to the surface. b, c) Side-on orientation and edge-on orientation, with the molecular plane perpendicular to the surface plane along the x and y direction, respectively as defined in Figure B.3 a.

calculations. Since the adsorption energy of the conformer is not directly correlated to its stability in gas phase, all 356 most stable conformers were incorporated into this monolayer model and relaxed, to identify the energetically stable monolayer configuration. The relative adsorption energies of all stable HB238 monolayer structures are shown in Figure S3 with reference to the absorption energy of the most stable structure, i.e. -3.51 eV. We identified the conformer which resulted in the most stable monolayer configuration upon adsorption. The conformer which is ranked as the most stable structure by CREST features a planar π-conjugated backbone with n-butyl group arranged in a compact fashion (Figure S4). The n-butyl groups do not significantly disrupt the planarity of the molecule.

In the starting geometry, the monolayer models have the following characteristics: The adsorption positions of the four molecules in the stable monolayers are obtained by rotations around the molecule's center of mass from 0° to 355° in steps of 5°(See Figure B.5) and have P1 slab group symmetry.

We investigate these structural features using electronic structure and charge density calculations in the next sections.

The adsorption energy of the HB238 conformer with respect to the angular rotation (25°) is shown in Figure B.5. The most stable monolayer adsorption structure obtained in this way is denoted as model A (see Figure B.6). The variations in adsorption energies can be attributed to the interplay of various factors, such as electrostatic and dispersion interactions between HB238 and Ag(100), molecular reorganization, and intermolecular interaction between the HB238 molecules.

The calculated adsorption energies of -3.51 eV/molecule for the monolayer (model A) and -3.07 eV for the single HB238 molecule adsorbed on Ag surface indicate attractive interaction between the HB238 molecules in the monolayer by hydrogen bonding, dispersion and electrostatic interaction. For comparison, a much smaller adsorption energy of a thiophene molecule on the Ag(111) surface has previously been reported, -0.75 eV.[263] The E_{int} and E_{reorg} for Model A are -2.28 eV and 0.94 eV, respectively. The electrostatic contribution to the interaction energy is -0.96 eV and is dominated by the Ag-N interaction amounting to -0.88 eV. The electrostatic interaction of the HB238 molecules with the surface was calculated in the monopole approximation. The atomic charges were calculated using DDEC6 atomic population analysis in CHARGEMOL (Table of charges in SI).[264] The lateral stabilization can mainly be attributed to the intramolecular C-H$\cdots$N hydrogen bonding(see Figure S5). The average C-H$\cdots$N bond length is 3.2 Å, corresponding to the cusp of moderate and weak hydrogen bonding.[265] The lateral interactions stabilize the adsorption of the monolayer, however, a major contribution to the adsorption energy comes from the molecule to surface interaction. The reorganization energy is small indicating that a slight structure rearrangement is possible and therefore

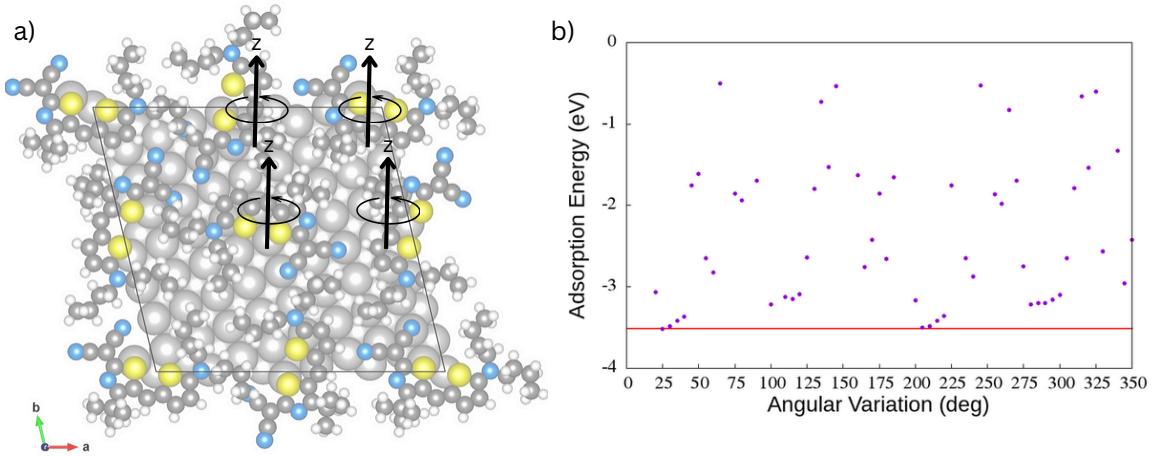


Figure B.5: a)Schematic picture of the simultaneous rotation of 4 HB238 molecules along the azimuthal angle defined with respect to the initial orientation obtained from CREST as in Figure B.3 a, around their respective center of mass. b)The variation of adsorption energy (in eV) of HB238 monolayers with respect to azimuthal angular variation (in degrees), indicating the energy landscape experienced by the rotating molecules. The adsorption energy of the monolayer is calculated upon simultaneous rotation of four molecules in the supercell. The red line represents the average adsorption energy of HB238 molecule in monolayer model A (25° rotation).

the azimithual orientations between 25 and 40° have similar stability (Figure B.5 b). Similar adsorption energies are obtained for angles between 205 and 240°, due to the symmetry of the Ag(100) surface. We also varied other arrangements of HB238 molecules to test different parallel and anti-parallel orientations the HB238 dipole vectors (See Figure B.7). The three structures, denoted as B, C and D, are less stable than model A and differ in adsorption energy by 1.1, 1.2 and 0.3 eV per HB238, respectively (See Table B.1. The structure C in Figure B.7 is 0.3 eV less stable than Abut fails to explain the orbital features observed in experiment by Kny et al.[82]

B.3.3 Scanning Tunneling Microscopy images

We calculated the STM images using the Tersoff-Hamann approximation in VASP.[175] Calculated STM images were rendered using the postprocessing script STM-2Dscan.py.[266] The energy range (EINT) for the calculation performed on the most stable adsorbate structure of model A was set from -0.3 to 0 eV with respect to Fermi energy, and the tip position was fixed at a distance of 9.5 Å from the surface. The experimental STM image was acquired at 300 K with a bias voltage (U_{bias}) of -0.3 V and a tunneling current (I_t) of 0.05 nA.

The calculated STM image captures the essential features of the experimental image, demonstrating that the bright spots originate from the tert-butyl groups and since little to no tunneling contrast is observed around the carbonitrile ends, the black adjacent spots arise from the empty space between

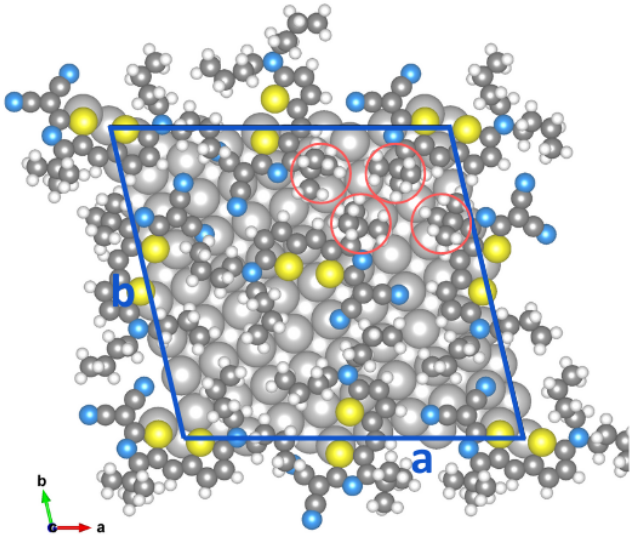


Figure B.6: The most stable adsorption model A of a HB238 monolayer adsorbed on Ag(100) optimized using PBE/D3 (E_{ads} = -3.51 eV). The positions of the tert-butyl groups are marked by red circles.

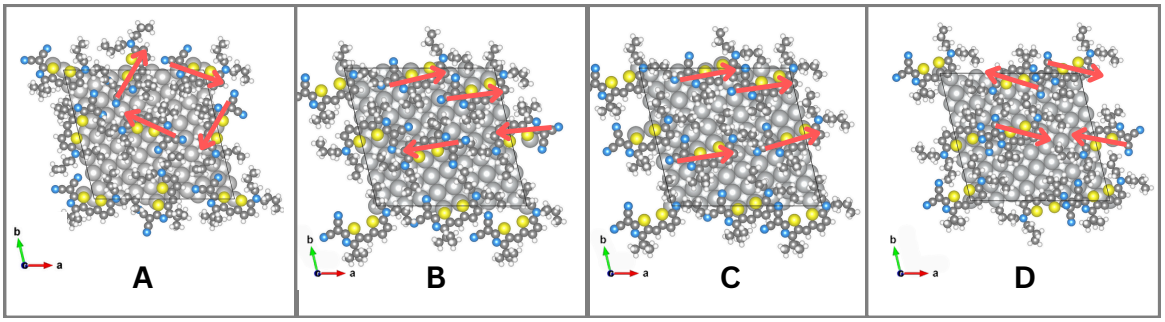


Figure B.7: Different tetrameric structures of HB238 where the dipole vectors are placed in parallel and anti-parallel fashion, optimized using preconLBGFS at PBE/D3 level. Red lines indicate the transition dipole vector of the HB238 molecule calculated in gas phase which serves as a visual guide for the HB238 orientations. [163]

Table B.1: The calculated adsorption energy (E_{ads}), reorganization energy (E_{reorg}), intermolecular interaction energy (E_{int}), intermolecular electrostatic interaction energy (E_{int}^{elec}), surface-molecule electrostatic interaction energy (E_{elec}), and dispersion energy (E_{disp}) in eV for HB238 monolayer adsorption models A, B, C, and D on Ag(100) surface (See Figure B.7).

Adsorption Model	A	B	C	D
E_{ads}	-3.51	-3.10	-3.19	-3.23
E_{reorg}	0.78	0.96	1.00	0.73
E_{int}	-0.81	-0.01	-0.08	-0.74
E_{elec}	-0.23	-0.18	-0.13	-0.16
E_{disp}	-2.78	-2.55	-2.60	-2.47

the HB238 molecules and carbonitrile groups (see Figure B.8).

Further analysis of the STM images reveals several similarities and differences between the theoretical and experimental structural model. A significant observation is that the carbonitrile groups of diagonally opposite HB238 molecules are positioned closer together than anticipated in the experimental model, contributing to the black contrast observed in the STM image (as shown in Figure B.8c).

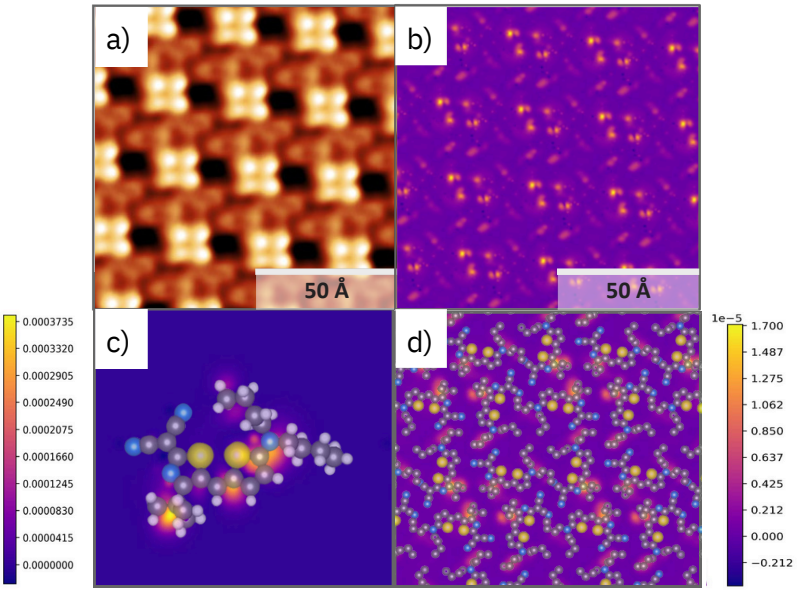


Figure B.8: Comparison of a) Experimental STM image of the α-phase (T = 300 K, U_{bias} = 3.0 V; I_t = 0.05 nA) adapted with permission from ref [4]. Copyright 2023 Royal Society of Chemistry; b) Calculated STM images (EINT = -0.3 to 0 eV, Tip position = 9.5 Å) for model A. ; c, d) Calculated STM images of a single adsorbed molecule and monolayer adsorbed on Ag surface superimposed on the hardsphere model of HB238 molecule and monolayer, respectively. The H atoms in d) are removed for clarity and distances are given by a scale bar. Calculated STM images were rendered using STM-2Dscan.py.[266]

B.3.4 Adsorption Height and Charge Density Analysis

We determined the adsorption height of the HB238 monolayer on the Ag surface by calculating the averaged vertical distance between the S atoms of HB238 and the Ag surface (Figure B.9). The adsorption height analysis was performed on the most energetically stable monolayer structure (model A) described above and the average adsorption height is given in Table B.2. The overall average adsorption height was determined to be 2.84 Å for the isolated molecule and 2.99 Å for the monolayer. The distance between the S atoms of HB238 and the Ag surface in both cases is less than the sum of their van der waals radii (3.52 Å) which implies covalent interactions, contributing to the overall stability of the monolayer.

Table B.2: The average heights of the acceptor and donor sulfur atoms adsorbed on the Ag surface (Å). Heights are measured with respect to the average Ag(100) surface plane.

Models	Average Adsorption Height	
	Donor S	Acceptor S
Single molecule (PBE/D3)	2.90	2.78
Monolayer (PBE/D3)	3.09	2.88

For a single HB238 molecule adsorbed on the Ag(100) surface (PBE/D3), the donor S atom is slightly higher (2.90 Å) than the acceptor S atom (2.78 Å). When considering the monolayer (PBE/D3), the average heights increase to 3.09 Å (donor) and 2.88 Å (acceptor). These variations in height can be attributed to the weakening of the molecule surface interactions due to intermolecular interactions between neighboring HB238 molecules and surface atoms. The interaction also leads to a slight rumpling (variation of the vertical position of topmost Ag atoms) of the surface by ≈ 0.2 Å. In the adsorbed state, S atoms are not positioned directly above Ag atoms, but with an angular deviation. This deviation varies between 8 and 28 degrees from the perfectly aligned atop placement (See Figure B.9 b). The average distance between S atoms within the molecule in single molecule and monolayer is 3.02 and 3.10 Å, respectively.

We performed the molecular charge analysis using DDEC6 population analysis in CHARGEMOL.[264] The resulting atomic charges are shown in Figure S6 (full Table in SI). The atomic charges range from -0.42 to -0.28 *e*. Notable charge accumulation can be observed on the C atoms of the tert-butyl groups which is in accordance with the observation of brighter lobes in the STM experiments (see Figure B.10). This is also observed for the gas phase molecule and is therefore not induced by adsorption. A slight depletion of the atomic charges for the S atoms is observed upon adsorption. The average atomic S charge in gas phase molecule and monolayer model A are 0.03 and 0.07 *e*respectively (See Table S3 and S4 in SI).

In the charge density difference (CDD) plot, we observe a clear and distinct distribution of charge depletion and charge accumulation over the entire HB238 monolayer on the Ag(100) surface (Figure B.11). The charge redistribution patterns suggest that the adsorption of the HB238 monolayer involves substantial changes in electron density and site-specific bonding interactions are present between the HB238 monolayer and the Ag surface. The charge depletion around the hydrogen (H) atoms is closer to the surface, indicating possible involvement in bonding interactions with the Ag surface. Concurrently, the attached carbon atoms exhibit an increased electron density upon adsorption. Additionally, a notable density redistribution was observed around the sulfur (S) atoms and the nitrogen

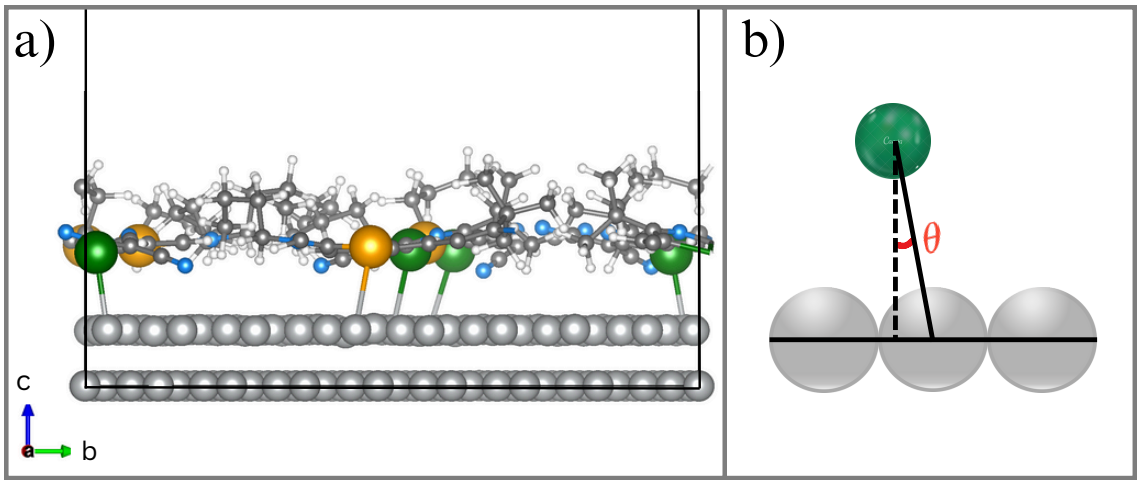


Figure B.9: a) HB238 monolayer adsorbed on the Ag(100) surface (side view of model A). The molecular structure of the monolayer is presented, showing the bonding between the donor or acceptor S atoms (represented as yellow and green spheres) of HB238 and the Ag atoms (depicted as silver spheres) of the surface. The radii of S atoms have been increased to 2 Å for better visibility. b) Representative picture of the vertical and actual S-Ag distance in monolayer shown using dotted and slanted black lines, respectively. The angle θ represents the deviation from S-Ag atop adsorption positions.

(N) atoms of the cyano group, implying the presence of strong polarization around these atoms. This polarization will further stabilize the monolayer on the surface beyond the simplified point charge interaction.

B.3.5 Projected Density of States

In order to employ a less empirical theoretical approach compared to the commonly used DFT+U approximation [268], we calculated the PDOS of adsorbed HB238 using various hybrid functionals including HSE06 [269], B3LYP[146, 222, 223] and r^2SCANh[**r2scan**](with 10% Fock exchange). Hybrid functionals are preferable over GGA functionals to calculate the electronic structures of molecules and condensed phases. The inherent self interaction error of GGA is diminished by the inclusion of Fock exchange in the corresponding exchange functional. In these electronic structure calculations, we applied a 450 eV cutoff energy and $1 \times 1 \times 1$ Monkhorst-Pack. A comparison of the PDOS obtained with PBE+U, HSE06, B3LYP and r^2SCANh is given in Figure B.12. We tested the convergence of calculated PDOS using higher cutoffs and different Projector augmented wave (PAW) potentials with B3LYP, but this does not have significant effects on the results. The experimental highest occupied molecular orbital (HOMO) and next HOMO (HOMO-1) energies for the HB238 monolayer model A are -1.77 eV and -2.71 eV, respectively.[82] The hybrid functionals improve the description of the HOMO energy with respect to the experimental UPS compared to conventional PBE

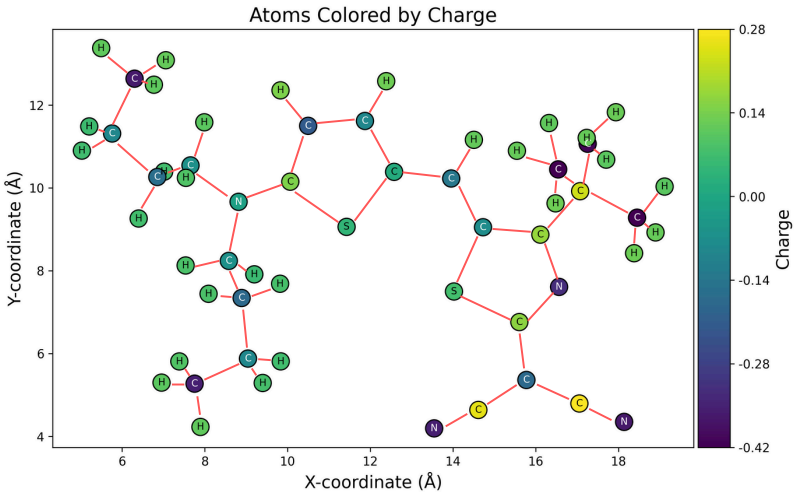


Figure B.10: Atomic charge distribution plot of one HB238 molecule adsorbed on the Ag(100) calculated using DDEC6 population analysis using CHARGEMOL.[264] All charges are in atomic units.

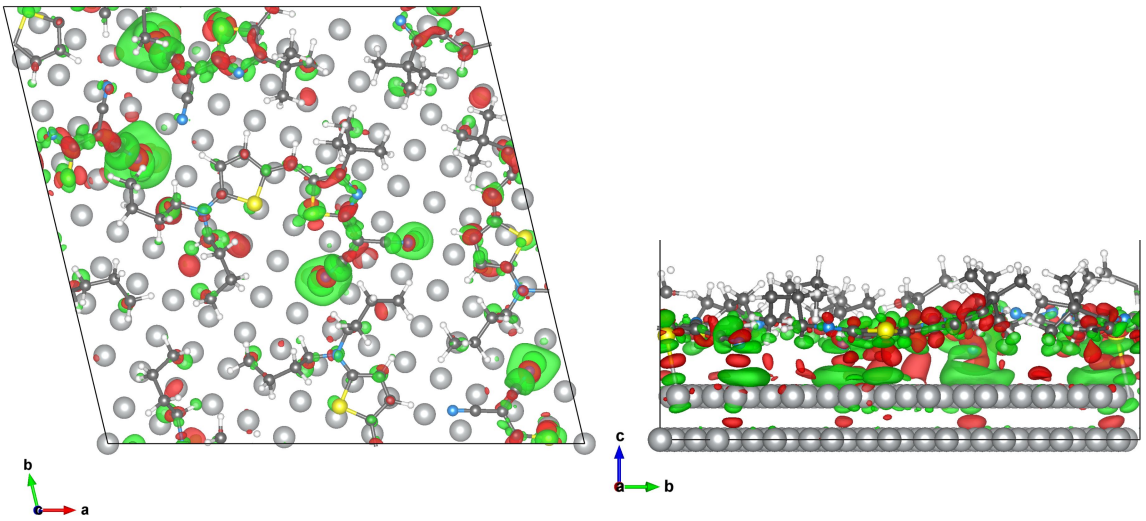


Figure B.11: Charge density difference (CDD) plot of HB238 monolayer on Ag(100) surface revealing the changes in electronic charge resulting from the adsorption process. Green regions indicate charge depletion, while red areas represent charge accumulation. CDD was calculated at r^2SCANh level and plots were drawn with VESTA at an isovalue of 0.001 electrons/au^3.[267]

+U (see Table B.3). B3LYP results have an average difference of 0.28 eV, while HSE06 and r^2SCANh have much smaller differences of 0.04 eV and 0.03 eV, respectively. For the HOMO-1 energy level, the discrepancies are 0.55 eV for B3LYP, 0.32 eV for HSE06, and 0.19 eV for r^2SCANh. Therefore, we recommend to employ the meta-GGA hybrid functional r^2SCANh for the calculation of electronic levels even for adsorption on metallic surfaces which is usually considered the domain of GGA. However, it has to be noted that the hybrid calculations are extremely demanding with VASP. They are feasible only on modern GPU architectures. The HOMO and LUMO energies for gas phase molecule are -3.41 eV and -2.05 eV, respectively calculated with respect to the vacuum energies. Both orbitals shift to lower energy levels upon adsorption (See Figure S9). We visualized the molecular orbitals using of the monolayer and gas phase molecule at r^2SCANh level using wannier90.[270] Lowest unoccupied molecular orbital (LUMO) and HOMO are mainly located on two diagonally opposite HB238 molecules in the monolayer configuration (See Figure B.13). The observed asymmetry in Figures 10 and 12 compared to the STM image (Figure 7a) arises from the C_2 symmetry of the adsorbed monolayer (also observed experimentally by Kny et al. [82], whereas the four bright lobes in the STM image result from the tert-butyl groups being positioned closer to the STM tip, enhancing their apparent intensity. Only with r^2SCANh we observe an additional molecular orbital close to Fermi level which corresponds to the LUMO (see Figure B.12). By comparison to the applied voltage in the STM experiment, we conclude that the observed contrasts are mainly due to this orbital. Since the orbital energy is below the Fermi level, this indicates a charge transfer from surface to HB238, as already seen from the CDD in Figure 10. This charge transfer further stabilizes the adsorbate structure.

Table B.3: Calculated HOMO and HOMO-1 energy levels (eV) of the HB238 monolayer adsorbed on the Ag(100) surface. Orbital energies were calculated relative the fermi level.

Method	HOMO	HOMO-1
PBE + U(4.5eV)	-1.35	-2.41
B3LYP	-2.05	-3.26
HSE06	-1.81	-3.03
r^2SCANh	-1.80	-2.90
Exp (UPS)[82]	-1.77	-2.71

B.3.6 X-ray absorption spectra

We compared calculated X-ray absorption spectra (XAS) and experimental XPS results because the calculation of XPS is presently not implemented in VASP. XAS corresponds to the excitation of core electrons to valence states, while XPS measures the binding energies of core electrons with respect to the Fermi level. Therefore, the absolute positions of the peaks are not comparable, and we focus on the relative positions of selected peaks. Here, we selected the S 2p states since their positions are sensitive to intramolecular charge transfer between the donor and the acceptor units of HB238, and between the Ag surface and the molecule. The calculations were performed using PBE/D3 with the standard PAW potentials, an energy cutoff of 400 eV, and a $2 \times 2 \times 1$ Monkhorst-Pack Grid.

A comparison of the calculated XAS spectra for the monolayer adsorbed on Ag and bulk HB238 (representing multilayers of HB238) reveals distinct differences in the S 2p orbitals. The crystallographic information for the HB238 bulk (CCDC 2073438) is shown in Table S5. In contrast, for

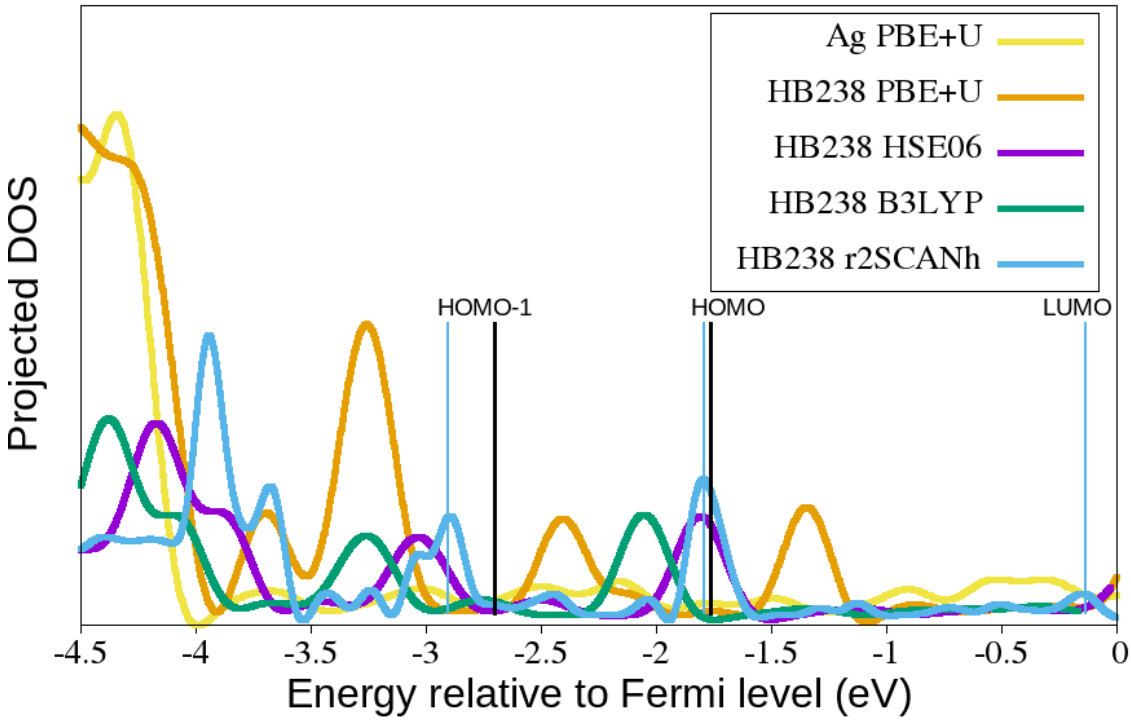


Figure B.12: Projected density of states plot for the Ag surface calculated using PBE +U(4.5eV) (light blue line), and monolayer of HB238 molecules adsorbed on an Ag(100) surface calculated using: PBE+U (orange), HSE06 (purple), B3LYP (green) and r^2SCANh (yellow). The experimental peaks for HOMO and HOMO-1 energies as seen in the UPS data [82] are indicated by the bold solid black lines. The calculated HOMO, HOMO-1 and LUMO using r^2SCANh are indicated with blue straight lines.

the HB238 bulk configuration, the XAS peaks of the donor and acceptor S are present at 160.6 and 160.2 eV, respectively, with a larger energy difference of 0.4 eV, reflecting more distinct electronic environments for the sulfur atoms within the bulk structure (See Figure D.10 a). In the monolayer, the calculated maxima of XA spectra of the donor and acceptor sulfur atoms are at 158.8 and 158.5 eV, respectively (See Figure D.10 b). The signals are very close in energy, with a small energy difference of 0.25 eV and a significant overlap. This indicates more equilibrated charge states among the sulfur atoms, likely due to the interaction with the Ag surface. This is slightly different from experiments, where the donor and acceptor S2p peaks coincide for the monolayer structure.[82] However, the trends that these peaks are closer in energy compared to the bulk coincide and hint to surface-mediated charge transfer from donor to acceptor units.

B.4 Conclusions

A theoretical investigation was performed for the monolayer adsorption and characterization of HB238 merocyanine on the Ag(100) surface. In agreement with previous experimental structure determinations, we identified a face-on adsorption configuration with the molecular plane parallel

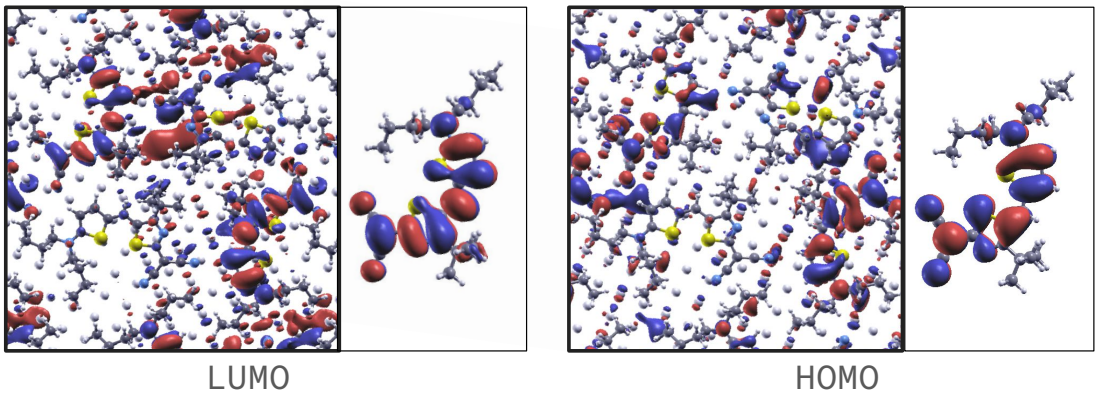


Figure B.13: Visualization of the LUMO (left) and HOMO (right) orbitals of HB238 monolayers on Ag(100) and gas phase HB238 molecule, calculated using r2SCANh method. The positive and negative orbital lobes are shown in red and blue color, respectively.

to the surface as the minimum structure. Although the HB238 structure is considerably changed upon adsorption, the energetic destabilization is relatively small due to the high flexibility of the molecule. The monolayer is stabilized by electrostatic, polarization and dispersion interactions between the HB238 molecules and the surface, and hydrogen bonds between the molecules leading to an arrangement that is different from the bulk structure. The calculated adsorption height analysis revealed that the donor sulfur atom of HB238 is positioned at approximately 3.09 Å above the Ag surface, while the acceptor sulfur atom is closer, positioned at 2.88 Å. The calculated PDOS was compared to experimental UPS results, and the meta-GGA hybrid function r^2SCANh shows a good correlation between the calculated and experimental electronic structures. The theoretical results indicate that the bright features of STM are due to the partly filled LUMO of the adsorbed HB238. The calculated and measured STM images are very similar indicating that the global minima search was successful and describes the experimental adsorbate structure. Furthermore, the calculated XAS spectra and CDD plot indicate that the donor and acceptor sulfur atoms equilibrate their net charge upon adsorption, and this charge redistribution is mediated by interactions with the Ag surface.

B.5 Supporting Information

Supporting Information (SI) Includes: Ag layer convergence tests, adsorption energy profiles of HB238 conformers on Ag(100), crystallographic parameters of the HB238 bulk structure, and relative energy profiles from VASP (CG method) and ASE (PreconLBFGS algorithm). Hydrogen bonding interactions in the HB238 monolayer, atomic charge distributions in the gas phase and adsorbed states,

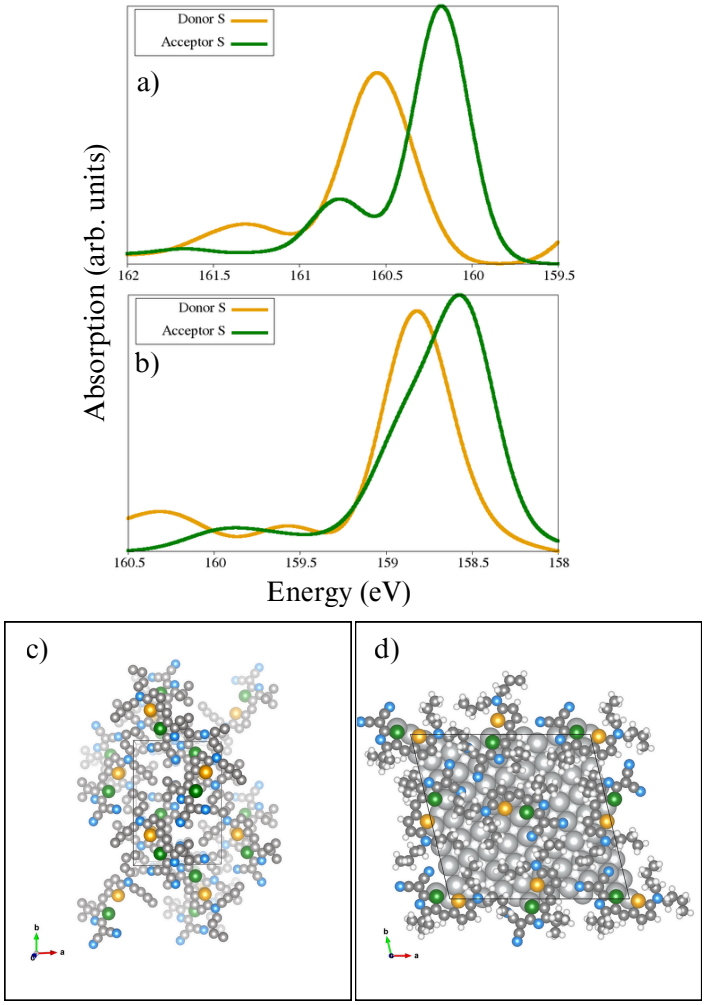


Figure B.14: Calculated X-ray absorption spectra for a) HB238 bulk crystal structure, and b) the monolayer adsorbed on Ag(100), model A. The images, c and d) represent the bulk and adsorption complex of HB238 where the donor and acceptor S atoms are visualized using orange and green color, respectively.

LUMO and HOMO of the monolayer, atomic charges of HB238, projected density of states (PDOS) using PBE, and the molecular orbital diagram of HB238 in gas phase and upon adsorption on Ag(100) are also available in SI.

B.6 Acknowledgements

The authors gratefully acknowledge funding from the DFG (RTG-2591 *TIDE Template-designed Organic Electronics*), and the computer resources from the HPC clusters Bonna, Bender and Marvin of the University of Bonn.

APPENDIX C

Machine Learning Accelerated Global Search for Adsorption Geometries of Merocyanine molecule on Hexagonal Boron Nitride

Ritu Tomar[†], Thomas Bredow[†]

Received: 31 October 2025
Published online: 17 February 2026

Adapted with permission [¶] from
R. Tomar and T. Bredow, *Machine Learning Accelerated Global Search for Adsorption Geometries of Merocyanine molecule on Hexagonal Boron Nitride*, *J. Comput. Chem.* **6** (2026), e70332
DOI:10.1002/jcc.70332

Own Contributions

- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to Wiley

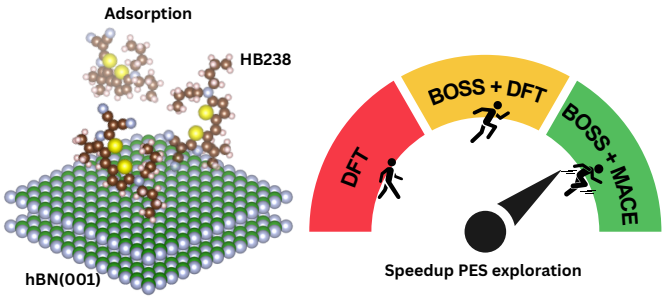


Figure C.1: Table of content graphic (ToC)

Abstract The adsorption of the merocyanine dye HB238 on hexagonal boron nitride (hBN) was investigated using a machine learning (ML) assisted global search strategy. A series of MACE machine learning interatomic potentials with higher-order equivariant message passing were finetuned on density functional theory (DFT) reference data for single and dimer adsorbate configurations, providing accurate surrogate models for the potential energy surface. The Bayesian Optimisation Structure Search (BOSS) was used to search over translational and rotational degrees of freedom of the adsorbed molecules, followed by full PBE/D3 optimisation of the most promising structures. The ML-accelerated search revealed that HB238 prefers to adsorb face-on hBN with the sulfur atoms located near hollow sites; however, the molecule exhibits no strong site selectivity, giving rise to a broad ensemble of configurations within energies 0.1 eV above the global minimum. When two HB238 molecules are adsorbed, they align parallel to each other and lie flat on the surface. Overall, our results demonstrate that combining finetuned ML potentials with Bayesian optimisation enables an efficient and accurate exploration of complex adsorption landscapes and provide fundamental insights into the physisorption of dipolar dyes on 2D insulators. This combined MACE×BOSS approach can be easily extended to investigate organic molecular aggregates on 2D surfaces.

C.1 Introduction

Understanding molecular adsorption on solid surfaces is a central problem in catalysis, sensor design, and material functionalization [184–187]. Two-dimensional (2D) materials such as hexagonal boron nitride (hBN) and graphene have attracted significant interest due to their exceptional mechanical strength, thermal stability, and wide band gap, making them ideal candidates for various surface chemistry applications [78, 190]. In particular, hBN and graphene provide atomically flat, chemically inert templates on which organic molecules can form highly ordered films with well-defined epitaxial relationships [188, 271, 272]. Such ordered growth has been reported for rod-like conjugated molecules such as para-hexaphenyl and pentacene, which align along high-symmetry directions of the hBN lattice [188, 189]. However, despite extensive work on related systems, adsorption of merocyanine dyes on hBN has not been studied experimentally or theoretically to our knowledge, leaving an important gap in understanding the templated growth of dipolar chromophores on insulating 2D materials.

The merocyanine dye HB238 (2-[5-(5-dibutylamino-thiophen-2-yl-methylene)-4-*tert*-butyl-5H-thiazol-2-ylidene]-malononitrile) is a prototypical donor–acceptor chromophore with a large ground-state dipole moment and tunable optoelectronic properties [81, 82]. On metallic substrates, HB238 has

been observed to chemisorb via its sulfur atoms, forming ordered assemblies such as chiral tetramers on Ag(100) [82]. In contrast, adsorption on inert van der Waals surfaces such as hBN(001) is expected to be governed by weaker dispersive and electrostatic interactions, potentially allowing the formation of ordered physisorbed monolayers without strong hybridization effects. Motivated by previous studies of merocyanine and spiropyran derivatives on graphene [273], we investigate here the fundamental adsorption behavior of HB238 on hBN.

Accurately predicting stable adsorption geometries of large and flexible molecules on surfaces is computationally demanding. The potential energy surface (PES) governing molecule–surface interactions is high-dimensional and rugged, containing numerous local minima [274]. Traditionally, global optimization and PES sampling methods such as simulated annealing [275], genetic algorithms [276], basin hopping [277], minima hopping [278], metadynamics [279], and Monte Carlo–minimization [280] have been used to address this challenge, but they are often inefficient for complex organic adsorbates.

Recent advances in machine learning (ML) have revolutionized this domain by providing accurate surrogate models for energy and force prediction at a fraction of the cost of density functional theory (DFT). In particular, ML-based interatomic potentials such as Message Passing Neural Network (MPNN) architectures have demonstrated DFT-level accuracy for a variety of systems [101, 281]. Among them, the MACE stands out by incorporating higher-order equivariant interactions and achieving state-of-the-art performance in capturing short-range physics in molecules and materials [101, 282]. Recent work has demonstrated that MACE can be effectively finetuned with DFT data to pre-relax adsorption geometries, discovering new stable configurations while reducing the computational cost by up to 75% compared to brute-force DFT screening [108].

To efficiently explore the PES using such ML surrogates, Bayesian Optimization (BO) has emerged as a powerful tool for global structure search. The BOSS framework [92, 181, 182] employs Gaussian Process (GP) regression to model the energy landscape and iteratively select new sampling points that balance exploration and exploitation [283]. BOSS has been successfully applied to identify low-energy adsorption configurations of organic molecules on metallic and 2D surfaces [181].

In this work, we present a hybrid MACE×BOSS framework that combines a MACE-MP-0 model, fine-tuned on PBE/D3 data, with the BOSS Bayesian optimization method to efficiently explore the adsorption landscape of the HB238 molecule on hBN. The MACE-MP-0 model, originally trained on PBE and PBE+U datasets without dispersion corrections, is finetuned here to include dispersion interactions, which play a crucial role in adsorption phenomena. Within this framework, BOSS performs a five-dimensional global optimization—two translational and three rotational degrees of freedom—to identify the most stable adsorption geometry. The MACE × BOSS approach achieves a substantial computational speed-up compared to direct DFT-based optimizations while retaining near-DFT accuracy. By combining MACE energy calculation and Bayesian optimization, we overcome the limitations of brute-force DFT screening, provide the first systematic study of HB238 adsorption on an insulating 2D substrate, and establish a broadly applicable framework for ML-accelerated exploration of molecule–surface interactions.

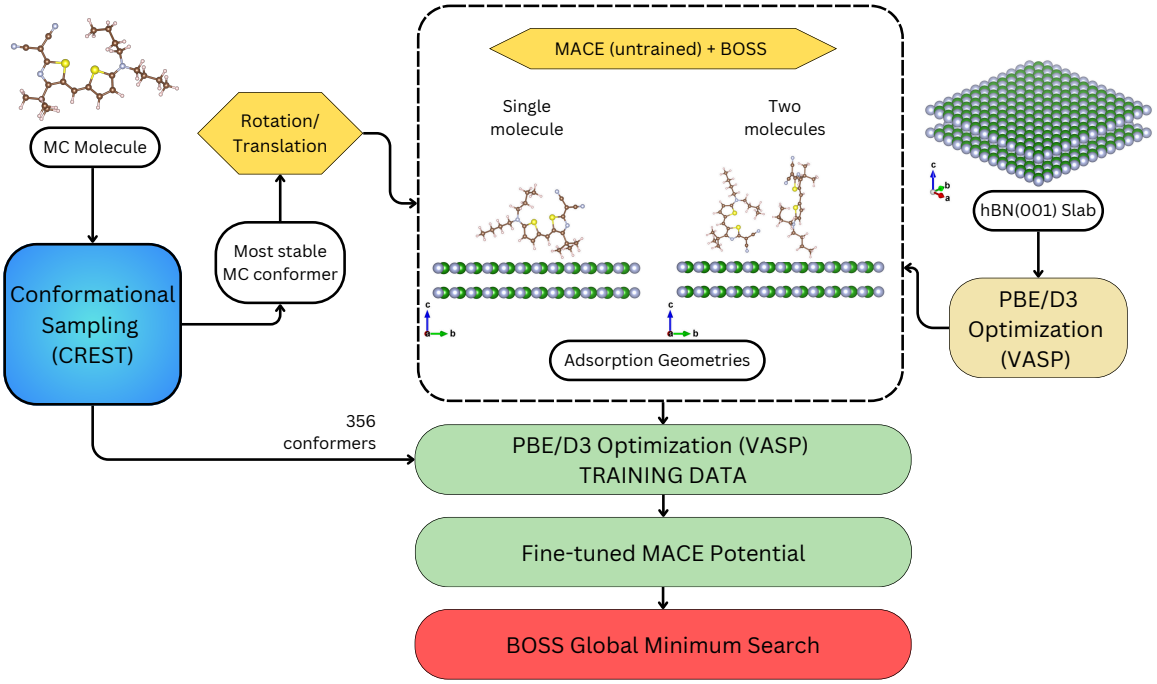


Figure C.2: Workflow of the global minimum search of the adsorption structure of HB238 molecule on hBN(001) surface.

C.2 Computational methods

C.2.1 Workflow

The global minimum search for the adsorption structure of the HB238 molecule on the hBN(001) surface is shown in Figure C.2. Initially, the conformational landscape of the HB238 molecule was explored using the Conformer-Rotamer Ensemble Sampling Tool (CREST)[137]. We performed iMTD-sMTD metadynamics-based conformer sampling with the GFN2-xTB[135] semiempirical method as implemented in CREST. After identifying the most stable gas-phase conformer among the 356 sampled structures, this structure is used as the adsorbate on hBN(001) surface for adsorption structure generation using MACE-MP-0 and BOSS. 1191 single molecule and 1163 dimer adsorption structures were generated and then optimized with PBE/D3[132, 220] in VASP,[128, 129] creating a comprehensive training dataset. This training dataset generation procedure is implemented as a Python module called "adsgen".The energy and gradients of the above mentioned PBE/D3 optimized structures structures were used to train the MACE[101] potential. Finally, the finetuned MACE potential, coupled with BOSS,[92] was applied to efficiently perform a global minimum search on the potential energy surface (PES).

C.2.2 Bayesian Optimization Structure Search

BOSS was employed to perform global optimization of HB238 adsorption configurations on the hBN(001) surface using MACE-trained potentials.[92] For single-molecule adsorption, a 5-dimensional configurational space was explored, consisting of x and y lateral translations from 0-2.5 Å on the

surface and three Euler rotational angles (α, β, γ) of the HB238 molecule. Each angle was sampled uniformly from 0° to 359°. The search was performed using the finetuned potentials as the surrogate model for the potential energy surface. Geometry optimization of each sampled structure was conducted using the PreconLBFGS[174] optimizer in ASE[284] with a force convergence threshold of 0.01 eV/Å. A total of 1000 configurations were evaluated, comprising 20 initial random samples and 980 Bayesian optimization iterations.

The dimer search explored a 10-dimensional configurational space comprising the combined translations and rotations of two HB238 molecules adsorbed on hBN. Each molecule was allowed two in-plane translations (x, y) and three Euler rotations (α, β, γ). To ensure comprehensive sampling while avoiding physically unrealistic placements, the translational bounds were derived from extremal positions of two HB238 molecules stacked in various orientations—vertical, side-by-side, rotated, diagonal, and overlapping — within the hBN unit cell (see Figures S1 and S2). From these reference configurations the centre-of-mass coordinates were restricted to $[4.00, 7.50]$ Å in x and $[2.00, 7.50]$ Å in y for Molecule 1, and $[4.00, 23.50]$ Å in x and $[4.00, 20.50]$ Å in y for Molecule 2. The rotational angles were sampled in discrete 15° intervals from 0° to 345°. For each configuration, the predicted α, β, γ values from BOSS were rounded to the nearest multiple of 15° to reduce the number of redundant configurations and maintain rotational symmetry. Both molecules were initially placed 2.5 Å above the hBN surface, and for vertically stacked configurations the second molecule was offset by an additional 2.5 Å above the first molecule. This vertical spacing eliminated unphysically small distances between the atoms of the two molecules and allowed us to discard implausible structures before optimisation.

C.2.3 MACE Training

Three independent MACE interatomic potentials were refitted to model the PES of HB238 adsorption on the hBN(001) surface. All training structures were optimized at PBE/D3 level prior to MACE model refinement. The first model, referred to as **MACE-1M**, was re-fitted on DFT energies and forces of 1191 structures consisting of single HB238 molecules adsorbed on the surface. The second model, **MACE-1M2M**, used an extended dataset of 2354 structures incorporating both single- and two-molecule adsorption configurations. The third model, **MACE-1M2M+C**, was trained on a combined dataset of 2660 structures, which included all previous adsorption configurations along with HB238 gas-phase conformers generated via CREST and optimized at PBE/D3 level. The relative energy of 356 conformers obtained with CREST is shown in Figure S3.

All models were re-fitted starting with the "medium" foundation model. The training was performed for a maximum of 2000 epochs with early stopping enabled (patience = 20). The optimizer used was Adam with the AMSGrad variant and a fixed learning rate of 0.001. The loss function was defined as a weighted sum of energy and force terms, with energy and force weights of 10.0 and 1000.0, respectively. Energy reference values were computed using per-atom averages.

The MACE models were trained with a batch size of 4 for both training and validation, leveraging CUDA-enabled GPUs in a distributed setting across four nodes, each equipped with four GPUs. To prevent overfitting, gradient clipping was applied with a threshold of 1.0. Exponential Moving Average (EMA) decay with a rate of 0.995 was employed to smooth parameter updates during backpropagation. The precision for all calculations was maintained at double precision (float64) to ensure numerical stability during gradient updates. Validation was performed after every epoch with a 5% holdout fraction, and the error was measured using Per Atom Mean Absolute Error (PerAtomMAE).

Each of the three refitted MACE models was subsequently integrated into a BOSS to perform global

structure optimization over the predicted PES, see previous section.

C.2.4 Density Functional Theory

All density functional theory (DFT) calculations were performed with VASP 6.5.1[128, 129], employing the Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional within the generalized gradient approximation together with Grimme's D3 dispersion correction[132]. Core–electron interactions were described using the projector augmented-wave (PAW) method[262, 285], with PAW-PBE pseudopotentials from the standard VASP library. A plane-wave kinetic energy cutoff of 450 eV was applied. Electronic occupancies were determined using Fermi–Dirac smearing with a width of 0.01 eV. The hBN(001) surface was modeled with a lattice constant of a=2.51 Å (optimized at the PBE/D3 level) as a periodic slab consisting of two atomic layers. The experimental lattice constant of hBN is 2.50 Å.[286] For adsorption studies, a 12 × 12 × 1 supercell was used for single-molecule adsorption and an 18 × 18 × 1 supercell for dimer configurations. Given the large surface dimensions, the Brillouin zone was sampled at the Γ-point only. Ionic relaxations were performed with the conjugate gradient (CG) algorithm, where only the adsorbed HB238 molecule was relaxed while the hBN surface atoms were kept fixed.

C.3 Results and Discussion

C.3.1 MACE fine-tuning

The fine-tuning performance of the three MACE potentials shown in Figure C.3 shows a clear trend of rapid convergence. For the single-adsorbate model (MACE-1M) the energy mean absolute error (MAE) drops below 1 meV/atom and the force MAE below 5 meV/Åwithin the first few hundred epochs, while the validation loss closely tracks the training loss throughout, indicating that the potential fits the single-molecule adsorption data without overfitting. When the training set is expanded to include two-molecule configurations (MACE-1M2M), the initial energy and force errors are larger because the model must learn dimer interactions, but both errors decline steadily and force errors approach those of the single-molecule model by the end of training and the energy errors drop below 2.5 meV/atom. The most comprehensive model (MACE-1M2M+C), which also includes gas-phase conformers, starts with the highest errors but quickly reduces the energy MAE to below 2.5 meV/atom and the force MAE to roughly 2 meV/Å, with a small oscillation around 500 epochs reflecting the assimilation of the conformer data; the training and validation losses remain aligned, demonstrating that the broader dataset does not compromise the model's accuracy.

C.3.2 Global Minima Search for Adsorption Structure with BOSS

To identify the most stable adsorption geometries of the HB238 molecule on the hBN surface we employed the BOSS[92] method, which combined a machine–learning potential with a Bayesian acquisition function to navigate the high-dimensional configurational space efficiently. For each configuration returned by BOSS, the energy provided by the fine-tuned MACE models was used to rank candidate structures, and the energy configurations were then re-evaluated with DFT to obtain accurate reference energies.

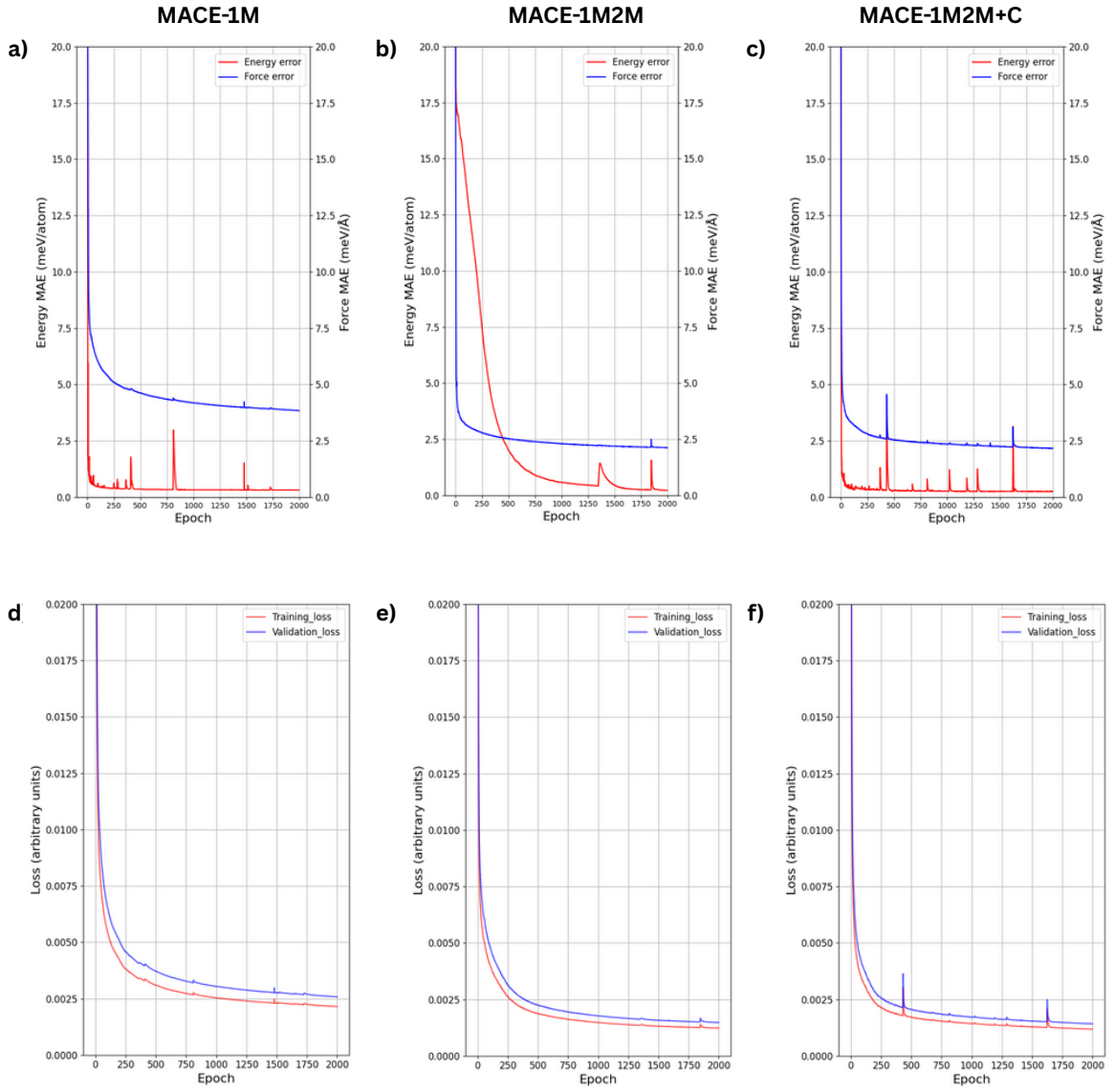


Figure C.3: Learning curves for the three refitted MACE potentials developed in this work. Panels (a–c) show the mean absolute error (MAE) of predicted energies in meV/atom (red) and forces in meV/Å (blue) versus training epoch for models trained on single adsorbate (MACE-1M), single molecule and dimer adsorbate (MACE-1M2M), and single, dimer and gas-phase conformer (MACE-1M2M+C) finetuning datasets, respectively. Panels (d–f) plot the corresponding training (red) and validation (blue) losses versus epoch for each model.

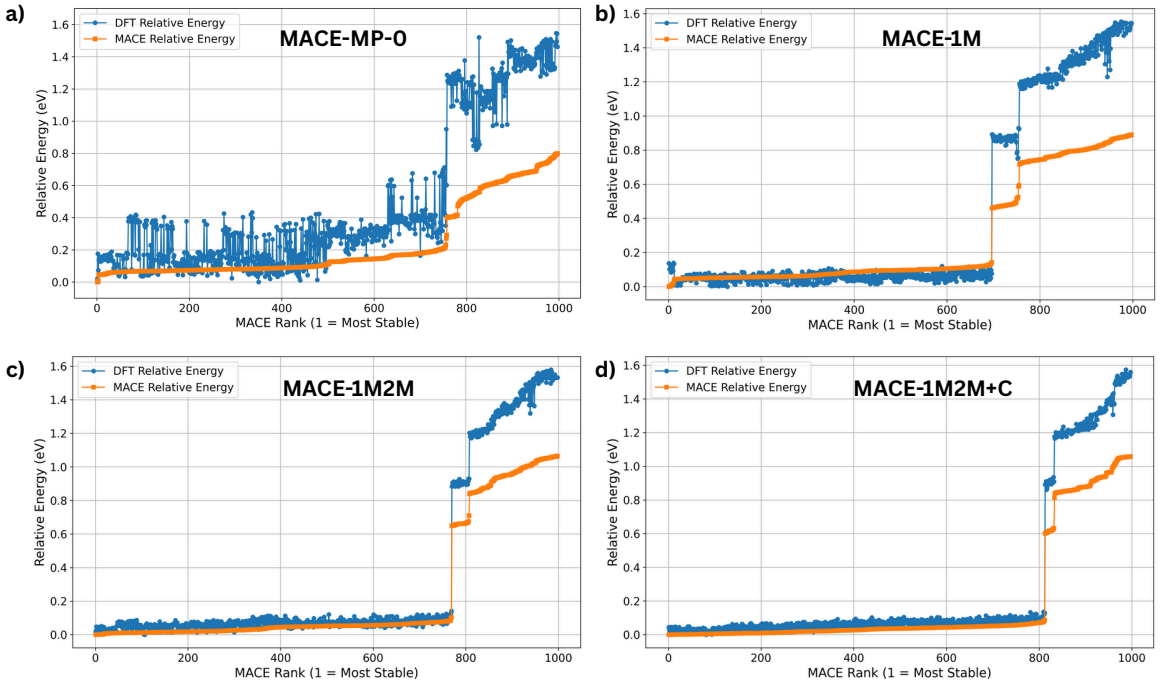


Figure C.4: Ranking of relative energies for single-molecule adsorption obtained with BOSS using the foundation model a)MACE-MP-0 and the fine-tuned models: b)MACE-1M, c)MACE-1M2M and d)MACE-1M2M+C. The surrogate energies (orange dots) are compared to respective DFT reference energies (blue dots).

Single-molecule adsorption

In the single-adsorbate case, the search space comprised two in-plane translations and three rotational degrees of freedom as described above. Figure C.4 compares the BOSS rankings obtained with the foundation model (MACE-MP-0) and the three fine-tuned models. Each panel plots the DFT relative energy (blue points) and the MACE surrogate potential energy (orange line) as a function of the MACE rank (lower rank corresponds to more stable structures). The PBE/D3 energies shown in Figure C.4 are obtained from single point calulations of the structure predicted from MACE×BOSS approach, obtained for the origional MACE-MP-0 and the three refitted models, respectively. The foundation model shows little correlation between the MACE-MP-0 and the PBE/D3 energies: PBE/D3 energies vary widely across the entire rank range and the MACE-MP-0 model fails to distinguish low-energy structures. In contrast, the fine-tuned models substantially improve the agreement. For the MACE–1M model, the first 700 ranked structures lie within 0.2 eV of the global minimum, and the finetuned model energy increases smoothly thereafter. This indicates that the model reliably identifies a large family of low-energy adsorption configurations. The MACE–1M2M and MACE–1M2M+C models exhibit similar trends, albeit with slightly smaller deviations between the MACE predicted and PBE/D3 energies at higher ranks.

Inspection of all structures with 0.2 eV energy difference above the lowest structure obtained with the three finetuned models shows that the HB238 molecule lies flat on the hBN surface. The most stable configuration was obtained using the MACE-1M2M+C model and is shown in Figure C.5. In this configuration, both sulfur atoms are positioned at the hollow sites. The distribution of sulfur

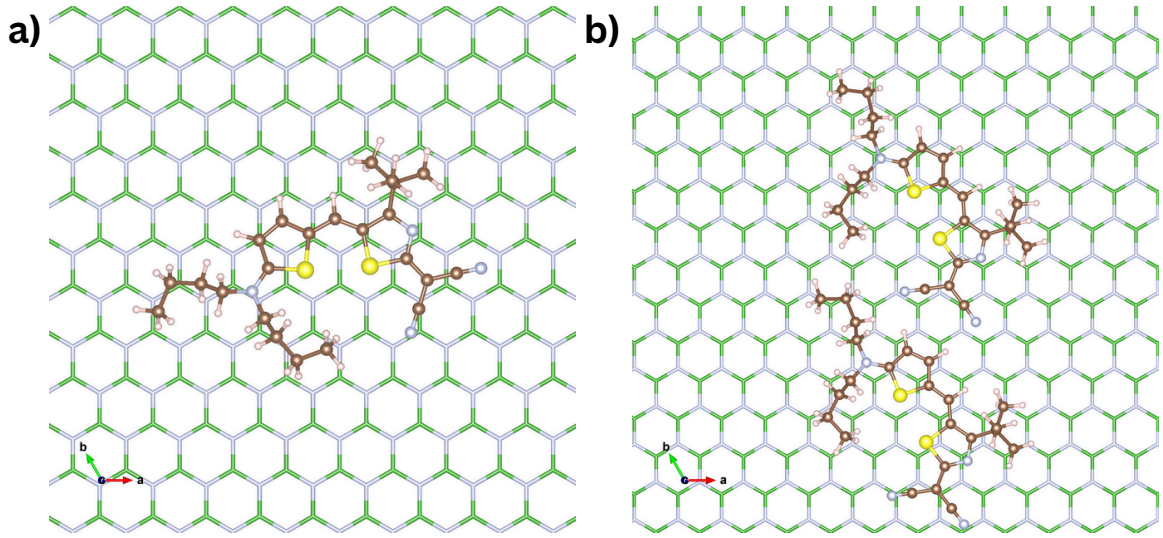


Figure C.5: Most stable adsorption motifs from a) 5D and (b) 10D BOSS search obtained using MACE-1M2M+C model and relaxed at the PBE/D3 level using DFT. The hBN slab is shown using hexagonal rings in blue and green color and the S, C , N and H atoms are shown using yellow, brown, light blue and light pink color respectively. Only the first layer of the two layer hBN slabs is shown.

positions extracted from all structures within ≤ 0.1 eV relative energy (773 structures) with respect to the global minimum (Figure C.6a) shows the HB238 molecule lacks site-specific bonding, i.e. the molecule is mobile on the surface. Ten representative monomer adsorption structures within 0.1 eV of the minimum are shown in Figure S4, with their adsorption energies summarized in Table S1.

Dimer adsorption

To study the adsorption of a second molecule as the second step of a deposition process, 10-dimensional configurational space was explored which included the combined translations and rotations of two HB238 molecules adsorbed on hBN. Figure C.7 illustrates the BOSS ranking for this case. The foundation model (MACE-MP-0) fails to correlate with the DFT energies and yields a wide spread of predicted low-energy structures. Fine-tuning on adsorption data dramatically improves the ranking. The MACE–1M model captures the low energy plateau but fails to identify the lowest energy structures. From 550 similar energy structures predicted by the model very few are actual minima, this is expected since the model did not contain any dimer structures. The MACE-1M2M and MACE–1M2M+C models improve the identification of dimer minima structures. Overall, the surrogate-guided search locates a rich set of 175 dimer configurations with energies within 0.1 eV of the global minimum.

The lowest-energy dimer identified by BOSS is shown in Figure C.5b. The two HB238 molecules adsorb flat on the surface and align parallel to each other. Ten representative structures within 0.1 eV of the minimum are shown in Figure S5, with their adsorption energies summarized in Table S2. The lateral interactions of two molecules on hBN surface stabilized by intermolecular hydrogen bonding as shown in Figure S6. As in the single-molecule case, the sulfur positions from the dimer configuration (175 structures within 0.1 eV) also do not show any site preference (Figure C.5b), but it is evident that the molecules do not prefer to stack on each other but adsorb flat on the surface, side by side.

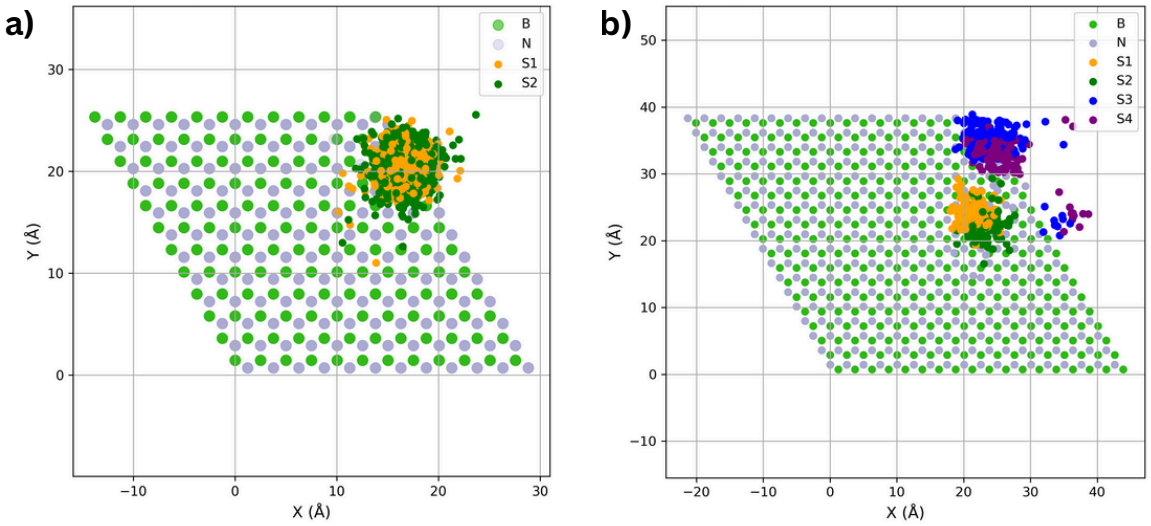


Figure C.6: Distribution of sulfur atom positions for all structures with 0.1 eV relative energy of the global minima obtained from the BOSS ensambles with MACE-1M2M+C. (a) Single-molecule search: the donor S (orange points) and acceptor atoms (green points) (b) Dimer search: four clusters (coloured points) correspond to the four sulfur atoms in the dimer, again localising near nitrogen sites.

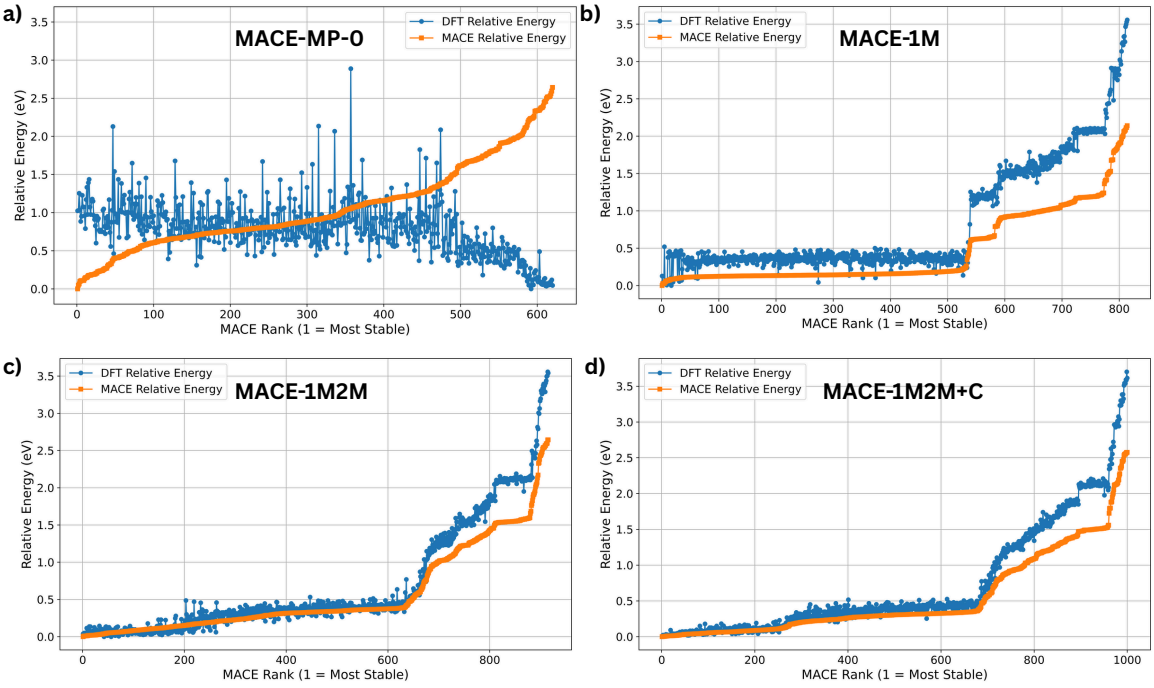


Figure C.7: Ranking of relative energies for 10D dimer adsorption obtained with BOSS using the foundation model a)MACE-MP-0 and the fine-tuned models: b)MACE-1M, c)MACE-1M2M and d)MACE-1M2M+C. The surrogate energies (orange dots) are compared to respective DFT (PBE/D3) reference energies (blue dots) obtained from single point calculations of the MACE×BOSS strucures.

C.4 Adsorption energy

Table C.1: Adsorption energy per molecule E_{ads} (eV) for the most stable monomer and dimer adsorption configuration calculated at PBE/D3 level

Adsorption Configuration	E_{ads} (eV)
Monomer	-1.69
Dimer	-1.95

The adsorption energies per molecule were computed as $E_{\text{ads}} = (E_{\text{HB238/hBN}} - E_{\text{hBN}} - N\,E_{\text{HB238}})$/N, where $E_{\text{HB238/hBN}}$ is the total DFT energy of the relaxed adsorbate–slab system, E_{hBN} is the energy of the bare slab, E_{HB238} is the energy of an isolated HB238 molecule in the gas phase, and N is the number of adsorbed molecules. A negative value indicates exothermic adsorption.

Table C.1 summarises the global minimum adsorption energies for the monomer and dimer configurations. E_{ads} of the most stable monomer adsorption motif is –1.69 eV, while the E_{ads} of most stable dimer configuration is –1.95 eV. The latter is 0.26 eV lower, indicating that lateral interactions between the two molecules provide an additional stabilisation. This extra stabilisation arises electrostatic interaction and hydrogen bonding between the cyano group of one molecule and the hydrogen atoms from the pi-conjugated backbone of the adjacent molecule (Figure S6). The stronger adsorption of HB238 compared to trans-azobenzene (–1.49 eV on hBN)[**Szitas2020AdsorptionRh111**] is due to the larger molecular size and polarizability of the merocyanine dye. Within the ensemble of 700 low-energy structures found by BOSS, the adsorption energies vary by less than 0.2 eV, highlighting the flatness of the potential energy surface and suggesting high mobility of the molecule on hBN.

C.5 Electronic Structure Analysis

The electronic interaction between HB238 and the hBN substrate was evaluated through molecular-projected density of states (PDOS) and charge density difference (CDD) analyses, as presented in Figure C.8.

The PDOS projected onto HB238 (Figure C.8a) indicates that the frontier orbitals remain sharp and well defined upon adsorption. The HOMO of the HB238 molecule is close to the Fermi level of the combined HB238/hBN system. The persistence of narrow peaks demonstrates that the molecular orbitals retain their localized character with negligible mixing into substrate states, consistent with weak electronic coupling to the hBN lattice.

The CDD maps (Figures C.8b and C.8c, side and top views) provide spatial insight into charge redistribution at an isovalue of $5\text{x}10^{-4}\text{e}/^{3}$. Only minor regions of charge accumulation and depletion are observed around cyano groups of HB238, while the hBN surface remains largely unaffected. This distribution confirms the absence of covalent bonding or significant charge transfer.

C.6 Conclusions

A machine-learning-accelerated workflow was developed to explore the high-dimensional adsorption landscape of the merocyanine dye HB238 on hexagonal boron nitride. MACE interatomic potentials,

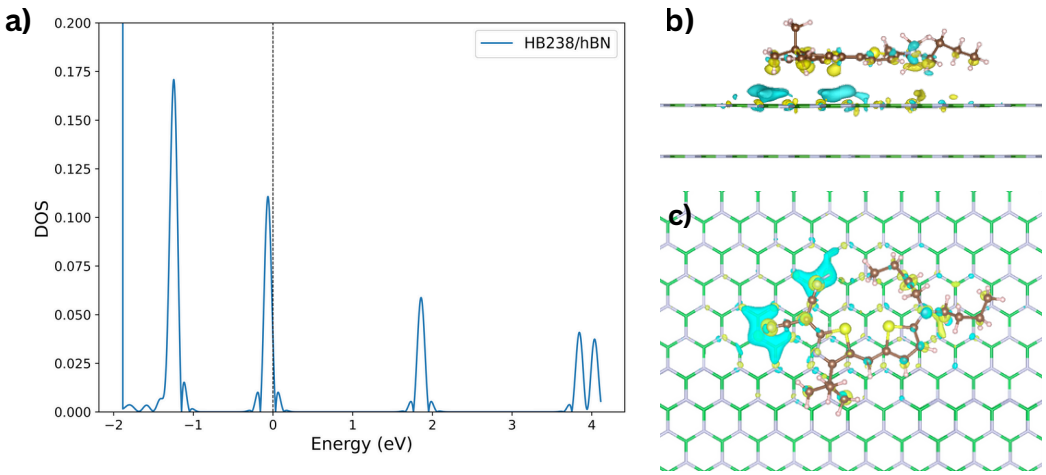


Figure C.8: a) PDOS of HB238 in the most stable adsorption motif on hBN surface, obtained from 5D BOSS search as shown in Figure C.5. The black dotted line indicates the Fermi level of the HB238/hBN system. b) Side view and c) top view of the charge density difference (CDD) at an isovalue of $5x10^{-4}e/^{3}$, illustrating charge redistribution upon adsorption. Yellow regions correspond to charge accumulation, while cyan regions denote charge depletion.

fine-tuned on PBE/D3 training data for single and dimer adsorbate configurations, were combined with the BOSS Bayesian optimisation framework to efficiently search the five- and ten-dimensional configurational spaces and identify stable adsorption motifs with chemical accuracy. The machine-learning surrogate potential enabled the sampling of thousands of candidate structures. The results indicate that HB238 adsorbs flat on hBN without strong site specificity, yielding a manifold of monomer and dimer adsorption configurations within 0.2 eV above the global minimum energy. Based on the present theoretical results it can be predicted that the growth of ordered monolayers of HB238 on hBN is only governed by intermolecular interaction and may be hampered by the lack of directing effects from the surface. This workflow is expected to be broadly applicable to other large organic adsorbates and to support the design of hybrid organic–2D material systems for optoelectronic and sensing applications.

C.7 Supporting Information

Supporting Information (SI) Includes: Adsorption configuration used to identify the bounds for BOSS 10D search, Illustration of the monomer and dimer HB238 adsorption structures on hBN surface with their adsorption energies.

C.8 acknowledgement

The authors gratefully acknowledge funding from the DFG (RTG-2591 *TIDE Template-designed Organic Electronics*), and the computer resources from the HPC clusters Bonna and Marvin of the University of Bonn.

APPENDIX D

HB238 Adsorption on Coinage Metal Surfaces: Insights from a combined Machine-Learning Interatomic Potential and Bayesian Optimization Framework for Interfacial Electronic Structure

Ritu Tomar[†], Thomas Bredow[†]

Received: 31 October 2025
Published online: 20 April 2026

Adapted with permission [¶] from
R. Tomar and T. Bredow, *HB238 Adsorption on Coinage Metal Surfaces: Insights from a combined Machine-Learning Interatomic Potential and Bayesian Optimization Framework for Interfacial Electronic Structure*, *J. Phys. Chem. C*, (2026)
DOI: 10.1021/acs.jpcc.5c07480

Own Contributions

- python scripting and implementation
- performing of all calculations
- interpretation of the results
- writing the manuscript

[†] Mulliken Center for Theoretical Chemistry, University of Bonn, 53115, Germany
[¶] Permission requests to reuse material from this chapter should be directed to the American Chemical Society

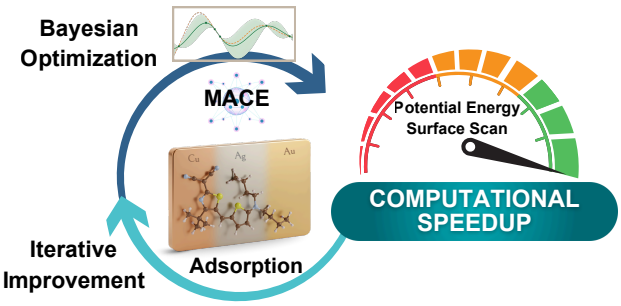


Figure D.1: Table of content graphic (ToC)

Abstract Organic donor–acceptor (D–A) molecules with permanent dipoles are attractive candidates for flexible electronics. Merocyanines represent a well-studied class of such systems, both experimentally and theoretically. In this study, the adsorption of a single merocyanine molecule on Cu(100), Ag(100), and Au(100) surfaces was investigated using a workflow that combines machine learning interatomic potentials with Bayesian optimisation. Interatomic potentials based on DFT data were fine-tuned for each substrate and embedded within a Bayesian optimisation scheme to efficiently explore lateral translations and rotations. Candidate geometries were subsequently refined via dispersion-corrected DFT relaxations, and the electronic structure was analysed through projected density of states and core-level spectra simulations.

The search identified flat-lying adsorption as the energetically preferred configuration across all coinage metals, with the strongest substrate–molecule coupling on Cu(100). Core-level spectra calculations showed that donor sulfur atoms consistently possessed higher binding energies than acceptor sulfur atoms, and that the donor–acceptor splitting narrowed upon adsorption, indicating significant charge redistribution at the interface. This study provides an efficient route to map complex adsorption landscapes and to elucidate how metal reactivity influences the electronic structure of dipolar organic molecules.

D.1 Introduction

Organic semiconducting molecules have emerged as suitable materials in the development of flexible and lightweight optoelectronic devices, with their molecular design offering extensive tunability in optical absorption, charge transport, and interfacial properties. [28] Within this class, donor–acceptor (D–A) molecules with large permanent dipole moments have gained attention due to their potential to control electronic level alignment at interfaces, a crucial factor in optimizing the performance of organic solar cells and transistors. [37, 38] The orientation and packing of such polar molecules in thin films significantly influences their optoelectronic properties, including the work function, exciton binding energy, and charge carrier mobility. [29, 36] Variations in molecular orientation from face-on to edge-on are shown in Figure D.2 for the merocyanine 2-[5-(5-dibutyl-amino-thiophene-2-yl-methylene)-4-tert-butyl-5H-thiazol-2-ylidene]-malononitrile, denoted as HB238. Such orientational differences can lead to substantial differences in energy level alignment, polarization screening, and interfacial dipole formation. [70, 287, 288]

Surface-sensitive experimental techniques including scanning tunneling microscopy (STM), low-energy electron diffraction (LEED), photoelectron spectroscopy (PES), X-ray photoelectron spectro-

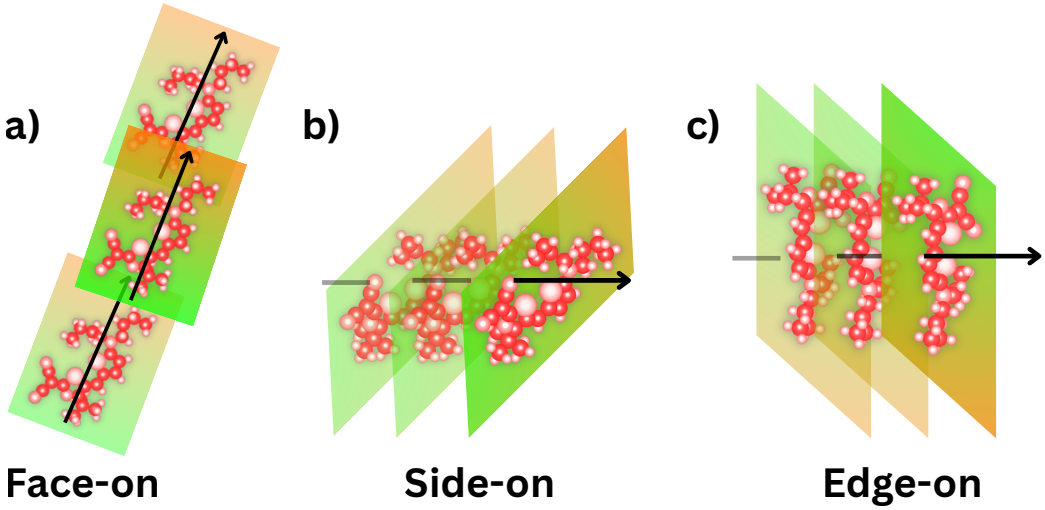


Figure D.2: Representative picture of molecular orientation and packing of the HB238 molecule in thin films. The molecular motif is represented in shades of red. Black arrows indicate the direction of the dipole moment vector. The color gradient from green to orange shows the orientation from acceptor to donor unit.

scopy (XPS) and X-ray absorption spectra (XAS) have offered extensive experimental insight into the nature of interactions between molecules and metal surfaces. [72, 73] However, these studies have predominantly focused on highly symmetric and rigid molecules such as rubrene, PTCDA, pentacene, and phthalocyanines, [289–292] which often form well-ordered layers on noble metal surfaces. These model systems have been widely used to investigate van der Waals interactions, epitaxial growth, and molecule-induced work function shifts. [293–295]

However, the growing complexity of organic semiconductors used in devices characterized by asymmetric structures, flexible side chains, and heteroatoms necessitates a deeper theoretical exploration of their adsorption behavior on technologically relevant surfaces. [296, 297] Merocyanine dyes represent a compelling class of such materials, exhibiting high dipole moments, intense optical absorption, and a rich variety of packing motifs. [59, 242, 298] The molecule HB238 exemplifies these features and is shown in Figure D.3. With its modular donor–acceptor structure, flexible π-conjugated core, and chemically active sulfur and nitrogen atoms, HB238 is particularly suited for studying the intricate balance between molecule–substrate and intermolecular interactions. [43, 81, 82, 299]

Recent experimental studies by Öcal et al. [81] and Kny et al. [82] have provided experimental insight into the interaction of HB238 with metal surfaces. Despite these important experimental efforts, a detailed theoretical understanding of the interface geometry and electronic structure using density functional theory (DFT) can offer valuable insight into the nature of molecule–metal bonding and the effect of intermolecular and molecule-surface interaction on the optoelectronic properties of HB238. [294, 300, 301] For similar systems, DFT has proven effective in reproducing adsorption energies and projected density of states (PDOS), and has been used to rationalize differences in reactivity among the coinage metals. [302, 303] Additionally, simulation of XAS allows comparison with element-specific experimental signals and helps characterize the local bonding environment. [299, 304, 305]

However, the high configurational flexibility of HB238 introduces a large and complex adsorption

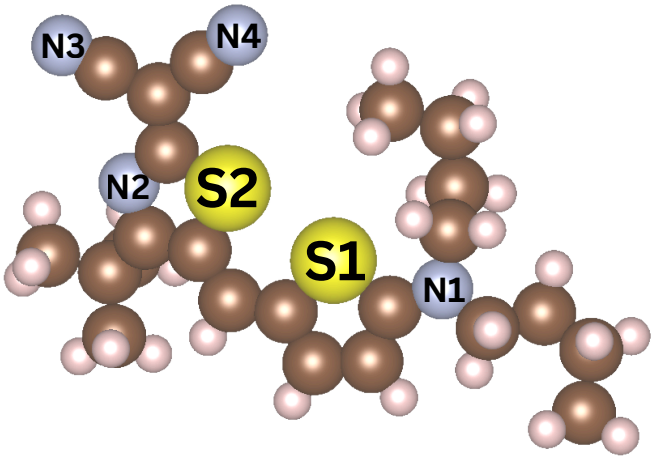


Figure D.3: Molecular structure of the merocyanine HB238. The donor sulfur atom (S1) is located in the thiophene ring, and the acceptor sulfur (S2) in the thiazole ring. The donor nitrogen (N1) is part of the dialkylamino group, while the acceptor nitrogens comprise N2 in the thiazole ring and N3 and N4 in the cyano groups. C, H, N, and S atoms are shown in brown, light pink, blue, and yellow color, respectively.

phase space, making global geometry exploration with standard DFT-based structure optimization techniques computationally unfeasible. To address this challenge, we employ the machine-learned interatomic potential (MLIP) MACE,[101] fine-tuned with DFT energies and forces of the HB238:metal systems, to evaluate a wide range of adsorption geometries. This MLIP is coupled with the Bayesian Optimization Structure Search (BOSS)[181] algorithm, which efficiently explores the five-dimensional configurational space involving lateral shifts and molecular rotations, enabling the identification of global and low-energy local minima. This approach complements experimental insight by enabling quantitative prediction of stable adsorption geometries and associated electronic features with high accuracy and efficiency.

In this work, we perform a comprehensive theoretical investigation of the HB238 molecule adsorbed on the (100) facets of Ag, Au, and Cu surfaces. Using the MACE/BOSS approach followed by dispersion corrected DFT calculations of selected structures, we optimize adsorption geometries and quantify adsorption energies, sulfur–metal bond lengths, and nitrogen–metal bonding. We analyze the PDOS to assess orbital hybridization and energy level shifts, and simulate XAS to probe the local electronic environment around key atoms. By comparing the results across the three substrates, we elucidate the role of metal reactivity in terms of interfacial bonding, molecular orientation, and electronic structure.

D.2 Computational Details

D.2.1 MLIP training

To facilitate efficient sampling and accurate evaluation of adsorption geometries for HB238 on coinage metal surfaces, we finetuned the original MACE-MP-0 model[180] separately for each metal

substrate—Cu(100), Au(100), and Ag(100).[101] The fine-tuned models allowed efficient sampling of the PES for HB238 adsorption on metal surfaces.

Each model was trained independently using DFT-generated datasets that included a diverse set of adsorption configurations with varying molecule–surface orientations. For Cu(100), a total of 1511 structures were used for training and 167 for testing. The Au(100) dataset comprised 1394 training structures and 154 test structures, while the Ag(100) model was trained on 1475 structures and evaluated on 163 test structures. All models achieved sub-meV accuracy in energy predictions and force errors on the order of a few meV/Å, as shown in Table D.1 with final test set MAE values for energy and forces of 0.7 meV/atom and 5.8 meV/Å, respectively, for Cu(100), 1.3 meV/atom and 2.3 meV/Å for Au(100), and 0.6 meV/atom and 2.4 meV/Å for Ag(100).

Fine-tuning was performed using the `medium` foundation model with mace-torch 0.3.13.[101] The models were trained for up to 2000 epochs based on a validation split of 5% structures from the training set. The loss function incorporated weighted contributions from both energy and force components, assigning a force weight of 200.0 and an energy weight of 10.0 to prioritize accurate force predictions during geometry optimization. The MLIP parameter optimizer used the AMSGrad method with a weight decay of 5×10^{-7} per epoch, and an exponential moving average (EMA) of the model parameters were maintained throughout fine-tuning with a decay factor of 0.995. A batch size of 4 was used for both training and validation, and the learning rate was fixed at 0.001. To ensure numerical stability, gradient clipping was applied with a threshold of 1.0. The Au model required extended fine-tuning upto 3000 epochs to achieve convergence. After epoch 2000, the energy weight in the loss function was increased from 10.0 to 1000.0 to further emphasize accurate prediction of adsorption energies.

Table D.1: Training dataset sizes and final mean averaged error (MAE) values for MACE-MP-0 models fine-tuned per metal surface.

2cSurface	No of Structures		MAE	
	Training set	Test Set	Energy [meV/atom]	Force [meV/Å]
Cu(100)	1511	167	0.7	5.8
Au(100)	1394	154	1.3	2.3
Ag(100)	1475	163	0.6	2.4

The convergence behavior of the MACE models during training is shown in Figure D.4, which reports the decay of per-atom MAE for energies and forces as a function of fine-tuning epoch. These fine-tuned MACE models for accelerated structure screening and initial relaxation, providing starting geometries that were subsequently refined using full DFT relaxations (See section DFT optimization) with top-layer relaxation enabled.

D.2.2 Bayesian Optimization of Adsorption Geometries

Bayesian Optimization of the HB238 molecule adsorbed on coinage metal surfaces was performed using the BOSS framework, with fine-tuned MACE models as the energy evaluator.

The search space for Bayesian optimization was defined over five configurational degrees of freedom: lateral translations (x and y) and molecular rotations (α, β, γ). The bounds for the lateral search along x and y direction were set according to the calculated lattice constants of each surface: [0, 4.07] Å for

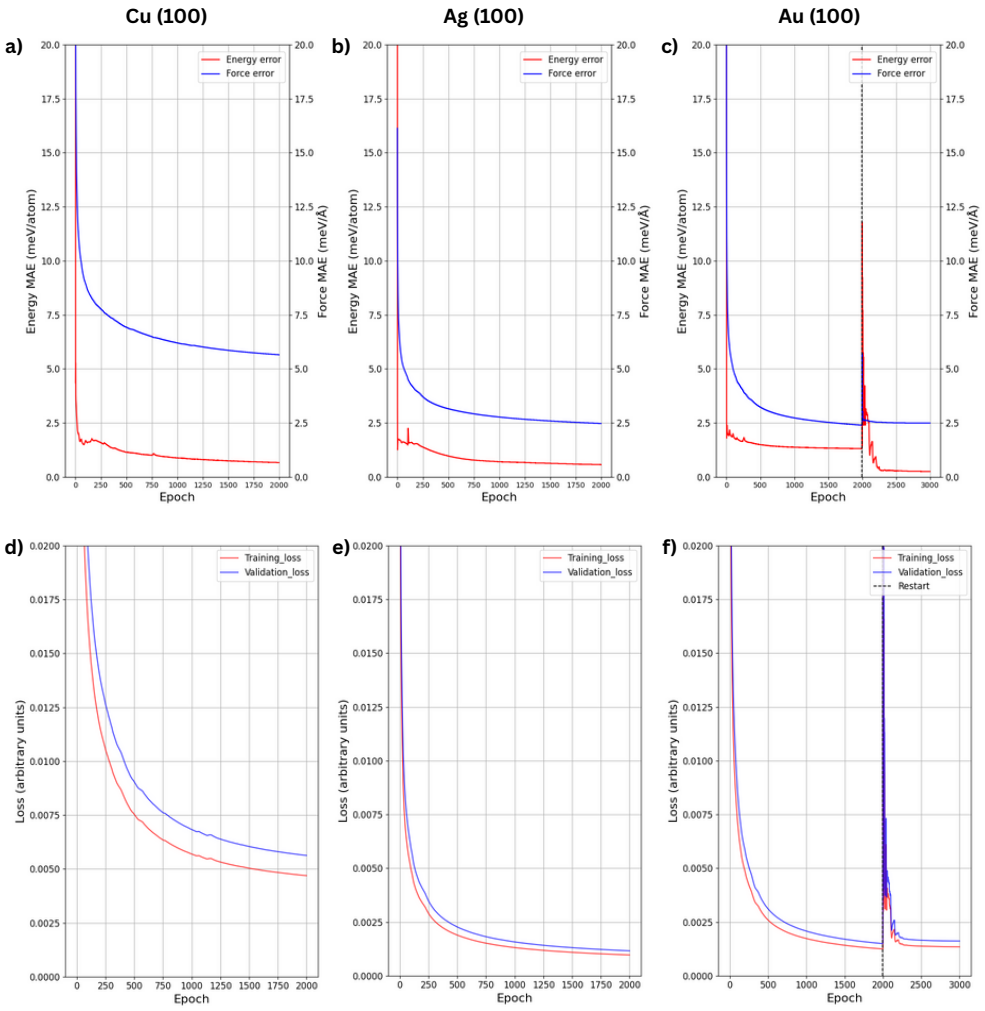


Figure D.4: a-c) Training error progression for the MACE-MP-0 models on Cu(100), Ag(100), and Au(100) surfaces, respectively. The mean absolute error (MAE) for energy (meV/atom) is shown in red, and the MAE for forces (meV/Å) is shown in blue, plotted over training epochs. (d–f) Training (red) and validation (blue) loss evolution over epochs during fine-tuning of the corresponding MACE models.

Ag(100), [0, 4.10] Å for Au(100), and [0, 3.56] Å for Cu(100). The details of measured and calculated lattice constants are given in Table S1. The rotational angles were allowed to vary in the range [0°, 359°] for all three axes. Each Bayesian optimization run was initialized with 20 randomly sampled configurations and refined through 980 adaptive sampling iterations using acquisition functions tailored to each degree of freedom.

Each candidate configuration was constructed by using a pre-relaxed Ag(100), Au(100), or Cu(100) surface, respectively, and the gas-phase optimized geometry of HB238 obtained with CREST.[137] Both the surface slab and gas phase molecule were first optimized at PBE/D3 level using VASP. The molecule was first rotated about its center of mass, then laterally shifted and vertically adjusted such that its lowest atom was positioned 2.5 Å above the topmost metal layer.

D.2.3 DFT optimization

All structural optimizations of the HB238 molecule adsorbed on metal (100) surfaces were carried out using the Vienna Ab initio Simulation Package (VASP, version 6.4.3) [128, 129]. The plane-wave-based density functional theory (DFT) calculations employed the generalized gradient approximation (GGA) using the Perdew–Burke–Ernzerhof (PBE) exchange-correlation functional [220], with the D3 dispersion correction [132]. A moderate energy cutoff of 400 eV was applied for the plane-wave basis set. In previous studies,[299] we have shown that this is sufficient of HB238 adsorption on metal surfaces.

Metal surfaces (Cu, Ag, Au) were modeled using a 3-layer slab with a 8×8 supercell of the primitive surface unit cell, comprising 192 atoms. The bottom two layers were kept fixed during relaxation of the global minima adsorption structures of HB238 on each surface. However, for DFT structures generated for fine-tuning, metal slabs were kept fixed. The lattice constant for each surface was individually optimized using PBE/D3 prior to constructing the adsorption models. In preliminary test calculations and a previous study, we have observed that three-layer models are sufficient for convergence of adsorption energies.[299]

For ionic relaxation, the conjugate-gradient algorithm was used with a time step of 0.5 fs. A force convergence criterion of 0.01 eV/Å was set. Electronic convergence was defined as 10^{-6} eV. The Gaussian smearing method was used with a smearing width of 0.2 eV, A $(1 \times 1 \times 1)$ k-point grid, i.e. the Γ-point approximation, and automated real-space projection settings were applied.

PDOS calculations were carried out using the r^{2}SCAN hybrid functional with an exchange fraction of 10% (r^{2}SCANh). Based on tests with various functionals, r^{2}SCANh trned out to be most accurate.[299] In these calculations a higher plane-wave energy cutoff of 500 eV was used. The number of electronic bands was set to 3200 to ensure proper convergence of unoccupied states near the Fermi level (E_F). Gaussian smearing was used for partial occupancies with $\sigma \approx 0.1$ eV). Electronic energy convergence was set to 10^{-5} eV. The DOS was evaluated on a fine energy grid of 6001 points over a window from –6.0 to 0.0 eV relative to the E_F.

X-ray Absorption Spectra Calculations

XAS were calculated using the real-space projector augmented wave (PAW)[285] method implemented in Grid-Based Projector-Augmented Waves (GPAW) version 25.11.0, [306] in conjunction with the ASE. [284]

The adsorption structures were centered with a vacuum region of 10 Å along the surface normal to avoid interactions between periodic images. The plane-wave mode with a 600 eV cutoff and Γ-point sampling was employed. A Fermi–Dirac smearing of 0.1 eV was applied to the band occupations.

Core-level excitations were treated following the delta Kohn Sham (ΔKS) approach for 2p states of S atoms. A transition-potential ground state was obtained by introducing a half-core hole on the selected atom, while reference energies were evaluated from a calculation without a core hole. The final state was computed using a full core hole with an added electron in the conduction manifold and a constrained local magnetic moment on the excited atom. In all calculations, domain-decomposition parallelization was employed for efficient scaling.

The excitation energy was obtained from the total energy difference between the full core-hole and reference states. X-ray absorption spectra were then generated using the GPAW xas module, where both stick spectra and Gaussian-broadened spectra (FWHM = 0.6 eV) were calculated. The spectra were aligned by shifting the energy scale such that the first transition matched the ΔKS excitation energy, and the total intensity was obtained by summing over the three Cartesian S2p components.

D.3 Adsorption Geometry and Interaction Strength

Using the above mentioned MACE/BOSS workflow, 1000 unique structures were generated for each surface and evaluated using single-point energy calculations with both the pre-trained MACE-MP-0 model and the fine-tuned metal-specific models. Figure D.5 compares the predicted energies with DFT reference values in order to demonstrate the improvement upon finetuning. The zero-energy reference point in each panel corresponds to the global minimum structure predicted by the MACE/BOSS search. In Figure D.5a, c and e, the blue points represent predicted energies from the general-purpose MACE-MP-0 model with D3 corrections, which shows substantial deviation and poor ranking ability as confirmed with DFT single point calculations. Fine-tuning the MACE models improved its agreement with DFT-calculated energies, allowing it to accurately rank structures and identify global minima.

These lowest-energy configurations are all flat-lying geometries where the interaction of the π-system of HB238 with the surface is maximized. The corresponding optimized structures are shown in Figure D.6. The representative side-on, and edge-on configurations are shown in Figure D.7 using DFT-based geometry re-optimizations (PBE/D3 in VASP). They are in general less stable than the flat-lying structures. Their adsorption energies are 1.5–3.5 eV higher than those of the flat structures, as shown in Table D.2. This significant energetic difference reflects both loss of interaction of the π-electrons with the metal and diminished S/metal bonding. Nonetheless, such orientations are useful for probing rotational barriers and understanding steric effects during monolayer formation. Notably, some side-on structures show downward-pointing cyano nitrogen atoms (N3/N4) that interact with the surface, contributing to modest stabilization on Cu(100), but far less so on Ag(100) or Au(100), where N–metal interactions are weaker.

Figure D.8 presents the lateral adsorption positions of the sulfur atoms in HB238 for all configurations within 0.1 eV of the global minimum on Cu(100), Ag(100), and Au(100) surfaces. On the Cu(100) surface, both the donor sulfur (S1) and acceptor sulfur (S2) preferentially bind near atop sites, indicating strong site-specific Cu–S interactions. While S2 consistently adsorbs close to atop positions across all surfaces, S1 deviates from atop adsorption on Ag(100), likely due to the larger lattice constant of 4.07 Å compared to Cu (3.56 Å) limiting simultaneous atop alignment of both sulfur atoms. Au(100) has the largest lattice constant of 4.10 Å and the sulfur atoms exhibit the largest displacement

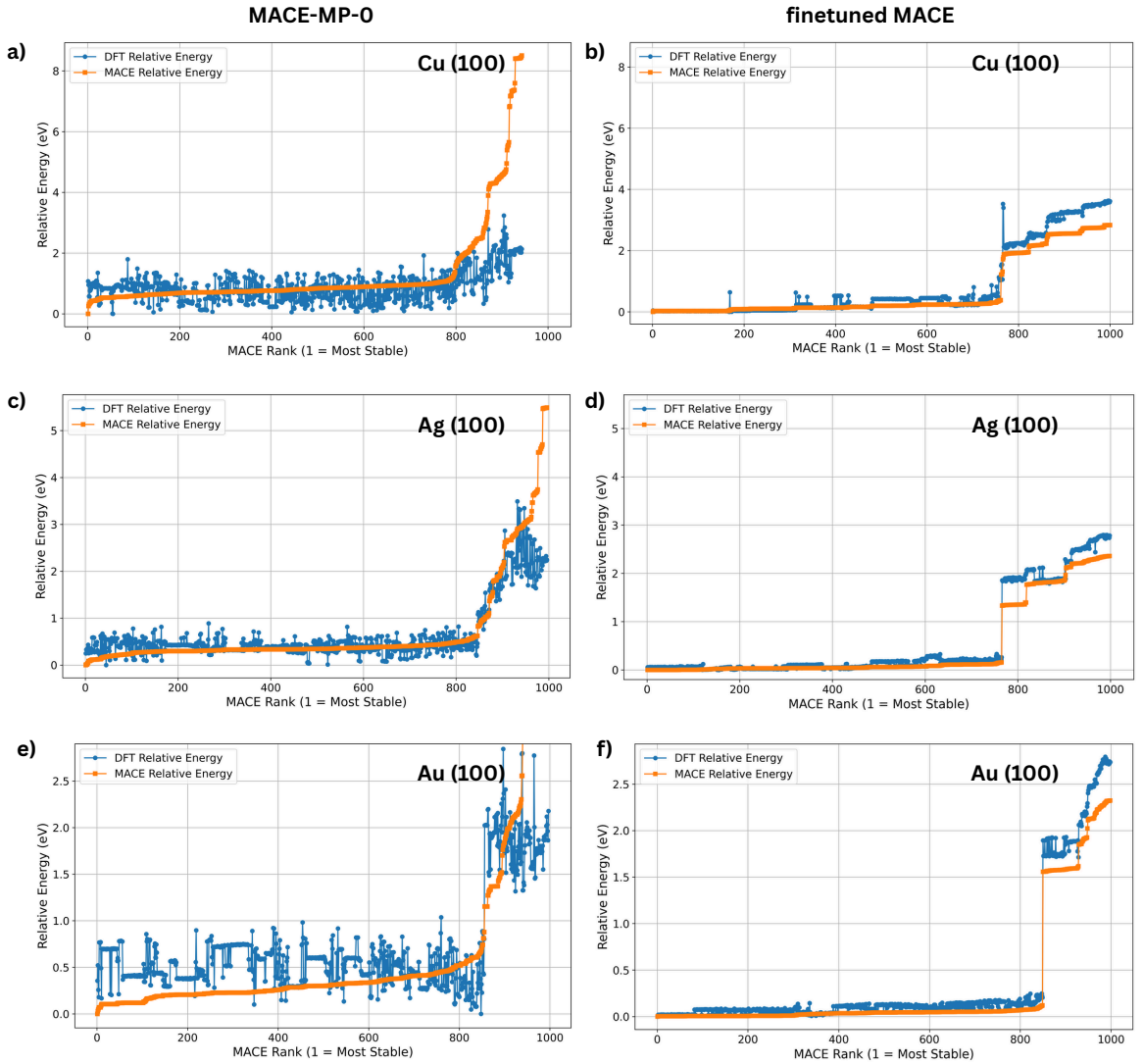


Figure D.5: Comparison of predicted relative energies from MACE models against DFT reference energies for 1000 HB238 adsorption geometries generated via BOSS. Panels a), c), and e) show predictions from the pre-trained MACE-MP-0 model; b), d), and f) show results from the fine-tuned metal-specific models. The zero point in each panel denotes the global minimum from the MACE/BOSS workflow.

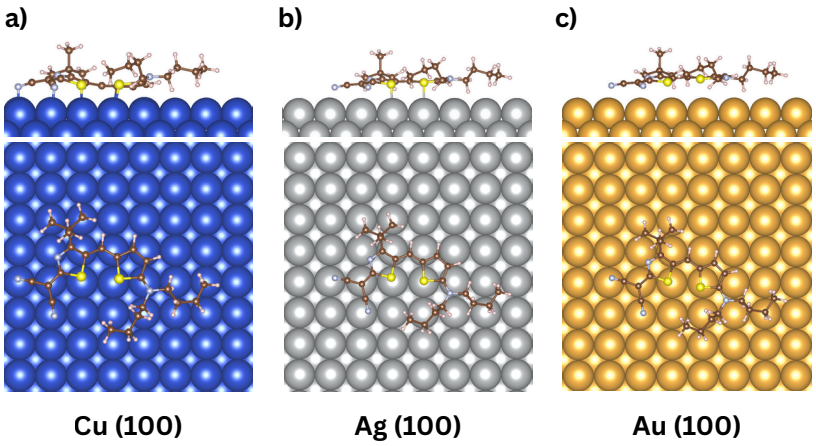


Figure D.6: Sideview(top) and top-view (bottom) of optimized flat-lying adsorption geometries of HB238 on a) Cu(100), b) Ag(100), and c) Au(100). Atom colors: sulfur (yellow), nitrogen (light blue), carbon (brown), hydrogen (light pink), and metal atoms—Cu (blue), Ag (silver), Au (gold).

Table D.2: Adsorption energies (in eV) of HB238 on Cu(100), Au(100), and Ag(100) for face-on orientation in Figure D.6 and different side-on and edge-on orientations as shown in Figure D.7.

Configuration	Cu(100)	Au(100)	Ag(100)
Face-on	–4.00	–3.23	–3.20
Side-on 1	–1.46	–1.41	–1.70
Side-on 2	–1.52	–1.28	–0.66
Edge-on 1	–0.34	–0.38	–0.29
Edge-on 2	–1.48	–1.10	–1.00

from atop sites on Au, consistent with weak Au–S bonding and reduced site specificity.

The most stable flat-lying structures predicted by the MACE/BOSS framework were further analyzed in terms of geometric registry, specifically the alignment of S and N atoms with surface sites, as shown in Figure D.6. The vertical distances, bond lengths, and lateral deviation angles between selected heteroatoms (S, N) and their nearest surface metal atoms are shown in Figure D.9 and summarized in Table D.3. On Cu(100), the larger deviation angles (θ) for N1 (15°) and N2 (14°) suggest strained positioning, while on Au and Ag these atoms exhibit even larger θ angles (20–24°), reflecting both weaker direct interaction and greater steric freedom. The N3/N4 atoms from the cyano groups lie closer to the surface on Cu (2.09–2.10 Å), implying secondary interactions that may enhance dipolar alignment. On Ag(100), S1 exhibits a deviation of just 1.9°, implying near-perfect atop adsorption. This lateral registry reflects a balance between surface corrugation and lattice constant: Cu has the smallest lattice constant (3.56 Å) and relatively stronger Cu-N interaction, leading to misalignment and strain on S atop positions, whereas Ag (4.07 Å) allows more relaxed matching.

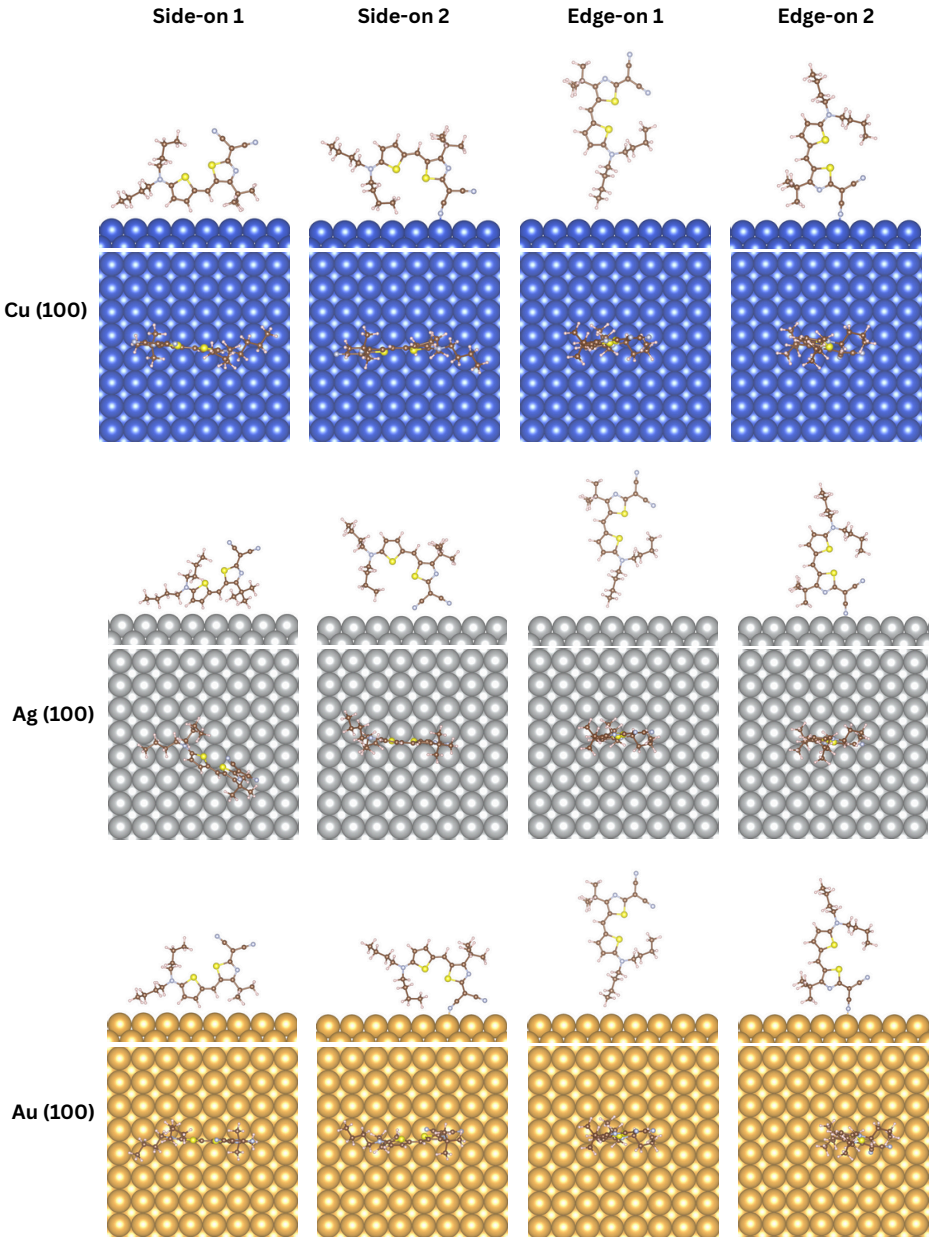


Figure D.7: Sideview(top) and top-view (bottom) of optimized HB238 adsorption geometries on Cu(100), Ag(100), and Au(100) in distinct orientations: Side-on 1, Side-on 2, Edge-on 1, and Edge-on 2. Each row corresponds to a different metal surface (top to bottom: Cu, Ag, Au).

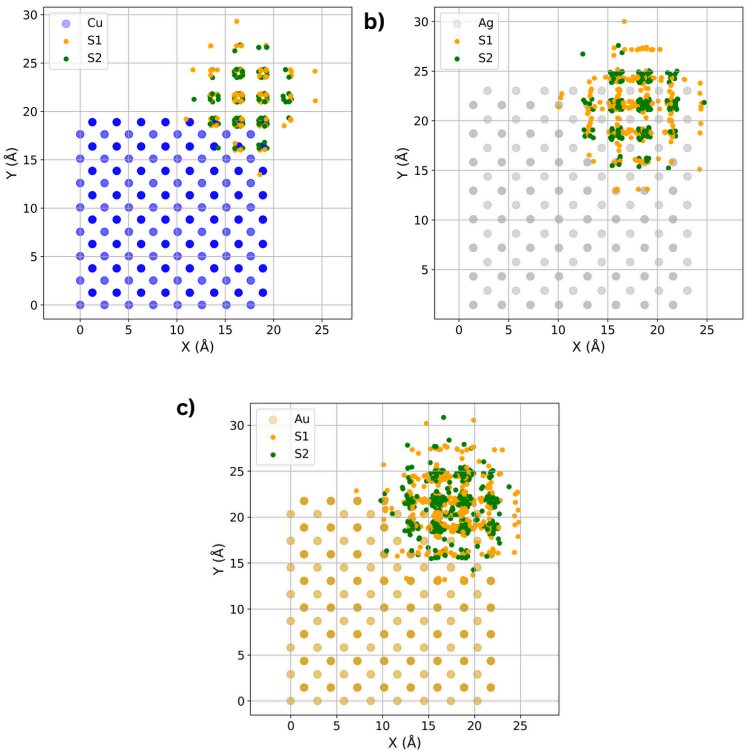


Figure D.8: a-c) Top-view projections of the donor (S1) and acceptor (S2) sulfur atom positions in HB238 adsorbed on Cu(100), Ag(100), and Au(100) surfaces, respectively. Each point corresponds to a local minimum configuration within 0.1 eV of the global minimum as obtained with BOSS/MACE. Donor(S1) and acceptor(S2) sulfur atoms are shown in orange and green, respectively. Surface metal atoms are displayed in blue (Cu), silver (Ag), and gold (Au).

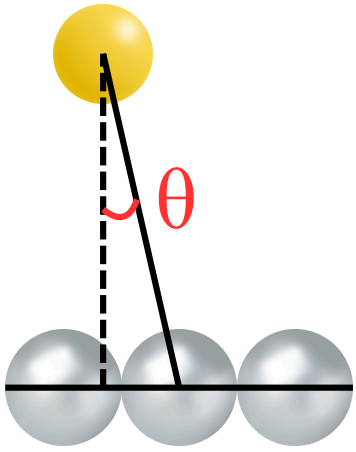


Figure D.9: Representative illustration of the vertical (dotted black line) and actual (solid black line) S–Ag bond distances in the monolayer. The angle θ denotes the deviation angle from the ideal atop adsorption position.

Table D.3: Vertical distances (height difference between each HB238 atom and its nearest metal atom), bond lengths, and lateral deviation angles (displacement from the ideal atop site) for donor and acceptor atoms in HB238 adsorbed in the face-on configuration. See Figure D.3 for atom numbering.

Surface	Atom	Vert. Distance (Å)	Bond Length (Å)	Angle (°)
Cu(100)				
	S1	2.35	2.41	14
	S2	2.31	2.35	11
	N1	3.06	3.17	15
	N2	3.16	3.25	14
	N3	2.09	2.12	9
	N4	2.10	2.12	6
Ag(100)				
	S1	2.88	2.88	2
	S2	2.72	2.75	8
	N1	3.29	3.49	20
	N2	3.17	3.38	20
	N3	2.44	2.53	15
	N4	2.39	2.46	13
Au(100)				
	S1	3.19	3.23	9
	S2	2.88	2.91	8
	N1	3.37	3.68	24
	N2	3.30	3.52	20
	N3	2.56	2.57	6
	N4	2.56	2.61	11

D.4 X-ray Absorption Spectra (XAS)

To investigate the site-specific electronic environments of the donor and acceptor sulfur atoms in HB238, sulfur L-edge S 2p XAS were computed using GPAW with the PBE/D3 functional. Separate core-hole simulations were performed for the thiophene donor(S1) and thiazole acceptor (S2) sulfur atoms in HB238 adsorbed on Cu(100), Au(100), and Ag(100) surfaces, as well as for the gas phase molecule. In the gas phase, the donor(S1) and acceptor(S2) sulfur atoms exhibit a well-separated ΔE_{XAS} of 0.44 eV, reflecting the strong internal dipole moment of the molecule. Upon adsorption on metal surfaces, this energy separation is significantly reduced.

On Cu(100), the calculated donor–acceptor core-level splitting is 0.29 eV, indicating that strong molecule–substrate interactions partially symmetrize the local chemical environments of the two sulfur atoms. For Ag(100), the splitting is slightly smaller at 0.26 eV, and Au(100) exhibits the lowest separation at 0.20 eV. These effects have previously been explained by surface-mediated charge transfer speculation and the formation of image charges on the surface[299] These values highlight the substrate sensitivity of ΔE_{XAS} and the coupling strength of the molecule-surface interaction.

Figure D.10 a shows the S 2p sulfur L-edge XAS spectra for HB238 adsorbed on the three coinage

metal surfaces, with separate contributions from the donor and acceptor sulfur atoms. The calculated peak positions and energy separations are summarized in Table D.4 and shown in Figure D.10 b. For the Au (100) surface, the calculated ΔE_{XAS}, 0.20 eV, is close to the experimentally measured value of 0.3eV. On the Ag (100) surface, however, the calculated ΔE_{XAS} is 0.26 eV, while the experimental value is zero. Possible reasons for the discrepancy are discussed below.

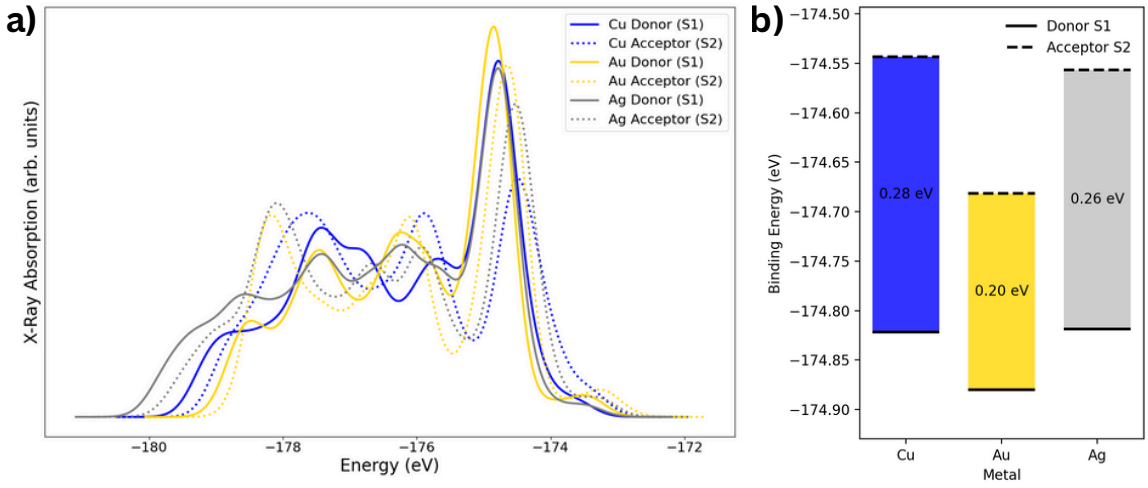


Figure D.10: a)Sulfur $L_{2,3}$-edge (S 2p) XAS spectra of a single HB238 molecule adsorbed on coinage metal surfaces. Solid lines represent the donor(S1) sulfur contributions, while dotted lines indicate the acceptor(S2) sulfur sites. Spectra are color-coded by substrate: blue for Cu(100), yellow for Au(100), and grey for Ag(100).b) Binding energy of the donor (S1) and acceptor (s2) of HB238 molecule adsorbed on Cu, Au and Ag surface.

Table D.4: Calculated sulfur sulfur $L_{2,3}$-edge (S 2p) XAS peak positions (in eV) for donor (S1) and acceptor sulfur(S2) atoms in HB238 adsorbed on metal surfaces and in the gas phase. ΔE_{XAS} refers to the donor–acceptor energy difference.

System	ΔE_{XAS} (eV)	$\Delta E^{exp}_{\mathrm{XPS}}$ (eV)
Cu(100)	0.29	-
Ag(100)	0.26	0.0
Au(100)	0.20	0.3
Gas phase	0.44	-

To further analyze the nature of the reduced ΔE_{XAS} on Ag(100), we compared the spectra of HB238 in two distinct adsorption scenarios: a single isolated molecule as studied here and a densely packed monolayer studied in our previous work.[299] While the monomer configuration retains much of the molecular electronic asymmetry, the monolayer shows a near-complete overlap between donor and acceptor sulfur peaks. This is illustrated in Figure D.11, where panel (b) shows the calculated L-edge S 2p spectra for the monolayer on Ag(100). The near-degeneracy of the two peaks in this configuration arises not only from surface-mediated charge transfer and metal–molecule screening, but also from lateral intermolecular interactions.

This finding has important implications for experimental interpretation. HB238 is known to form

well-ordered monolayers on Ag(100), and the XPS measurements by Öcal et al. [81] were conducted under such monolayer coverage conditions. Their S 2*p* spectra show no resolvable donor–acceptor splitting, consistent with our monolayer simulations. The agreement between the calculated XAS spectra and experimental binding energies supports the conclusion that both surface screening and intermolecular packing contribute to the observed spectral convergence.

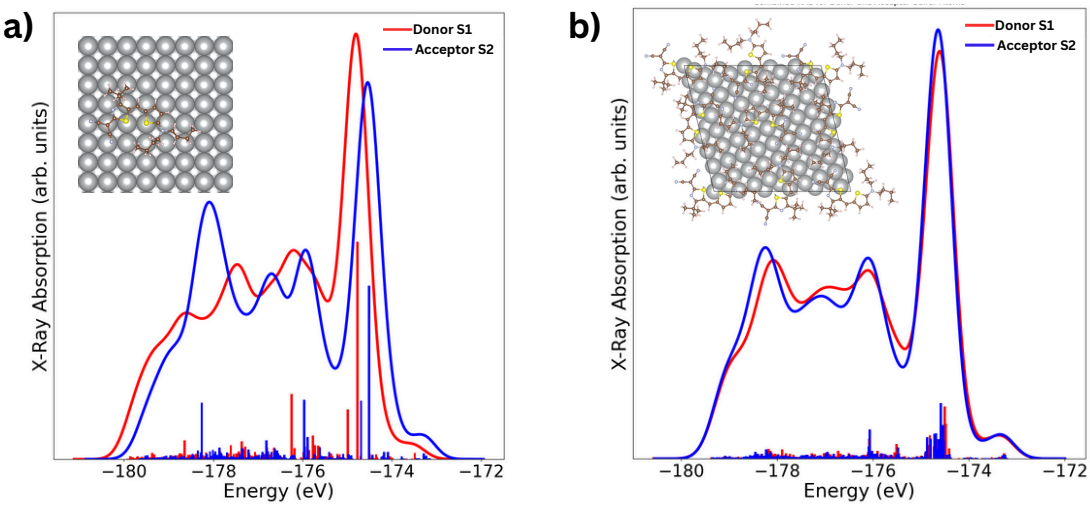


Figure D.11: Calculated sulfur $L_{2,3}$-edge (S 2p) XAS spectra showing (a) contributions from the donor(S1) and acceptor(S2) sulfur atoms of an isolated HB238 molecule adsorbed on Ag(100), and (b) the summed contributions from four donor(S1) and four acceptor(S2) sulfur atoms in the monolayer unit cell of HB238 adsorbed on the Ag surface. The donor and acceptor components are plotted separately to illustrate the effects of intermolecular interactions and surface screening in the monolayer model. The inset shows the monolayer adsorption geometry. Image adapted with permission from Tomar *et al.*[299]. Copyright © 2025 American Chemical Society.

D.5 Projected Density of States

The frontier orbitals of HB238 couple to the electronic states of the metal substrates, as observed from the projected density of states (PDOS) calculations which were performed for the relaxed, flat–lying adsorption geometries on Cu(100), Au(100) and Ag(100) surfaces as shown in Figure D.12. The calculations were performed with the r^2SCANh functional.

In the gas phase, the HOMO and LUMO levels of HB238 are located at –3.41 eV and –2.05 eV, respectively at r^2SCANh level.[299] Upon adsorption, these levels shift relative to the substrate E_F and broaden due to hybridisation with metal states.

The molecular orbitals are visualized in Figure D.13 and the frontier orbitals data is given in Table D.5. The HOMO is 1.1-1.6 eV below E_F and the LUMO is 2.0-2.9 eV above E_F. The HOMO–LUMO separations referenced to the substrate E_F is in the range of 3.3–4.1 eV. This is significantly larger than the calculated gas-phase HOMO–LUMO gap of 1.44 eV. The increase reflects a renormalisation of the frontier levels induced by the metallic environment. Together, the PDOS,

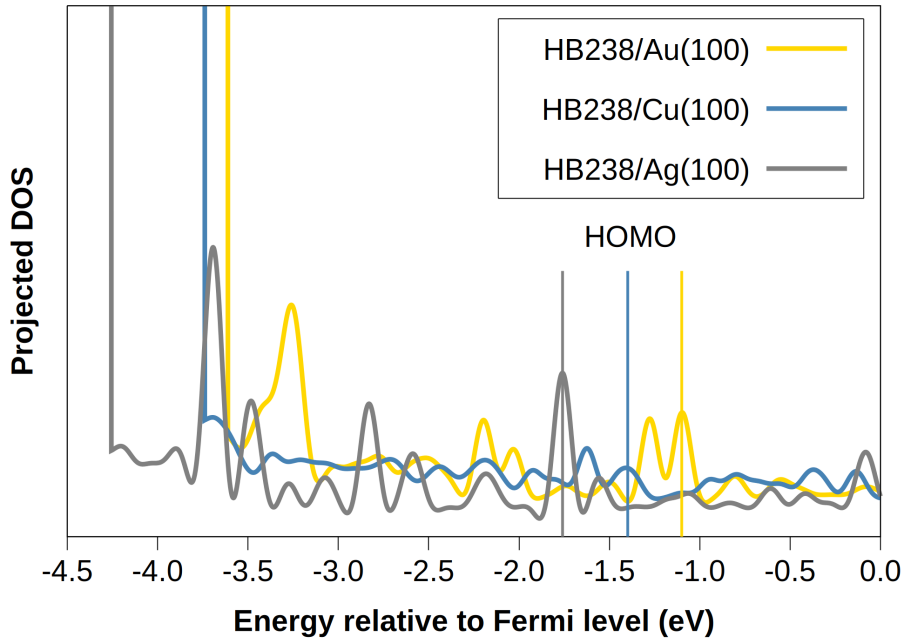


Figure D.12: Projected density of states of a single HB238 molecule adsorbed on Au(100), Cu(100) and Ag(100) surface respectively with respect to the Fermi level calulated using r^2SCANh functional.

Table D.5: Calculated HOMO and LUMO energies of HB238 molecules adsorbed on different coinage metal surfaces. The values are presented with respect to the Fermi level of each metal surface. All calculations are performed with r^2SCANh.

Surface	HOMO (eV)	LUMO (eV)	Gap (eV)
Cu(100)	-1.39	1.98	3.37
Ag(100)	-1.64	2.49	4.13
Au(100)	-1.11	2.86	3.97

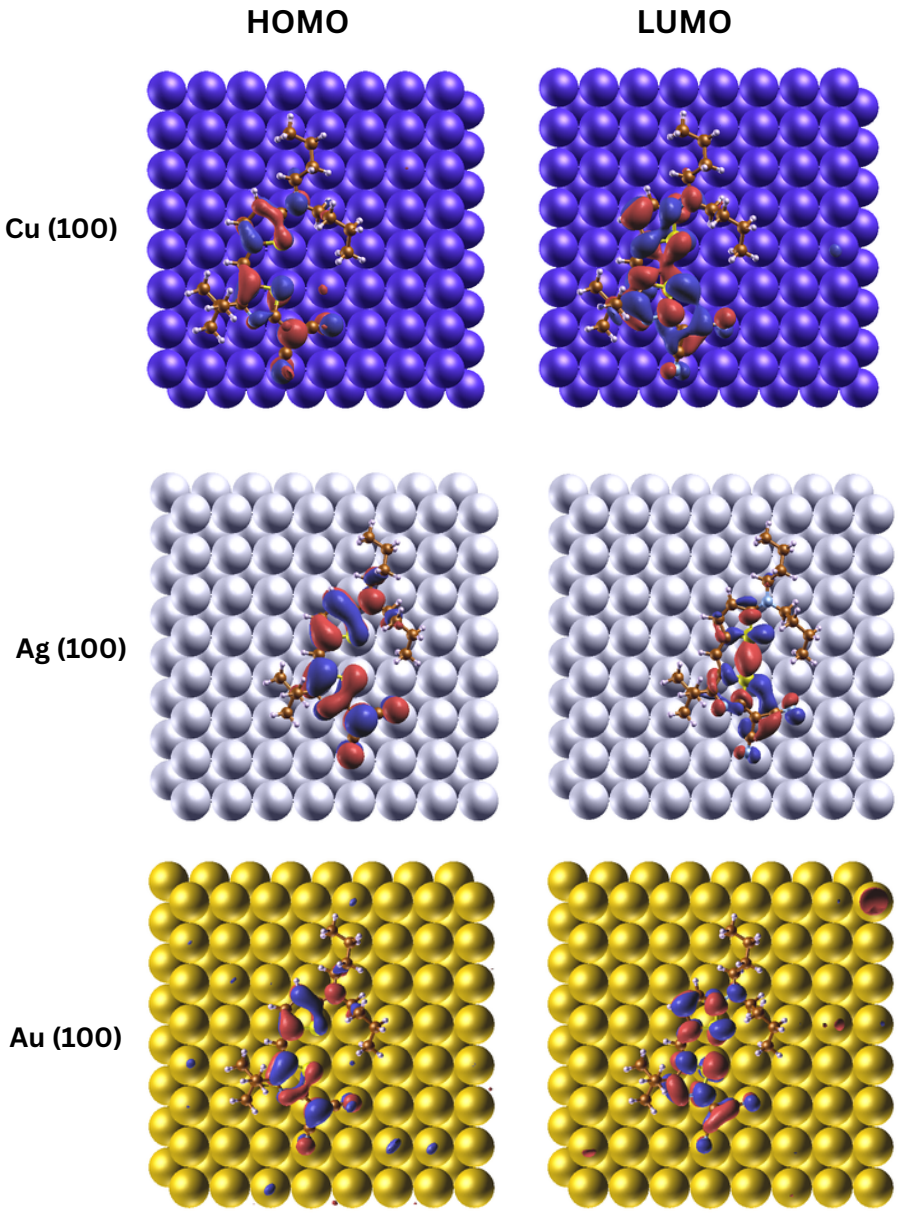


Figure D.13: HOMO and LUMO orbitals of HB238 molecule adsorbed on Cu (top), Ag (middle) and Au (bottom) respectively. All calculations are performed with r^2SCANh

orbital visualisations and energy alignments demonstrate how different surfaces governs both the degree of hybridisation and the renormalised level alignment of HB238.

D.6 Conclusion

In this study, we investigated the adsorption behavior and electronic structure of an isolated merocyanine molecule HB238 on Cu(100), Au(100), and Ag(100) surfaces using the MACE/BOSS framework, projected density of states and core-level spectroscopy simulations. Geometry optimizations revealed that the molecule adopts a flat-lying configuration on all surfaces, with the strongest adsorption energy observed on Cu(100), consistent with shorter adsorption distances. The site-specific bonding strength of sulfur atoms on metal surfaces follows the trend: Cu > Ag > Au.

Sulfur $L_{2,3}$-edge (S 2p) XAS spectra demonstrated a progressive reduction in donor–acceptor spectral splitting upon adsorption in the series Cu → Ag → Au. The largest splitting is observed on Cu(100), suggesting partial neutralization of the molecular dipole through surface-mediated charge redistribution. PDOS calculations further show that the band gap of the HB238 molecule increases upon adsorption on coinage metal surfaces.

Together, our results highlight how metal surface reactivity governs the degree of electronic perturbation in adsorbed dipolar molecules, with implications for the design of molecular electronic interfaces and optoelectronic devices.

D.7 acknowledgement

The authors gratefully acknowledge funding from the DFG (RTG-2591 *TIDE Template-designed Organic Electronics*), and the computer resources from the HPC clusters Bonna, Bender and Marvin of the University of Bonn.

D.8 Supporting Information

The experimental and calulated lattice constants for metal bulks and the calculated sulfur XAS spectra of HB238 molecule on HB238 molecule are available in SI.

Bibliography

[1] S. R. Forrest, *Electronic Appliances on Plastic*, Nature **428** (2004) 911, ISSN: 00280836.

[2] M. C. Petty, *Organic and Molecular Electronics: from Principles to Practice*, John Wiley & Sons, 2019.

[3] S. Ogawa, *Organic Electronics Materials and Devices*, Springer, 2015.

[4] J. H. Kwon et al., *Highly efficient, reliable, and ultraflexible bio-organic light-emitting diode patch*, npj Flexible Electronics **9** (2025) 1, ISSN: 23974621.

[5] H. Jeon et al., *Nanomechanical Insights into Enhancing Flexibility and Reliability in Deformable Optoelectronic Devices*, ACS Appl. Mater. Interfaces **17** (2025) 42568.

[6] F. Mohamad, M. A. M. Teridi, P. Chelvanathan, M. N. M. Nasir and Ibrahim, *Recent development and challenges of flexible perovskite transparent conductive oxides in optoelectronic applications*, Mater. Sci. Semicon. Proc **196** (2025) 109577.

[7] W. Lowrie et al., *Organic photovoltaics: The current challenges*, J. Chem. Phys **158** (2023) 110901.

[8] E. K. Solak and E. Irmak, *Advances in organic photovoltaic cells: a comprehensive review of materials, technologies, and performance*, RSC Adv. **13** (2023) 12244.

[9] G. Zhang et al., *Renewed Prospects for Organic Photovoltaics*, Chem. Rev. **122** (2022) 14180.

[10] O. Inganäs, *Organic Photovoltaics over Three Decades*, Adv. Mater. **30** (2018) 1800388, ISSN: 15214095.

[11] B. Geffroy, P. le Roy and C. Prat, *Organic light-emitting diode (OLED) technology: Materials, devices and display technologies*, Polym. Int. **55** (2006) 572, ISSN: 09598103.

[12] R. Liguori, F. Nunziata, S. Aprano and M. G. Maglione, *Overcoming Challenges in OLED Technology for Lighting Solutions*, Electronics **13** (2024) 1299, ISSN: 2079-9292.

[13] G. Mandal, J. Bauri and R. B. Choudhary, *Conjugated polymeric nanocomposite-based light-generating active materials for OLED applications: A review*, Mater. Sci. Eng. B **303** (2024) 117271.

[14] G. Hong et al., *A Brief History of OLEDs—Emitter Development and Industry Milestones*, Adv. Mater. **33** (2021) 2005630, ISSN: 15214095.

[15] X. Ren et al.,
Organic Field-Effect Transistor for Energy-Related Applications: Low-Power-Consumption Devices, Near-Infrared Phototransistors, and Organic Thermoelectric Devices,
Adv. Energy Mater. **8** (2018) 1801003, ISSN: 16146840.

[16] P. B. Bennett et al., *Bottom-up graphene nanoribbon field-effect transistors*,
Appl. Phys. Lett. **103** (2013) 253114, ISSN: 00036951.

[17] F. Yang et al., *Emerging Opportunities for Ferroelectric Field-Effect Transistors: Integration of 2D Materials*, Adv. Funct. Mater. **34** (2024) 2310438, ISSN: 16163028.

[18] Z. Chen et al., *π-Conjugated Polymers for High-Performance Organic Electrochemical Transistors: Molecular Design Strategies, Applications and Perspectives*,
Angew. Chem. Int. Ed. **64** (2025) e202423023.

[19] G. Li, R. Zhu and Y. Yang, *Polymer solar cells*, Nat. Photonics **6** (2012) 153.

[20] J. Kim et al., *Interface Control in Organic Electronics Using Mixed Monolayers of Carboranethiol Isomers Terms of Use*, Nano Lett **14** (2014) 2946.

[21] A. K.-Y. Jen, *Interface engineering of organic electronics*, J. Mater. Chem. (2010) 2491.

[22] X. Sun, C.-A. Di and Y. Liu,
Engineering of the dielectric-semiconductor interface in organic field-effect transistors †,
J. Mater. Chem. **20** (2010) 2599.

[23] Y. Xie et al., *Improving the Efficiency of Layer-by-Layer Organic Photovoltaics to Exceed 19% by Establishing Effective DonorAcceptor Interfacial Molecular Interactions*,
ACS Appl. Mater. Interfaces **17** (2025) 15741.

[24] Z. He, K. Asare-Yeboah and S. Bi, *Diketopyrrolopyrrole based organic semiconductors: pioneering advances in organic electronic devices*, Discov. Electron **2** (2025) 27.

[25] R. Ruvalcaba et al., *On-Surface Reactions of Electronically Active Self-Assembled Monolayers for Electrode Work Function Tuning*, ACS Materials Lett. **7** (2025) 1421.

[26] D. Gkeka et al.,
Tuning Hole-Injection in Organic-Light Emitting Diodes with Self-Assembled Monolayers,
ACS Appl. Mater. Interfaces **16** (2024) 39728.

[27] R. Otero, A. L. Vázquez de Parga and J. M. Gallego,
Electronic, structural and chemical effects of charge-transfer at organic/inorganic interfaces,
Surf. Sci. Rep **72** (2017) 105.

[28] V. Coropceanu et al., *Charge transport in organic semiconductors*,
Chem. Rev. **107** (2007) 926, ISSN: 00092665.

[29] Q. Li and Z. Li, *Molecular Packing: Another Key Point for the Performance of Organic and Polymeric Optoelectronic Materials*, Acc. Chem. R. **53** (2020) 962, ISSN: 15204898.

[30] J. Vanderspikken, W. Maes and K. Vandewal, *Wavelength-Selective Organic Photodetectors*,
Adv. Funct. Mater. **31** (2021), ISSN: 16163028.

[31] A. Liess et al.,
Organic Thin Film Transistors Based on Highly Dipolar Donor-Acceptor Polymethine Dyes,
Adv. Funct. Mater. **25** (2015) 44, ISSN: 1616301X.

[32] S. Nizamuddin et al., *A review on the development of covalently connected donor-acceptor molecular materials for single-component organic solar cells*, J. Mater. CHem. A **13** (2025) 11114.

[33] H. Bürckstümmer, *Merocyanine Dyes for Solution-Processed Organic Bulk-Heterojunction Solar Cells*, PhD Thesis (2011).

[34] K. W. Chen et al., *Merocyanines for vacuum-deposited small-molecule organic solar cells*, Organic Electronics **26** (2015) 319, ISSN: 15661199.

[35] F. Würthner and K. Meerholz, *Systems chemistry approach in organic photovoltaics*, Chem. Eur. J. **16** (2010) 9366, ISSN: 09476539.

[36] U. Hörmann et al., *Voc from a morphology point of view: The influence of molecular orientation on the open circuit voltage of organic planar heterojunction solar cells*, J. Phys.Chem. C **118** (2014) 26462, ISSN: 19327455.

[37] N. Koch et al., *Molecular orientation dependent energy levels at interfaces with pentacene and pentacenequinone*, Org. Electron. **7** (2006) 537, ISSN: 15661199.

[38] S. Duhm et al., *Orientation-dependent ionization energies and interface dipoles in ordered molecular assemblies*, Nat. Mater. **7** (2008) 326, ISSN: 14764660.

[39] L. G. S. Brooker et al., *Studies in the Cyanine Dye Series. XI.1 The Merocyanines*, J. Am. CHem. Soc. **73** (1951) 5326.

[40] W. T. Simpson, *A Mathematical Treatment of the Color of the Merocyanine Dyes*, J. Am. Chem. Soc **73** (2002) 5359, ISSN: 15205126.

[41] C. Reichardt, *Solvatochromic Dyes as Solvent Polarity Indicators*, Chem. Rev. **94** (2002) 2319, ISSN: 00092665.

[42] A. Mishra, R. K. Behera, P. K. Behera, B. K. Mishra and G. B. Behera, *Cyanines during the 1990s: A Review*, Chem. Rev. **100** (2000) 1973.

[43] A. V. Kulinich, A. A. Ischenko, A. K. Chibisov and G. V. Zakharova, *Effect of electronic asymmetry and the polymethine chain length on photoprocesses in merocyanine dyes*, J. Photochem. Photobiol. A **274** (2014) 91.

[44] A. Arjona-Esteban, M. Stolte and F. Würthner, *Conformational Switching of π-Conjugated Junctions from Merocyanine to Cyanine States by Solvent Polarity*, Angew. Chem. **128** (2016) 2516.

[45] F. Würthner, G. Archetti, R. Schmidt and H.-G. Kuball, *Solvent Effect on Color, Band Shape, and Charge-Density Distribution for Merocyanine Dyes Close to the Cyanine Limit*, Angew. Chem. **120** (2008) 4605, ISSN: 00448249.

[46] A. V. Kulinich, A. A. Ishchenko and U. M. Groth, *Electronic structure and solvatochromism of merocyanines NMR spectroscopic point of view*, Spectrochim. Acta A **68** (2007) 6.

[47] H. Bürckstümmer et al., *Tailored merocyanine dyes for solution-processed BHJ solar cells*, J. Mater. Chem. **20** (2010) 240, ISSN: 0959-9428.

[48] A. V. Kulinich and A. A. Ishchenko, *Merocyanines: Electronic Structure and Spectroscopy in Solutions, Solid State, and Gas Phase*, Chem. Rev. **124** (2024) 12086, ISSN: 15206890.

[49] F. Würthner, T. E. Kaiser and C. R. Saha-Möller, *J-Aggregates: From Serendipitous Discovery to Supramolecular Engineering of Functional Dye Materials*, Angewandte Chemie International Edition **50** (2011) 3376, ISSN: 14337851.

[50] M. Kasha, H. R. Rawls and M. A. El-Bayoumi, *The Exciton Model In Molecular Spectroscopy*, Pure Appl. Chem. **11** (1965) 371, ISSN: 13653075.

[51] F. C. Spano, *The Spectral Signatures of Frenkel Polarons in H-and J-Aggregates*, Acc. Chem. Res. **43** (2010) 429.

[52] N. J. Hestand and F. C. Spano, *Expanded Theory of H- and J-Molecular Aggregates: The Effects of Vibronic Coupling and Intermolecular Charge Transfer*, Chem. Rev. **118** (2018) 7069, ISSN: 15206890.

[53] A. Eisfeld and J. S. Briggs, *The J- and H-bands of organic dye aggregates*, Chem. Phys. **324** (2006) 376, ISSN: 0301-0104.

[54] N. J. Hestand and F. C. Spano, *Expanded Theory of H-and J-Molecular Aggregates: The Effects of Vibronic Coupling and Intermolecular Charge Transfer*, Chem. Rev. **118** (2018) 7069.

[55] A. Liess et al., *Ultranarrow Bandwidth Organic Photodiodes by Exchange Narrowing in Merocyanine H- and J-Aggregate Excitonic Systems*, Adv. Funct. Mater. **29** (2019), ISSN: 1616-301X.

[56] B. L. Cannon et al., *Coherent Exciton Delocalization in a Two-State DNA-Templated Dye Aggregate System*, J. Phys. Chem. A **121** (2017) 6905.

[57] T. Schembri, L. Kolb, M. Stolte and F. Würthner, *Polarized, color-selective and semi-transparent organic photodiode of aligned merocyanine H-aggregates*, J. Mater. Chem. C **12** (2024) 4948, ISSN: 2050-7534.

[58] M. Más-Montoya, R. A. J. Janssen, M. Más-Montoya and R. A. J. Janssen, *The Effect of H- and J-Aggregation on the Photophysical and Photovoltaic Properties of Small Thiophene–Pyridine–DPP Molecules for Bulk-Heterojunction Solar Cells*, Adv. Funct. Mater. **27** (2017) 1605779, ISSN: 1616-3028.

[59] A. Arjona-Esteban et al., *Influence of Solid-State Packing of Dipolar Merocyanine Dyes on Transistor and Solar Cell Performances*, J. Am.Chem. Soc. **137** (2015) 13524, ISSN: 15205126.

[60] A. Liess, M. Stolte, T. He and F. Wü rthner, *Single-crystal field-effect transistors of a highly dipolar merocyanine dye †*, Mater. Horiz **3** (2016) 77.

[61] M. S. A. Mamun et al., *Chemistry of the photoisomerization and thermal reset of nitro-spiropyran and merocyanine molecules on the channel of the MoS 2 field effect transistor*, Phys. Chem. Chem. Phys **23** (2021) 27281.

[62] L. Böhner et al., *Influencing optical and charge transport properties by controlling the molecular interactions of merocyanine thin films*, Org. Chem. Front **12** (2025) 1098.

[63] J.-C. Ribierre et al., *Simultaneous Edge-on to Face-on Reorientation and 1D Alignment of Small π-Conjugated Molecules Using Room-Temperature Mechanical Rubbing*, Adv. Func. Mater. **28** (2018) 170738.

[64] S. Malik, N. Fujita, M. Numata, K. Ogura and S. Shinkai, *Creation of supramolecular assemblies from a dipolar dye molecule by the template effect of 1,3-glucan polysaccharide*, J. Mater. Chem. **20** (2010) 9022.

[65] D. A. Kunkel et al., *Self-assembly of strongly dipolar molecules on metal surfaces*, J. Chem. Phys **142** (2015) 101921, ISSN: 0021-9606.

[66] P. Li and Z. H. Lu, *Interface Engineering in Organic Electronics: Energy-Level Alignment and Charge Transport*, Small Sci **1** (2021) 2000015, ISSN: 26884046.

[67] N. Chakravarthi et al., *A simple structured and efficient triazine-based molecule as an interfacial layer for high performance organic electronics*, Energy Environ. Sci. **9** (2016) 2595, ISSN: 1754-5706.

[68] D. Ji, L. Li, H. Fuchs and W. Hu, *Engineering the Interfacial Materials of Organic Field-Effect Transistors for Efficient Charge Transport*, Acc. Mater. Res. **2** (2021) 159, ISSN: 26436728.

[69] H. Ishii, K. Sugiyama, E. Ito and K. Seki, *Energy Level Alignment and Interfacial Electronic Structures at Organic/Metal and Organic/Organic Interfaces*, Adv. Mater. **11** (1999) 605, ISSN: 00036951.

[70] S. Winkler et al., *Impact of molecular dipole moments on fermi level pinning in thin films*, J. Phys. Chem. C **118** (2014) 11731, ISSN: 19327455.

[71] S. Krause, M. Stolte, F. Würthner and N. Koch, *Influence of Merocyanine Molecular Dipole Moments on the Valence Levels in Thin Films and the Interface Energy Level Alignment with Au(111)*, J. Phys. Chem. C **117** (2013) 19031, ISSN: 19327447.

[72] R. J. Maurer et al., *Adsorption structures and energetics of molecules on metal surfaces: Bridging experiment and theory*, Prog. Surf. Sci. **91** (2016) 72, ISSN: 00796816.

[73] C. Stadler, S. Hansen, I. Kröger, C. Kumpf and E. Umbach, *Tuning intermolecular interaction in long-range-ordered submonolayer organic films*, Nat. Phys. **5** (2009) 153, ISSN: 17452481.

[74] L. Óvári et al., *Hexagonal boron nitride on metal surfaces as a support and template*, Surf. Sci. Rep. **79** (2024) 100637, ISSN: 0167-5729.

[75] J. Cai, *Monolayer hexagonal boron nitride. An ultra-thin insulator*, PhD thesis, 2021.

[76] F. Karlický et al., *Adsorption of Organic Molecules to van der Waals Materials: Comparison of Fluorographene and Fluorographite with Graphene and Graphite*, J. Chem. Theory Comput. **13** (2017) 1328, ISSN: 15499626.

[77] G. Melani et al., *Donors, acceptors, and a bit of aromatics: electronic interactions of molecular adsorbates on hBN and MoS 2 monolayers - Physical Chemistry Chemical Physics (RSC Publishing) DOI:10.1039/D2CP01502A*, Phys. Chem. Chem. Phys. **24** (2022) 16671.

[78] J. D. Albar et al., *Adsorption of Hexacontane on Hexagonal Boron Nitride*, J. Phys. Chem. C **122** (2018) 27575.

[79] S. Sippola, M. Todorović and E. Peltola, *First-Principles Structure Search Study of 17-β-Estradiol Adsorption on Graphene*, ACS Omega **9** (2024) 34684, ISSN: 24701343.

[80] S. Maier and M. Stöhr, *Molecular assemblies on surfaces: towards physical and electronic decoupling of organic molecules*, Beilstein J. Nanotechnol **12** (2021) 956.

[81] B. Öcal, P. Weitkamp, K. Meerholz and S. Olthof, *Chemical interaction and molecular growth of a highly dipolar merocyanine molecule on metal surfaces: A photoelectron spectroscopy study*, Surf. Sci. (2025) 122690.

[82] A. J. Kny et al., *Chiral self-organized single 2D-layers of tetramers from a functional donor-acceptor molecule by the surface template effect*, Nanoscale (2023) 10319, ISSN: 2040-3364.

[83] C. Gharbi, Y. Ajili, D. Ben Abdallah and M. Hochlaf, *Sodium isocyanide–Helium potential energy surface and astrophysical applications*, Theor. Chem. Acc. **140** (2021) 1, ISSN: 14322234.

[84] V. G. Ruiz, W. Liu and A. Tkatchenko, *Density-functional theory with screened van der Waals interactions applied to atomic and molecular adsorbates on close-packed and non-close-packed surfaces*, Phys. Rev. B **93** (2016) 035118, ISSN: 24699969.

[85] N. Altman and M. Krzywinski, *The curse(s) of dimensionality this-month*, Nat. Methods **15** (2018) 399, ISSN: 15487105.

[86] S. Mannan, V. Bihani, N. M. A. Krishnan and J. C. Mauro, *Navigating energy landscapes for materials discovery: Integrating modeling, simulation, and machine learning*, MGE Advances **2** (2024) e25, ISSN: 2940-9489.

[87] J. Rogers et al., *Bypassing the computational bottleneck of quantum-embedding theories for strong electron correlations with machine learning*, Phys. Rev. Research **3** (2021) 013101, ISSN: 26431564.

[88] V. Obersteiner, M. Scherbela, L. Hörmann, D. Wegner and O. T. Hofmann, *Structure Prediction for Surface-Induced Phases of Organic Monolayers: Overcoming the Combinatorial Bottleneck*, Nano Lett. **17** (2017) 4453, ISSN: 15306992.

[89] F. Musil et al., *Machine learning for the structure–energy–property landscapes of molecular crystals*, Chem. Sci. **9** (2018) 1289, ISSN: 2041-6539.

[90] P. L. Kang, C. Shang and Z. P. Liu, *Large-Scale Atomic Simulation via Machine Learning Potentials Constructed by Global Potential Energy Surface Exploration*, Acc. Chem. Res. **53** (2020) 2119, ISSN: 15204898.

[91] F. Lu et al., *Fast Near Ab Initio Potential Energy Surfaces Using Machine Learning*, J. Phys. Chem. A **126** (2022) 4013, ISSN: 15205215.

[92] M. Todorović, M. U. Gutmann, J. Corander and P. Rinke, *Bayesian inference of atomistic structure in functional materials*, npj Computational Materials **5** (2019), ISSN: 20573960.

[93] D. M. Packwood and T. Hitosugi, *Rapid prediction of molecule arrangements on metal surfaces via Bayesian optimization*, Appl. Phys. Express **10** (2017) 065502, ISSN: 1882-0786.

[94] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse and M. Bokdam, *Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference*, Phys. Rev. Lett. **122** (2019) 225701, ISSN: 10797114.

[95] J. Järvi, P. Rinke, M. Todorović and S. Maier, *Detecting stable adsorbates of (1S)-camphor on Cu(111) with Bayesian optimization "Molecular assemblies on surfaces-towards physical and electronic decoupling of organic molecules"*, Beilstein J. Nanotechnol **2020** (2020) 1577.

[96] L. Fang, E. Makkonen, M. Todorović, P. Rinke and X. Chen, *Efficient Amino Acid Conformer Search with Bayesian Optimization*, J. Chem. Theory Comput. **17** (2021) 1955, ISSN: 15499626.

[97] L. Fang, X. Guo, M. Todorović, P. Rinke and X. Chen, *Exploring the Conformers of an Organic Molecule on a Metal Cluster with Bayesian Optimization*, J. Chem. Inf. Model. **63** (2023) 745, ISSN: 1549960X.

[98] J. Baumsteiger, L. Celiberti, P. Rinke, M. Todoroví and C. Franchini, *Exploring noncollinear magnetic energy landscapes with Bayesian optimization*, Digital Discovery **4** (2025) 1639.

[99] S. Packman, N. Riva and P. Rodriguez-Fernandez, *Bayesian Methods for Magnetic and Mechanical Optimization of Superconducting Magnets for Fusion*, J. Fusion Energy **44** (2025) 1, ISSN: 15729591.

[100] *Research | BOSS https://sites.utu.fi/boss/research/*.

[101] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner and G. Csanyi, "MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields", *Advances in Neural Information Processing Systems*, ed. by S. Koyejo et al., vol. 35, Curran Associates, Inc., 2022 11423.

[102] R. Drautz, *Atomic cluster expansion for accurate and transferable interatomic potentials*, Phy. Rev. B **99** (2019), ISSN: 24699969.

[103] Y. Lysogorskiy et al., *Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon*, npj Comput. Mater. **7** (2021) 1, ISSN: 20573960.

[104] V. L. Deringer, M. A. Caro and G. Csányi, *A general-purpose machine-learning force field for bulk and nanostructured phosphorus*, Nat. Comm. **11** (2020) 1, ISSN: 20411723.

[105] Z. El-Machachi, M. Wilson and V. L. Deringer, *Exploring the configurational space of amorphous graphene with machine-learned atomic energies*, Chem. Sci. **13** (2022) 13720, ISSN: 20416539.

[106] G. Yang, Y. B. Liu, L. Yang and B. Y. Cao, *Machine-learned atomic cluster expansion potentials for fast and quantum-accurate thermal simulations of wurtzite AlN*, J. Appl. Phys. **135** (2024) 85105, ISSN: 0021-8979.

[107] L. Hörmann, W. G. Stark and R. J. Maurer, *Machine learning and data-driven methods in computational surface and interface science*, npj Comput. Mater. **11** (2025) 1, ISSN: 20573960.

[108] S. Jeschke, K. Wilson and A. F. Lee, *Accelerating Computational Modeling of Reactant Adsorption through a Combined MACE+DFT Approach: Furfural on Cu Surfaces*, J. Phys. Chem. C **129** (2025) 11948.

[109] J. S. Jestilä, N. Wu, F. Priante and A. S. Foster, *Accelerated Lignocellulosic Molecule Adsorption Structure Determination*, J. Chem. Theory Comput. **20** (2024) 2297, ISSN: 15499626.

[110] D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction*, John Wiley and Sons, 2009 1, ISBN: 9780470373170.

[111] C. Fiolhais, F. Nogueira and M. Marques, *A Primer in Density Functional Theory*, Springer, 2003.

[112] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd, Wiley-VCH, 2001.

[113] F. Jensen, *Introduction to Computational Chemistry*, John wiley & sons, 2017.

[114] P. Atkins and R. Friedman, *Molecular Quantum Mechanics*, 4th ed., Oxford University Press, 2005.

[115] P. Fleming, *Quantum Chemistry with Applications in Spectroscopy*.

[116] P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev. B (1964).

[117] A. J. H. Frinjs, C. C. M. Rindt and S. V. Gaastra-Nedea, "Modeling thermochemical reactions in thermal energy storage systems ", *Advances in Thermal Energy Storage Systems*, 2nd ed., Woodhead Publishing Series in Energy, 2021 497.

[118] W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys Rev A **140** (1965) 1133.

[119] F. Della Sala, E. Fabiano and L. A. Constantin, *Kinetic-energy-density dependent semilocal exchange-correlation functionals*, Int. J. Quantum Chem. **116** (2016) 1641, ISSN: 1097461X.

[120] N. Mardirossian and M. Head-Gordon, *Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals*, Mol. Phys. **115** (2017) 2315, ISSN: 13623028.

[121] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons Ltd, 2004.

[122] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, 2004.

[123] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Harcourt College Publishers, 1976, ISBN: # 0-03-083993-9.

[124] J. Kohanoff, *Electronic Structure Calculations for Solids and Molecules*, Cambridge University Press, 2006.

[125] P. E. Blöchl, C. J. Först and J. Schimpl, *The Projector Augmented Wave Method: ab-initio molecular dynamics with full wave functions*, Bull. Mater. Sci **26** (2003) 33.

[126] D. R. Hamann, M. Schlüter and C. Chiang, *Norm-Conserving Pseudopotentials*, Phys. Rev. Lett. **43** (1979) 1494, ISSN: 00319007.

[127] H. J. Monkhorst and J. D. Pack, *Special points for Brillonin-zone integrations*, Phys. Rev. B **13** (1976) 5188.

[128] G. Kresse and J. Hafner, *Ab Initio Molecular Dynamics for Liquid Metals*, Phys. Rev. B **47** (1993) 558.

[129] G. Kresse and J. Furthmüller, *Efficiency of Ab-initio Total Energy Calculations for Metals and Semiconductors Using a Plane-wave Basis Set*, Comput. Mater. Sci. **6** (1996) 15.

[130] A. D. Becke, *Density-fnnctional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A **38** (1988).

[131] J. P. Perdew, K. Burke and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77** (1996) 3865.

[132] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, J. Chem. Phys **132** (2010), ISSN: 00219606.

[133] E. Caldeweyher, C. Bannwarth and S. Grimme, *Extension of the D3 dispersion coefficient model*, J. Chem. Phys. **147** (2017) 34112, ISSN: 00219606.

[134] S. Grimme, C. Bannwarth and P. Shushkov, *A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1-86)*, J. Chem. Theory Comput. **13** (2017) 1989, ISSN: 15499626.

[135] C. Bannwarth, S. Ehlert and S. Grimme, *GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions*, J. Chem. Theory Comput. **15** (2019) 1652, ISSN: 15499626.

[136] M. Elstner et al., *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, Phys. Rev. B **58** (1998) 7260.

[137] P. Pracht, F. Bohle and S. Grimme, *Automated exploration of the low-energy chemical space with fast quantum chemical methods*, Phys. Chem. Chem. Phys. **22** (2020) 7169, ISSN: 14639076.

[138] E. Runge and E. K. U. Gross, *Density-Functional Theory for Time-Dependent Systems*, Physical Review Letters **52** (1984) 997.

[139] R. Van Leeuwen, *Causality and Symmetry in Time-Dependent Density-Functional Theory*, Phys. Rev. Lett. **80** (1998) 1280.

[140] M. E. Casida, " Time-Dependent Density Functional Response Theory for Molecules", *Recent Advances in Density Functional Methods*, ed. by D. P. Chong, vol. 1, 1995 155.

[141] M. Petersilka, U. J. Gossmann and E. K. U. Gross, *Excitation Energies from Time-Dependent Density-Functional Theory*, Phys. Rev. Lett **76** (1996) 1212.

[142] N. T. Maitra, *Perspective: Fundamental aspects of time-dependent density functional theory*, J. Chem. Phys. **144** (2016), ISSN: 00219606.

[143] M. E. Casida, *Time-dependent density-functional theory for molecules and molecular solids*, J. Mol. Struct. THEOCHEM **914** (2009) 3, ISSN: 01661280.

[144] J. M. Herbert, "Density-functional theory for electronic excited states", *Theoretical and Computational Photochemistry: Fundamentals, Methods, Applications and Synergy with Experimental Approaches*, Elsevier, 2023, chap. 3 69, ISBN: 9780323917384.

[145] F. Neese, *Software update: The ORCA program system—Version 5.0*, Wiley Interdisciplinary Reviews: Computational Molecular Science **12** (2022), ISSN: 17590884.

[146] P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields*, J. Phys. Chem. **98** (1994) 11623.

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

[148] R. Bauernschmitt and R. Ahlrichs, *Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory*, Chem. Phys. Lett. **256** (1996) 454.

[149] V. Barone and M. Cossi, *Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model*, J. Phys. Chem. A **102** (1998) 1995.

[150] A. V. Marenich, C. J. Cramer and D. G. Truhlar, *Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions*, J. Phys. Chem. B **113** (2009) 6378, ISSN: 15206106.

[151] S. Ghosh and K. Bhattacharyya, *Origin of the Failure of Density Functional Theories in Predicting Inverted SingletTriplet Gaps*, J. Phys. Chem. A **2022** (2022) 1378.

[152] A. Prlj, B. F. E. Curchod, A. Fabrizio, L. Floryan and C. C. Corminboeuf, *Qualitatively Incorrect Features in the TDDFT Spectrum of Thiophene-Based Compounds*, J. Phys. Chem. Lett **6** (2015) 13.

[153] L. Bernasconi et al., *First-principles optical response of semiconductors and oxide materials*, Physical Review B **83** (2011) 195325, ISSN: 10980121.

[154] L. Hedin, *New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem*, Phys. Rev **139** (1965) 796.

[155] F. Aryasetiawan and O. Gunnarsson, *The GW method*, Rep. Prog. Phys **61** (1998) 237.

[156] M. Rohlfing and S. G. Louie,
Electron-hole excitations and optical spectra from first principles,
Phys. Rev. B **62** (2000) 4927.

[157] G. Onida, L. Reining and A. Rubio,
Electronic excitations: density-functional versus many-body Green's-function approaches,
Rev. Mod. Phys. **74** (2002) 601.

[158] X. Blase, I. Duchemin and D. Jacquemin, *The Bethe-Salpeter equation in chemistry: relations with TD-DFT, applications and challenges*, Chem. Soc. Rev **47** (2018) 1039.

[159] L. Bernasconi, R. Webster, S. Tomic and N. M. Harrison, *Optical response of extended systems from time-dependent Hartree-Fock and time-dependent density-functional theory*,
J. Phys.: Conf. Ser. **367** (2012) 012001.

[160] S. Tomić, L. Bernasconi, B. G. Searle and N. M. Harrison,
Electronic and optical structure of wurtzite CuInS2, J. Phys. Chem. C **118** (2014) 14478,
ISSN: 19327455.

[161] A. Živković, N. H. De Leeuw, B. G. Searle and L. Bernasconi,
Electronic Excitations in Copper Oxides: Time-Dependent Density Functional Theory Calculations with a Self-Consistent Hybrid Kernel, J. Phys. Chem. C **124** (2020) 24995,
ISSN: 19327455.

[162] S. Brahms and S. Nikitine,
Intrinsic absorption and reflection of cuprous oxide in the 2.5 to 6.5 eV region,
Solid State Communications **3** (1965) 209, ISSN: 0038-1098.

[163] R. Tomar, L. Bernasconi, D. Fazzi and T. Bredow,
Theoretical Study on the Optoelectronic Properties of Merocyanine-Dyes,
J. Phys. Chem. A **127** (2023) 9661, ISSN: 15205215.

[164] L. Yu, Q. Yan and A. Ruzsinszky,
Key role of antibonding electron transfer in bonding on solid surfaces,
Phys. Rev. Mater. **3** (2019) 092801.

[165] N. T. Nguyen, D. Q. Ho and N. T. Trung, *Theoretical insights into the adsorption and gas sensing performance of Fe/Cu-adsorbed graphene †*,
Phys. Chem. Chem. Phys **26** (2024) 14276.

[166] A. Chakraborthy et al., *A Critical Review of the Use of Graphene-Based Gas Sensors*,
Chemosensors **10** (2022), ISSN: 22279040.

[167] S. Kaneko et al.,
Site-Selection in Single-Molecule Junction for Highly Reproducible Molecular Electronics,
J. Am. Chem. Soc. **138** (2016) 1294, ISSN: 15205126.

[168] A. Jeindl et al., *Nonintuitive Surface Self-Assembly of Functionalized Molecules on Ag(111)*,
ACS Nano **15** (2021) 6723.

[169] D. M. Bylander, L. Kleinman and S. Lee,
Self-consistent calculations of the energy bands and bonding properties of B,zC,
Phys. Rev. B **42** (1990) 15.

[170] D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*,
Mathematical Programming **45** (1989) 503.

[171] J. Nocidal and S. J. Wright, *Numerical Optimization*, 2nd, Springer, 2000 136.

[172] D. C. Liu and J. Nocedal,
On the Limited Memory BFGS Method for Large Scale Optimization,
Mathematical Programming **45** (1989) 503.

[173] B. G. Pfrommer, M. Cô, S. G. Louie and M. L. Cohen,
Relaxation of Crystals with the Quasi-Newton Method, J. Comp. Phys. **131** (1997) 233.

[174] D. Packwood et al., *A universal preconditioner for simulating condensed phase materials*,
J. Chem. Phys. **144** (2016), ISSN: 00219606.

[175] J. Tersoff and D. R. Hamann, *Theory of the scanning tunneling microscope*,
Phys. Rev. B **31** (1985) 805.

[176] F. Karsai, M. Humer, E. Flage-Larsen, P. Blaha and G. Kresse, *Effects of electron-phonon coupling on absorption spectrum: K edge of hexagonal boron nitride*,
Phys. Rev. B **98** (2018) 235205, ISSN: 24699969.

[177] M. Unzog, A. Tal and G. Kresse, *X-ray absorption using the projector augmented-wave method and the Bethe-Salpeter equation*, Physical Review B **106** (2022), ISSN: 24699969.

[178] R. Jacobs et al.,
A practical guide to machine learning interatomic potentials – Status and future,
Current Opinion in Solid State and Materials Science **35** (2025) 101214, ISSN: 13590286.

[179] J. Nigam, S. Pozdnyakov, G. Fraux and M. Ceriotti, *Unified theory of atom-centered representations and message-passing machine-learning schemes*,
Journal of Chemical Physics **156** (2022), ISSN: 10897690.

[180] I. Batatia et al., *A foundation model for atomistic materials chemistry*, (2024).

[181] J. Järvi, P. Rinke, M. Todorović and S. Maier, *Detecting stable adsorbates of (1S)-camphor on Cu(111) with Bayesian optimization "Molecular assemblies on surfaces-towards physical and electronic decoupling of organic molecules"*, Beilstein J. Nanotechnol **11** (2020) 1577.

[182] S. Sippola, M. Todorović and E. Peltola,
First-Principles Structure Search Study of 17-β-Estradiol Adsorption on Graphene,
ACS Omega **9** (2024) 34684, ISSN: 24701343.

[183] V. Barone and M. Cossi, *Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model*, J. Phys. Chem. A **102** (1998) 1995.

[184] H. Ahmad et al., *First-principles investigation of Sc and Ti-decorated hBN monolayers as adsorbents and gas sensors for SF6 decomposition products*, Chemical Physics **595** (2025),
ISSN: 03010104.

[185] M. I. Khan et al.,
Exploring the sensing potential of Fe-decorated h-BN toward harmful gases: a DFT study,
RSC Advances **14** (2024) 7040, ISSN: 20462069.

[186] R. Ponce-Pérez, G. H. Cocoletzi and N. Takeuchi,
Two-dimensional boron nitride structures functionalization: first principles studies,
Journal of Molecular Modeling **22** (2016), ISSN: 09485023.

[187] M. Neek-Amal, J. Beheshtian, A. Sadeghi, K. H. Michel and F. M. Peeters,
Boron nitride monolayer: A strain-tunable nanosensor,
Journal of Physical Chemistry C **117** (2013) 13261, ISSN: 19327447.

[188] A. Matković et al., *Epitaxy of highly ordered organic semiconductor crystallite networks supported by hexagonal boron nitride OPEN*, Sci. Rep **6** (2016) 38519.

[189] Y. Zhang et al., *Probing Carrier Transport and Structure-Property Relationship of Highly Ordered Organic Semiconductors at the Two-Dimensional Limit*,
Phys. Rev. Lett **116** (2016) 016602.

[190] K. Zuo et al., *Ultrahigh resistance of hexagonal boron nitride to mineral scale formation*,
Nat. Comm. **13** (2022) 4523.

[191] F. Gao and E. Reichmanis, *Introduction: Emerging Materials for Optoelectronics*,
Chem. Rev. **123** (2023) 10835, ISSN: 15206890.

[192] T. Tan et al., *2D Material Optoelectronics for Information Functional Device Applications: Status and Challenges*, Adv. Sci. **7** (2020) 2000058, ISSN: 2198-3844.

[193] O. Ostroverkhova, *Organic Optoelectronic Materials: Mechanisms and Applications*,
Chem. Rev. **116** (2016) 13279, ISSN: 15206890.

[194] L. X. Chen, *Organic Solar Cells: Recent Progress and Challenges*,
ACS Energy Lett **4** (2019) 2537, ISSN: 23808195.

[195] D. Xiang, X. Wang, C. Jia, T. Lee and X. Guo,
Molecular-Scale Electronics: From Concept to Function, Chem. Rev. **116** (2016) 4318,
ISSN: 15206890.

[196] I. Burghardt, *First-Principles Simulations of Molecular Optoelectronic Materials: Elementary Excitations and Spatiotemporal Dynamics*,
J. Chem. Theory Comput. **21** (2025) 3297, ISSN: 15499626.

[197] B. Champagne, M. Guillaume and F. Zutterman,
TDDFT investigation of the optical properties of cyanine dyes,
Chemical Physics Letters **425** (2006) 105, ISSN: 00092614.

[198] A. J. (F.-W.-U. B. Kny,
Adsorption and Ordering of the Merocyanine HB238 on the Ag(100) Surface,
PhD thesis, 2020 67.

[199] H. Bürckstümmer et al., *Efficient solution-processed bulk heterojunction solar cells by antiparallel supramolecular arrangement of dipolar donor-acceptor dyes*,
Angew. Chem. Int. Ed. **50** (2011) 11628, ISSN: 14337851.

[200] S. R. Marder et al.,
A Unified Description of Linear and Nonlinear Polarization in Organic Polymethine Dyes,
Science **265** (1994) 632.

[201] K. Meerholz, B. Volodlin, Sandalphon, B. Kippelen and N. Peyghambarian,
A photorefractive polymer with high optical gain and diffraction efficiency near 100%,
Nature **371** (1994) 497.

[202] S. R. Marder et al., *Large First Hyperpolarizabilities in Push-Pull Polyenes by Tuning of the Bond Length Alternation and Aromaticity*, Science **263** (1994) 511.

[203] V. Parthasarathy et al., *Combination of Cyanine Behaviour and Giant Hyperpolarisability in Novel Merocyanine Dyes: Beyond the Bond Length Alternation (BLA) Paradigm*,
Chem. Eur. J. **21** (2015) 14211, ISSN: 15213765.

[204] F. Würthner et al., *Merocyanine Dyes in the Cyanine Limit: A New Class of Chromophores for Photorefractive Materials*, Angew. Chem. Int. Ed. Engl. **36** (1997) 2765, ISSN: 1521-3773.

[205] C. B. Gorman and S. R. Marder, *An investigation of the interrelationships between linear and nonlinear polarizabilities and bond-length alternation in conjugated organic molecules*,
Proc. Natl. Acad. Sci. USA **90** (1993) 11297.

[206] F. Würthner, R. Wortmann and K. Meerholz,
Chromophore Design for Photorefractive Organic Materials, Chem. Phys. Chem. **3** (2002) 17,
ISSN: 1439-4235.

[207] B. Tirri et al., *A combined Monte Carlo/DFT approach to simulate UV-vis spectra of molecules and aggregates: Merocyanine dyes as a case study*,
J. Comput. Chem **42** (2021) 1054, ISSN: 1096987X.

[208] C. Brü and B. Engels, *Benchmarking Ground-State Geometries and Vertical Excitation Energies of a Selection of P-Type Semiconducting Molecules with Different Polarity*,
J. Chem. Phys. A **119** (2015) 12876.

[209] I. Ciofini, C. Adamo and H. Chermette, *Effect of self-interaction error in the evaluation of the bond length alternation in trans-polyacetylene using density-functional theory*,
J. Chem. Phys. **123** (2005) 224103.

[210] D. Jacquemin, E. A. Perpète, I. Ciofini and C. Adamo,
Assessment of recently developed density functional approaches for the evaluation of the bond length alternation in polyacetylene, Chemical Physics Letters **405** (2005) 376,
ISSN: 00092614.

[211] H. Mustroph et al., *The effects of substituents and solvents on the ground-state π-electronic structure and electronic absorption spectra of a series of model merocyanine dyes and their theoretical interpretation*, Chem. Eur. J **18** (2012) 8140, ISSN: 09476539.

[212] C. Brückner et al., *Structure-Property Relationships for Exciton and Charge Reorganization Energies of Dipolar Organic Semiconductors: A Combined Valence Bond Self-Consistent Field and Time-Dependent Hartree-Fock and DFT Study of Merocyanine Dyes*,
J. Phys. Chem. C **119** (2015) 17602.

[213] M. Guillaume, B. B. Champagne and F. Zutterman, *Investigation of the UV/Visible Absorption Spectra of Merocyanine Dyes Using Time-Dependent Density Functional Theory*, J. Phys. Chem. A **110** (2006) 13007.

[214] A. V. Kulinich, E. K. Mikitenko and A. A. Ishchenko, *Synthesis, electronic structure and spectral fluorescent properties of vinylogous merocyanines derived from 1,3-dialkyl-benzimidazole and malononitrile*, Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy **171** (2017) 317, ISSN: 13861425.

[215] G. Zhao, Y. Yang, Zhang Chengyuan, Y. Song and Y. Li, *The theoretical study of excited-state intramolecular proton transfer of N, N,-bis (salicylidene)-(2-(34 -diaminophenyl) benzothiazole)*, J. Lumin **230** (2021) 117741.

[216] E. Salpeter and H. Bethe, *A Relativistic Equation for Bound-State Problems*, Phys. Rev. **84** (1951) 1232.

[217] B. O. Roos, P. R. Taylor and P. E. Sigbahn, *A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach*, Chemical Physics **48** (1980) 157, ISSN: 03010104.

[218] R. Berraud-Pache, F. Neese, G. Bistoni, R. Róbert and I. Izsák, *Unveiling the Photophysical Properties of Boron-dipyrromethene Dyes Using a New Accurate Excited State Coupled Cluster Method*, J. Chem. Theory Comput. **16** (2020) 564.

[219] R. Webster, L. Bernasconi and N. M. Harrison, *Optical properties of alkali halide crystals from all-electron hybrid TD-DFT calculations*, J. Chem. Phys. **142** (2015) 214705, ISSN: 00219606.

[220] J. P. Perdew, K. Burke and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77** (1996) 3865.

[221] C. Adamo and V. Barone, *Toward reliable density functional methods without adjustable parameters: The PBE0 model.*, J. Chem. Phys. **110** (1999) 4775.

[222] A. D. Becke, *Density-functional thermochemistry. III. The role of exact exchange*, J. Chem. Phys. **98** (1993) 5648.

[223] C. Lee, e. Yang and R. G. Parr, *Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B **37** (1988) 785.

[224] S. H. Vosko, L. Wilk and M. Nusair, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Can. J. Phys. **58** (1980) 1200.

[225] T. Yanai, D. P. Tew and N. C. Handy, *A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP)*, Chemical Physics Letters **393** (2004) 51, ISSN: 00092614.

[226] J.-D. Chai and M. Head-Gordon, *Systematic optimization of long-range corrected hybrid density functionals*, J. Chem. Phys. **128** (2008), ISSN: 0021-9606.

[227] S. Grimme and F. Neese,
Double-hybrid density functional theory for excited electronic states of molecules,
J. Chem.Phys. **127** (2007) 154116.

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

[229] F. Neese, *Software update: the ORCA program system, version 4.0*,
WIREs Comput Mol Sci **8** (2018), ISSN: 1759-0876.

[230] F. Bruneval et al.,
MOLGW 1: Many-body perturbation theory software for atoms, molecules, and clusters,
Computer Physics Communications **208** (2016) 149, ISSN: 00104655.

[231] T. H. Dunning, *Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen*, J. Chem. Phys. **90** (1989) 1007.

[232] F. Bruneval, S. M. Hamed and J. B. Neaton,
A systematic benchmark of the ab initio Bethe-Salpeter equation approach for low-lying optical excitations of small organic molecules, J. Chem. Phys. **142** (2015) 244101,
ISSN: 00219606.

[233] R. Dovesi et al., *CRYSTAL14: A program for the ab initio investigation of crystalline solids*,
Int. J. Quant. Chem. **114** (2014) 1287, ISSN: 1097461X.

[234] J. Laun and T. Bredow, *BSSE-corrected consistent Gaussian basis sets of triple-zeta valence with polarization quality of the fifth period for solid-state calculations*,
J. Comput. Chem., **43** (2022) 839, ISSN: 1096987X.

[235] B. Kirtman, F. L. Gu and D. M. Bishop,
Extension of the Genkin and Mednis treatment for dynamic polarizabilities and hyperpolarizabilities of infinite periodic systems. I. Coupled perturbed Hartree–Fock theory,
J. Chem. Phys **113** (2000) 1294, ISSN: 0021-9606.

[236] M. Ferrero, M. Rérat, R. Orlando and R. Dovesi, *The calculation of static polarizabilities of 1-3D periodic compounds. the implementation in the crystal code*,
J. Comput. Chem. **29** (2008) 1450, ISSN: 01928651.

[237] S. Grimme and A. Hansen,
A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects,
Angew. Chem. Int. Ed. **54** (2015) 12308, ISSN: 1433-7851.

[238] C. A. Bauer, A. Hansen and S. Grimme, *The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure*,
Chem. Eur. J. **23** (2017) 6150, ISSN: 0947-6539.

[239] D. Bialas, C. Zhong, F. Würthner and F. C. Spano, *Essential states model for merocyanine dye stacks: bridging electronic and optical absorption properties*,
J. Phys. Chem. C **123** (2019) 18654, ISSN: 19327455.

[240] A. Arjona-Esteban et al., *Influence of Solid-State Packing of Dipolar Merocyanine Dyes on Transistor and Solar Cell Performances*, J. Am. Chem. Soc. **137** (2015) 13524.

[241] J. Hoche et al., *The origin of the solvent dependence of fluorescence quantum yields in dipolar merocyanine dyes*, Chem. Sci. **10** (2019) 11013.

[242] N. Gildemeister et al.,
Understanding the structural and charge transport property relationships for a variety of merocyanine single-crystals: a bottom up computational investigation,
J. Mater. Chem. C **9** (2021) 10851, ISSN: 2050-7526.

[243] L. Fu, E. Yaschenko, L. Resca and R. Resta, *Hartree-Fock approach to macroscopic polarization: Dielectric constant and dynamical charges of KNbO3*,
Phys. Rev. B **57** (1998) 6967, ISSN: 0163-1829.

[244] P. I. Djurovich, E. I. Mayo, S. R. Forrest and M. E. Thompson, *Measurement of the lowest unoccupied molecular orbital energies of molecular organic semiconductors*,
Organic Electronics **10** (2009) 515, ISSN: 15661199.

[245] S. Tadayyon et al., *Reliable and reproducible determination of work function and ionization potentials of layers and surfaces relevant to organic light emitting diodes*,
Organic Electronics **5** (2004) 199, ISSN: 15661199.

[246] D. Jacquemin et al., *Assessment of the efficiency of long-range corrected functionals for some properties of large compounds*, J. Chem. Phys. **126** (2007) 144105.

[247] L. Goerigk and M. Casanova-Paéz, *The Trip to the Density Functional Theory Zoo Continues: Making a Case for Time-Dependent Double Hybrids for Excited-State Problems*,
Aust. J. Chem. **74** (2021) 3, ISSN: 14450038.

[248] F. Santoro and D. Jacquemin,
Going beyond the vertical approximation with time-dependent density functional theory,
WIREs Comput. Mol. Sci. **6** (2016) 460, ISSN: 17590884.

[249] E. Wigner, *On the Quantum Correction For Thermodynamic Equilibrium*,
Phys. Rev. (1932) 749.

[250] C. Fang, B. Oruganti and B. Durbeej, *How method-dependent are calculated differences between vertical, adiabatic, and 0-0 excitation energies?*, J. Phys. Chem. A **118** (2014) 4157,
ISSN: 15205215.

[251] H. C. Georg, K. Coutinho and S. Canuto,
Solvent effects on the UV-visible absorption spectrum of benzophenone in water: A combined Monte Carlo quantum mechanics study including solute polarization,
J. Chem. Phys. **126** (2007) 034507.

[252] K. Coutinho and S. Canuto, *Solvent effects in emission spectroscopy: A Monte Carlo quantum mechanics study of the n]* shift of formaldehyde in water*, J. Chem. Phys **113** (2000) 9132.

[253] S. Grimme,
Semiempirical hybrid density functional with perturbative second-order correlation,
J. Chem. Phys **124** (2006) 034108.

[254] R. Send, O. Valsson and C. Filippi, *Electronic Excitations of Simple Cyanine Dyes: Reconciling Density Functional and Wave Function Methods*,
J. Chem. Theory Comput. **7** (2011) 444.

[255] P. Boulanger, D. Jacquemin, I. Duchemin and X. Blase, *Fast and accurate electronic excitations in cyanines with the many-body bethe-salpeter approach*, J. Chem. Theory Comput. **10** (2014) 1212, ISSN: 15499626.

[256] C. Faber, P. Boulanger, C. Attaccalite, I. Duchemin and X. Blase, *Excited states properties of organic molecules: From density functional theory to the GW and Bethe-Salpeter Green's function formalisms*, Phil.Trans.R.Soc.A **372** (2014) 20130271, ISSN: 1364503X.

[257] M. Piantek et al., *Reversing the thermal stability of a molecular switch on a gold surface: Ring-opening reaction of nitrospiropyran*, J. Am. Chem. Soc. **131** (2009) 12729, ISSN: 00027863.

[258] A. Riemann, L. Browning and H. Goff, *Conformational behavior of naphtho-merocyanine dimers on Au(111)*, Surf. Sci. **709** (2021) 121837, ISSN: 0039-6028.

[259] A. Das et al., *Conformational Selectivity of Merocyanine on Nanostructured Silver Films: Surface Enhanced Resonance Raman Scattering (SERRS) and Density Functional Theoretical (DFT) Study nanostructured surfaces can be further explored for energy efficient and economical separation of geometrical isomers*, Front. Chem. **10** (2022) 902585.

[260] A. Riemann, L. Rankin and D. Henry, *Atomic Charge Dependency of Spiropyran/Merocyanine Adsorption as a Precursor to Surface Isomerization Reactions*, ACS Omega **9** (2024) 798, ISSN: 24701343.

[261] B. Hourahine et al., *DFTB+, a software package for efficient approximate density functional theory based atomistic simulations*, J. Chem. Phys **152** (2020) 124101.

[262] G. Kresse and D. Joubert, *From Ultrasoft Pseudopotentials to the Projector Augmented-wave Method*, Phys. Rev. B **59** (1999) 1758.

[263] S. Moussadeq et al., *Insight into the effect of alloying on the adsorption of benzene on Ag(100) surface: DFT calculations*, Surf. Sci. **734** (2023) 122321, ISSN: 00396028.

[264] N. G. Limas and T. A. Manz, *Introducing DDEC6 atomic population analysis: Part 4. Efficient parallel computation of net atomic charges, atomic spin moments, bond orders, and more*, RSC Advances **8** (2018) 2678, ISSN: 20462069.

[265] G. A. Jeffrey, *An introduction to hydrogen bonding*, Oxford University Press, 1997.

[266] S. Leung, *STM-2DScan.py*, https://github.com/ShuangLeung/STM_2DScan (2019).

[267] K. Momma and F. Izumi, *VESTA3 for three-dimensional visualization of crystal, volumetric and morphology data*, J. Appl. Crystallogr. **44** (2011) 1272.

[268] G. Marek and H. Jürgen, *CO adsorption on Cu(1 1 1) and Cu(0 0 1) surfaces: Improving site preference in DFT calculations*, Surf. Sci. **590** (2005) 117.

[269] J. Heyd, G. E. Scuseria and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential*, J. Chem. Phys. **118** (2003) 8207, ISSN: 00219606.

[270] A. A. Mostofi et al.,
An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions,
Comput. Phys. Commun. **185** (2014) 2309, ISSN: 00104655.

[271] F. H. Farwick zum Hagen et al., *Structure and Growth of Hexagonal Boron Nitride on Ir(111)*,
ACS Nano **10** (2016) 11012.

[272] R. Schäfer, P. Weitkamp, O. Erdene-Ochir, K. Meerholz and K. Lindfors,
Polarization-controlled strong light-matter interaction with templated molecular aggregates,
29 (2025).

[273] O. Bushman and A. Riemann, *Merocyanine and Spiropyran Adsorption on Graphene*,
ACS Omega **10** (2025) 27082.

[274] H. Jung, L. Sauerland, S. Stocker, K. Reuter and J. T. Margraf,
Machine-learning driven global optimization of surface adsorbate geometries,
npj Computational Materials **9** (2023) 114.

[275] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, *Optimization by Simulated Annealing*,
Science **220** (1983) 4598.

[276] R. L. Johnston,
Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries,
J. Chem. Soc., Dalton Trans. **3** (2003) 4193, ISSN: 14779226.

[277] D. J. Wales and J. P. K. Doye, *Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms*,
J. Phys. Chem. A (1997) 5111.

[278] S. Goedecker, *Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems*, J. Chem. Phys. **120** (2004) 9911,
ISSN: 00219606.

[279] A. Laio and M. Parrinello, *Escaping free-energy minima*,
Proc. Natl. Acad. Sci **20** (2002) 12562.

[280] Z. Li and H. A. Scheraga,
Monte Carlo-minimization approach to the multiple-minima problem in protein folding,
Proc. Natl. Acad. Sci **84** (1987) 6611.

[281] B. Mortazavi, X. Zhuang, T. Rabczuk and A. V. Shapeev, *Atomistic modeling of the mechanical properties: the rise of machine learning interatomic potentials*,
Mater. Horiz. **10** (2023) 1956.

[282] D. Péter Kovács, I. Batatia, E. Sára Arany, G. Csányi and J. Chem Phys, *Evaluation of the MACE force field architecture: From medicinal chemistry to materials science*,
J. Chem. Phys. **159** (2023) 044118.

[283] Y. Wu, A. Walsh and A. M. Ganose,
Race to the bottom: Bayesian optimisation for chemical problems,
Digital Discovery **3** (2024) 1086.

[284] A. Hjorth Larsen et al.,
The atomic simulation environment—a Python library for working with atoms,
Journal of Physics: Condensed Matter **29** (2017) 273002, ISSN: 0953-8984.

[285] P. E. Blöchl, *Projector augmented-wave method*, Phys. Rev. B **50** (1994) 17953.

[286] R. S. Pease, *An X-ray study of boron nitride*, Acta Cryst. **5** (1952) 356, ISSN: 0365-110X.

[287] H. Yamane et al.,
Does the molecular orientation induce an electric dipole in Cu-phthalocyanine thin films?,
J. Appl. Phys. **99** (2006) 093705, ISSN: 00218979.

[288] T. Schultz et al., *Reliable Work Function Determination of Multicomponent Surfaces and Interfaces: The Role of Electrostatic Potentials in Ultraviolet Photoelectron Spectroscopy*,
Adv. Mater. Interfaces **4** (2017) 1700324, ISSN: 21967350.

[289] H. Ding and Y. Gao, *Electronic structure at rubrene metal interfaces*,
Appl. Phys. A Materi. Sci. Process. **95** (2009) 89, ISSN: 09478396.

[290] S. Duhm et al., *PTCDA on Au(1 1 1), Ag(1 1 1) and Cu(1 1 1): Correlation of interface charge transfer to bonding distance*, Org. Electron. **9** (2008) 111, ISSN: 15661199.

[291] P. G. Schroeder, C. B. France, J. B. Park and B. A. Parkinson,
Energy level alignment and two-dimensional structure of pentacene on Au(111) surfaces,
J. Appl. Phys. **91** (2002) 3010, ISSN: 00218979.

[292] Z. H. Cheng et al.,
Epitaxial growth of iron phthalocyanine at the initial stage on Au(111) surface,
J. Phys. Chem. C **111** (2007) 2656, ISSN: 19327447.

[293] E. Umbach, M. Sokolowski and R. Fink,
Substrate-interaction, long-range order, and epitaxy of large organic adsorbates,
Appl. Phys. A **63** (1996) 565.

[294] L. Romaner, D. Nabok, P. Puschnig, E. Zojer and C. Ambrosch-Draxl, *Theoretical study of PTCDA adsorbed on the coinage metal surfaces, Ag(111), Au(111) and Cu(111)*,
New J. Phys. **11** (2009) 053010, ISSN: 13672630.

[295] P. S. Bagus, V. Staemmler and C. Wöll,
Exchangelike Effects for Closed-Shell Adsorbates: Interface Dipole and Work Function,
Phys. Rev. Lett. **89** (2002), ISSN: 10797114.

[296] Q. Wei et al., *A-DAD-A non-fullerene acceptors for high-performance organic solar cells*,
Sci. China Chem. **63** (2020) 1352, ISSN: 18691870.

[297] J. Noh and M. Hara, *Nanoscopic evidence for dissociative adsorption of asymmetric disulfide self-assembled monolayers on Au(111)*, Langmuir **16** (2000) 2045, ISSN: 07437463.

[298] F. Würthner, *Dipole-Dipole Interaction Driven Self-Assembly of Merocyanine Dyes: From Dimers to Nanoscale Objects and Supramolecular Materials*,
Accounts of Chemical Research **49** (2016), ISSN: 15204898.

[299] R. Tomar, A. J. Kny, M. Sokolowski and T. Bredow,
Modeling the Monolayer Formation of Merocyanine HB238 on the Ag(100) Surface,
J. Phys. Chem. C (2025), ISSN: 19327455.

[300] P. A. Derosa, A. G. Zacarias and J. M. Seminario, "Application of Density Functional Theory to the Study and Design of Molecular Electronic Devices: the Metal-Molecule Interface",
Rev. Mod. Quant. Chem. 2002 1537.

[301] N. T. Maitra, *Charge transfer in time-dependent density functional theory*, J. Phys.: Condens. Matter **29** (2017) 423001.

[302] Z. Yang, R. Wu and J. A. Rodriguez, *First-principles study of the adsorption of sulfur on Pt(111): S core-level shifts and the nature of the Pt-S bond*, Phys. Rev. B **65** (2002) 1554091, ISSN: 01631829.

[303] G. Witte, S. Lukas, P. S. Bagus and C. Wöll, *Vacuum level alignment at organic/metal junctions: "cushion" effect and the interface dipole*, Appl. Phys. Lett. **87** (2005) 1, ISSN: 00036951.

[304] M. .-. Lemieux and A. .-. S. Tremblay, *Densities of states, projected densities of states, and transfer-matrix methods from a unified point of view*, Phys. Rev. B. **36** (1987) 1463.

[305] J. Gao, D. Nandi and M. Gupta, *Density functional theory - projected local density of states - based estimation of Schottky barrier for monolayer MoS2*, J. Appl. Phys. **124** (2018) 014502, ISSN: 10897550.

[306] J. Enkovaara et al., *Electronic structure calculations with GPAW: A real-space implementation of the projector augmented-wave method*, J. Phys. Condens. Matter **22** (2010) 253202, ISSN: 09538984.

List of Figures

List of Tables