

Toward Design of Novel Materials for Organic Electronics

Pascal Friederich, Artem Fediai, Simon Kaiser, Manuel Konrad, Nicole Jung,
and Wolfgang Wenzel*

Materials for organic electronics are presently used in prominent applications, such as displays in mobile devices, while being intensely researched for other purposes, such as organic photovoltaics, large-area devices, and thin-film transistors. Many of the challenges to improve and optimize these applications are material related and there is a nearly infinite chemical space that needs to be explored to identify the most suitable material candidates. Established experimental approaches struggle with the size and complexity of this chemical space. Herein, the development of simulation methods is addressed, with a particular emphasis on predictive multiscale protocols, to complement experimental research in the identification of novel materials and illustrate the potential of these methods with a few prominent recent applications. Finally, the potential of machine learning and methods based on artificial intelligence is discussed to further accelerate the search for new materials.

1. Introduction

Organic semiconductors have seen an increasing research interest ever since the invention of the organic light-emitting diode (OLED) in 1987 by Tang and VanSlyke.^[1] Current and future applications of organic semiconductors range from commercially available OLED displays^[2] and infrared applications^[3,4] over potentially printable organic^[1,5] and hybrid

organic/inorganic perovskites solar cells^[6] to printable electronic circuits based on organic field-effect transistors (OFETs).^[7] While OLED displays outperform their inorganic counterparts in terms of energy efficiency,^[8] scientific and technical challenges concerning the stability and processability of the organic materials used in large-area OLEDs and organic solar cells remain. New challenges arise from applications, such as displays on flexible substrates, OLED lightning, large area displays as well as for printable or solution processable larger area solar cells.^[8]

Many of the remaining challenges are material related, e.g., the low mobility of charge carriers in organic materials in general and in amorphous organic semiconductors in particular. There are other


materials related issues, such as limited OLED life times due to unstable blue hosts and emitters,^[9] low fill factors, and therefore reduced power conversion efficiencies of organic solar cells,^[10,11] low conductivity and high costs of organic charge transport layers of perovskite solar cells^[6,12,13] and low conductivity and hard processability of crystalline OFET materials.^[14] Conductivity and injection can be improved by doping the organic thin films with molecular dopants with high electron affinities (p-type)^[15] or low ionization energies (n-type).^[16] The doping mechanism of organic materials is in many cases not well understood,^[17] making material and device optimization a costly experimental endeavor.

The development of better materials is presently based on chemical insight, in part guided by theoretical understanding, or the experimental screening of large numbers of compounds. Given the size of the potentially available chemical space this remains a costly and time-consuming approach. Recent successes in experimental design of novel materials and concepts include the development of a stable strong molecular n-type dopant,^[16] a study about the quantitative relation between interaction parameter, miscibility, and function of conjugated polymer donors and small-molecule acceptors for bulk heterojunctions as used in organic solar cells^[18] the development of a universal strategy for ohmic hole injection into organic semiconductors with high ionization energies^[19] and many others.^[20–25] Another example is the development of strategies to harvest triplet excitons in organic light emitting diodes. For this purpose, novel classes of emitter molecules were developed, which include thermally activated delayed fluorescence (TADF)-based molecules,^[26–32] rotationally accessed spin-state inversion,^[33] and radical-based emitters.^[34]

Dr. P. Friederich, Dr. A. Fediai, S. Kaiser, M. Konrad, Prof. W. Wenzel
Institute of Nanotechnology (INT)
Karlsruhe Institute of Technology (KIT)
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen, Germany
E-mail: wolfgang.wenzel@kit.edu

Dr. P. Friederich
Department of Chemistry
University of Toronto
80 St. George Street, M5S 3H6 Toronto, Ontario, Canada

Dr. N. Jung
Institute of Organic Chemistry (IOC)
Karlsruhe Institute of Technology (KIT)
Fritz-Haber-Weg 6, 76131 Karlsruhe, Germany

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adma.201808256>.

© 2019 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1002/adma.201808256

This experimentally driven approach is increasingly augmented by model-based screening of compounds in various application domains. Examples are the materials genome project^[35] for battery materials, the Clean Energy Project for organic solar cell materials,^[36] the computational high-throughput screening of potential 2D materials^[37] or screening efforts for metal–organic frameworks.^[38] With respect to organic semiconductors, recent applications of computational materials design include the systematic development of design rules to improve the electron mobility of small molecule organic semiconductors,^[39] a successful demonstration of band structure engineering of organic semiconductors by tuning the quadrupole moment of the molecules^[40] and the discovery of new materials for thermally activated delayed fluorescence OLEDs in a high-throughput virtual screening approach.^[41]

In order to systematically design new materials in a targeted exploration of the chemical space, a systematic understanding of the relations between the chemical structure of molecules and their thin film properties is necessary. One challenge underlying all theoretical efforts to simulate organic electronics applications and to computationally design new materials is the inherent multiscale nature of all models of organic electronics (see **Figure 1**),^[42–45] Thin film and device properties not only depend on the molecular structure of the constituent materials but also on the microscopic arrangement of molecules as well as on the mesoscale structure formation (Figure 1a).^[46–49] The quantitative prediction of thin film properties like the charge carrier mobility (Figure 1b)^[39,50–53] or device properties (Figure 1c)^[54] requires knowledge about the electronic structure of each individual molecule in a disordered system comprising of thousands of molecules.^[55–57] The formation of the morphology depends on weak intermolecular forces acting on long time scales during film preparation. Presently there are no methods available or emerging that would unify treatment of all necessary length- and timescales into one simulation method. In the absence of such a method the only available solution to this challenge are tightly interconnected multiscale simulation protocols which connect the microscopic description of the electronic structure of single molecules to the device scale where multiple materials are used in sophisticated multilayer device architectures.

The material design challenge does not end with the prediction of one or several lead candidate materials. Chemical synthesis and experimental characterization of computationally designed materials is necessary to validate and improve the applied algorithms as well as to test the performance of devices comprising the novel materials.^[58–60] Due to the complexity of chemical reactions, molecules designed computationally might be only available with high synthetic effort and high investments in resources and time. It is therefore of high importance for the design process, to include information about the synthetic accessibility of the target compounds. There are two general procedures that enable scientists to evaluate the accessibility of chemical compounds: the forward prediction of a reaction outcome as well as the analysis of possible reaction pathways yielding the desired compound (retrosynthesis, see Figure 1d).^[61] While reaction prediction is often used if the reagents or parts of the reagents are known, retrosynthesis can be applied if a distinct target molecule is required. Unfortunately, both methods are not readily available to nonspecialists and the



Pascal Friederich, born 1988 in Bad Friedrichshall, Germany, graduated 2016 with a Ph.D. in physics from Karlsruhe Institute of Technology, where he worked on multiscale models of organic semiconductors. After research visits at Georgia Tech and Harvard University, he is now working as a postdoctoral researcher at the University of Toronto, where he is combining multiscale models with machine learning methods to design new materials for organic electronics applications.



Wolfgang Wenzel, born 1963 in Berlin, Germany, graduated 1989 with a Ph.D. in condensed matter physics from Ohio State University. After two postdoctoral appointments, he joined the Institute of Nanotechnology at Karlsruhe Institute of Technology, where he leads a group on materials simulation and design and teaches as a professor at the department of physics. His research interests center on the development of predictive multiscale simulation methods for structure formation and materials design.

check for accessibility of a chemical compound has to be done manually in collaboration with a chemical synthesis expert. This state of the art hinders in silico development of new materials in many domains and for diverse applications. Therefore, many scientists worldwide in academia^[62] and industry^[63,64] are working on solutions to provide synthetic information automatically without prior consultation of chemistry experts.

Here, we will give an overview over state-of-the-art theoretical methods and approaches that address the multiscale problem as well as on the application of such methods to the design of new materials and the optimization of devices. We will in particular discuss the accuracy requirements of methods to be used in predictive multiscale frameworks and ways to address the synthetic accessibility of predicted compounds. We will finish with a discussion of newly emerging methods for property prediction, synthesis prediction, and material design based on machine learning algorithms.

2. Available Methods

The aim of computational design of new materials can only be reached if the prediction of materials properties is accurate

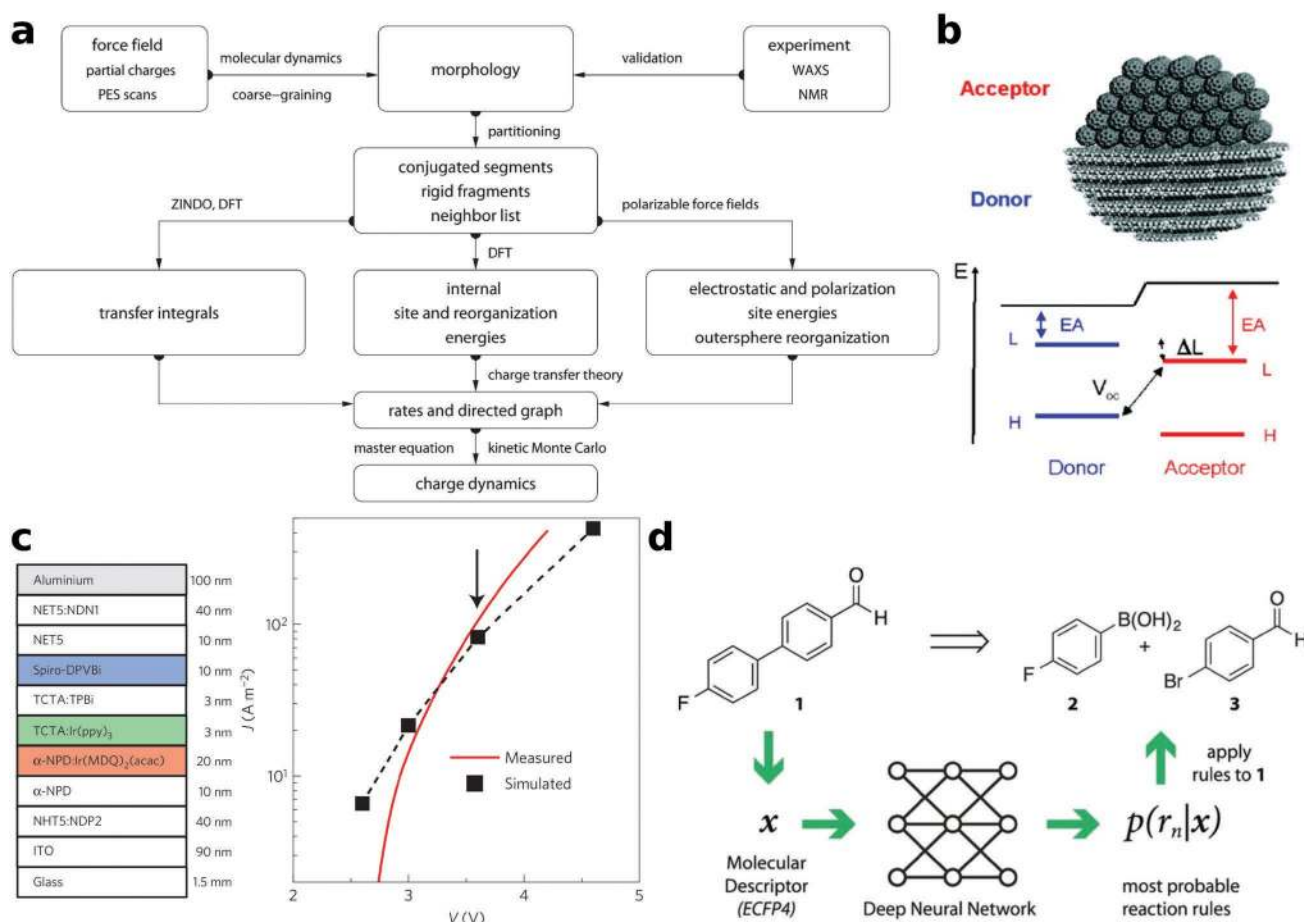


Figure 1. a) Multiscale workflow scheme for charge transport simulations. b) Models of Interface morphology and energy levels of an organic solar cell. c) Stack architecture and simulated/experimental IV characteristics of a white OLED. d) Prediction of a synthesis route (suitable retrosynthesis) using deep neural networks. a) Reproduced with permission.^[55] Copyright 2011, American Chemical Society. b) Reproduced with permission.^[45b] Copyright 2011, American Chemical Society. c) Reproduced with permission.^[132] Copyright 2013, Springer Nature. d) Reproduced with permission.^[147] Copyright 2017, Wiley-VCH.

enough not only to distinguish different classes of materials but also different closely related materials within one class. Computational models to select new materials therefore have to be fine grained enough to fully account for the effects of the chemical structure on materials properties. At the same time, many properties of materials and devices rely on phenomena on larger length scales, spanning hundreds of nanometers or, equivalently, millions of molecules. Multiscale simulation methods are one possible solution to address the challenges stemming from these requirements.^[45,53] Such multiscale methods combine models that cover different length scales, starting from a fine level (e.g., electronic structure of single molecules) to a more coarse level (e.g., mesoscopic transport phenomena). The conceptually simplest linkage between these methods and scales is parameter passing, i.e., the parameterization of the methods on larger scales by system specific input from the methods on the smaller scaled. Multiscale simulation methods to model charge transport in organic semiconductors must address the morphology of the material, the electronic properties of the molecules as well as mesoscale phenomena, such as charge transport or device operation on a coarse grained level. Methods to address these challenges will be discussed in the next sections.

2.1. Morphology

Organic electronic devices are manufactured from amorphous or crystalline organic semiconducting materials, which are either polymers or small molecules.^[65,66] Processing methods and conditions, e.g., vapor deposition or solution processing, influence materials properties and thus have to be taken into account when modeling the morphologies of organic semiconductors. Many amorphous thin films used in devices are not in equilibrium, but in a supercooled state.^[18] Various properties of these materials depend on the detailed morphology, which in turn varies with processing conditions and kinetics. Information on the morphology can be obtained from either calculations or experiments. Especially bottom-up processes such as self-assembly, phase separation of material mixtures, or crystallization of small molecules and polymers are challenging to describe in computational methods with atomistic resolution due to the large time and length scales involved.^[67,68] Small variations in the chemical structure of molecules^[39] as well as the solvents used in solution processed thin films^[69] can also alter the morphology of the material. Quantitative models thus have to be sensitive to small variations in chemical composition

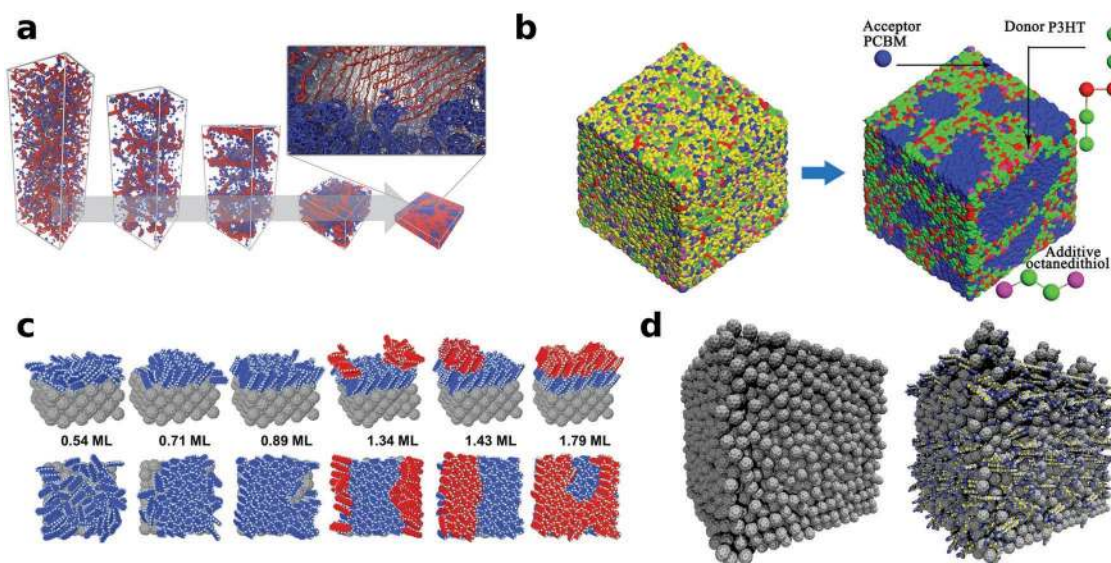


Figure 2. a) Snapshots of coarse grained molecular dynamics solvent evaporation simulation for a P3HT-PCBM mixture. Inset: Atomistic structure resulting from back-mapping. b) Dissipative particle dynamics simulation of phase separation in a P3HT-PCBM mixture. c) Molecular dynamics simulation of the deposition and self-assembly process of pentacene molecules. d) Pure and mixed amorphous morphologies of organic molecules generated by the DEPOSIT method. a) Reproduced with permission.^[81] Copyright 2015, American Chemical Society. b) Reproduced under the terms of the CC-BY 4.0 Creative Commons Attribution International License (<http://creativecommons.org/licenses/by/4.0/>).^[82] Copyright 2015, The Authors, published by Springer Nature. c) Reproduced with permission.^[46] Copyright 2011, Wiley-VCH. d) Reproduced with permission.^[88] Copyright 2015, Wiley-VCH.

as well as to collective effects of many thousand molecules on a larger time and length scale.

Accurate treatment of organic molecules in thin films and their interfaces requires accurate all-atom force fields. Experimental data to parameterize the force fields is often lacking, because density and radial distribution functions are difficult to measure and often not available for the molecules of interest. One widely pursued approach has been the development of generalized force fields that cover entire material classes. Well-known generic atomistic force fields are the OPLS family,^[70] GROMOS,^[71] and GAFF.^[72] These force fields can also be used as starting point for molecule specific parametrizations to match desired experimental quantities or ab initio data, e.g., a dihedral potential parametrized using density functional theory (DFT) calculations.^[73] A wider parameter space is offered by Class 2 force fields which include additional cross terms for the intermolecular degrees of freedom, e.g., the COMPASS force field.^[74]

In contrast to the aforementioned force fields, which fix the molecular topology throughout the simulation, bond order potentials, such as ReaxFF,^[75,76] allow dynamic bond breakage and formation. They require extensive parameterization from ab initio calculations but represent an alternative to expensive ab initio molecular dynamics simulations for modeling reactive systems.

An alternative to the conventional analytical expressions used to define force fields is the application of artificial neural networks for the calculation of intra- and interatomic potentials. Recent work includes the general ANI-1 potential for the prediction of DFT energies^[77] or a molecule specific parameterization of the internal energies for branched molecules with correlated dihedral potentials.^[78]

One way of overcoming the time and length scale limitations of fully atomistic molecular dynamics simulations is the

use of coarse graining (CG) techniques.^[79] There are several approaches to reduce the number of degrees of freedom by combining functional groups to superatoms while preserving correct distribution functions between the molecules, e.g., the parametrization of a coarse grained force field for the P3HT-polymer using iterative Boltzmann inversion (IBI) or force matching.^[67,80] For solution processed films, coarse graining techniques were applied to solvent evaporation simulations of P3HT-PCBM mixtures with subsequent backmapping to the atomistic model (see **Figure 2a**).^[81] Also for the P3HT-PCBM system, the dissipative particle dynamics method,^[82] which was originally developed for liquids, was used to model the influence of different processing parameters on the mesoscopic structure, including the blend ratio of the two components, the temperature and the solvent additive (see **Figure 2b**). However, presently available coarse grained methods struggle to describe nonequilibrium properties quantitatively. Recent developments, such as the Mori-Zwanzig dissipative particle dynamics method, aim to reproduce the equilibrium as well as kinetic properties of the underlying atomistic model. In this approach, the equations of motion of the CG beads are approximated by the Mori-Zwanzig projection and the conservative and fluctuating force components can be obtained in a bottom-up procedure from atomistic dynamics. The applicability of the method could be shown for a model system of star polymers.^[83] An alternative method incorporates kinetic information by adjusting the masses and force field parameters of the CG model within standard Langevin dynamics.^[84] An example for a general coarse grained force field is MARTINI, which is also applicable to small organic molecules.^[85,86]

Another coarse grained approach to overcome the time scale challenge and to simulate mesoscale phase separation as observed

in bulk heterojunctions of organic solar cells is the representation of the thin film in an discretized lattice model, where each voxel represents one of the materials or has a continuous concentration of a material.^[87,88] Atomistic simulations at smaller scales or constitutive equations have to be employed to parameterize the interaction between voxels as well as the time scale of diffusion and demixing. Back-mapping to atomistically resolved structures from discrete representations remains an open challenge.

In order to generate structures of vacuum processed thin films, it is necessary to model the growth process to obtain characteristic properties such as the anisotropic alignment of molecules with respect to the substrate.^[89,90] Molecular dynamics simulations of the deposition and surface rearrangement processes at the substrate surface can be used to model some aspects of the growth process, such as collective effects (see Figure 2c).^[46,47] However, realistic deposition rates and large samples of vapor deposited films are difficult to access in molecular dynamics simulations. Furthermore, molecules with several soft torsional degrees of freedom span a complex phase space which increases the sampling requirements. An efficient alternative approach is the DEPOSIT method (see Figure 2d), which directly simulates thin film deposition molecule-by-molecule.^[48] A Monte Carlo annealing algorithm samples molecule positions and orientations as well as internal degrees of freedom such as dihedral rotations of one molecule while molecules deposited earlier are constrained. It could be demonstrated that the experimentally observed anisotropy of the orientations of the molecules in vapor deposited thin films can be quantitatively predicted using the DEPOSIT method.^[49,91]

2.2. Electronic Structure and Quantum Embedding Methods

Models of amorphous organic semiconductors not only require information about the morphology but also details about the electronic structure of the resulting material. There is a large variety of methods that solve the Schrödinger equation in an approximative way to analyze the electronic structure of single molecules or periodic structures, starting from computationally inexpensive semiempirical methods such as PM7^[92] or DFTB^[93] to computationally much more demanding Post-Hartree-Fock methods^[94–96] or multi determinant methods like MRCI and CASPT2.^[97,98] DFT, one of the most widely used methods for the calculation of the electronic structure of molecular materials, balances computational cost and accuracy.^[99,100] Efficient implementations and increasing computational resources enable DFT calculations for large systems with several hundred atoms^[101] and the efficient screening of many millions of molecular compounds.^[102] For the quantitative prediction of highest occupied molecular orbital (HOMO) as well as lowest unoccupied molecular orbital (LUMO) energies of molecular materials and solids, GW has proven to yield accurate and reliable results at feasible computational costs.^[103,104] Efficient implementations of the GW method are available, e.g., in BerkeleyGW^[105] and FIESTA.^[106]

The investigation of electronic properties of amorphous thin films poses several challenges to standard electronic structure methods. First, large system sizes with $\mathcal{O}(1000)$ molecules are required to obtain sufficient statistics to quantitatively describe

disordered bulk materials. Second, the presence of localized charge carriers and excitons in typical organic semiconductors requires the accurate description of single—potentially charged or excited—molecules in an amorphous environment. To address these issues, a number of solutions were proposed which either self-consistently couple single-molecule quantum chemical calculations^[56] (see Figure 3a,b) or employ classical embedding schemes such as polarizable force fields to quantum mechanical calculations^[104,107] (see Figure 3c) in order to account for environment effects and intermolecular interaction. One of the advantages of these hybrid solutions is the flexibility in choice of the electronic structure method to describe the central molecule and the embedding. The methods can be applied to large systems and it is possible to confine charge carriers or excitons to selected molecules in a system.^[57]

For a quantitative calculation of properties such as the charge carrier mobility, highly accurate electronic structure methods are indispensable. To illustrate the severity of the problem, we note that the charge carrier mobility exponentially depends on the square of the energy disorder parameter σ , which is defined as the width of the distribution of available electronic states for holes or electrons in an amorphous material:^[108,109]

$$\mu \propto \exp(-C\beta^2\sigma^2) \quad (1)$$

where C is a percolation dependent model parameter and $\beta = 1/k_B T$ is the inverse product of temperature T and Boltzmann constant k_B . Due to this strong dependence, small errors in the calculation of the energy disorder lead to large errors in the predicted charge carrier mobility, demanding high accuracy of all models involved in the simulation workflow. We note that the accuracy of the energy disorder not only depends on the quantum mechanical method used in the electronic structure evaluation but also on the methods used to model the embedding,^[57] the morphology generation, the underlying force fields, and their parameterization (see Figure 3d about different parameters and their influence on the charge carrier mobility amorphous organic semiconductors),^[78] as well as on the consideration of dynamic effects in the model.^[110,111] Recent advances toward quantitative electronic structure methods were made using methods like GW^[104] or self-consistently tuned range-separated functionals.^[112]

Combining such approaches with accurate embedding schemes, it becomes possible to calculate solid state properties, such as thin film ionization energies or electron affinities^[17,104,113] energy disorder parameters for electrons, holes, or excited states of a pristine material or a mixture of materials, distributions of electronic couplings for charge or exciton transport^[114a] and optical film properties such as absorption spectra.^[115,116]

2.3. Mesoscale

Many properties of organic semiconductor devices are crucially influenced by mesoscale effects. Examples are percolation effects for charge transport, the formation of ordered and disordered domains in polymers, of grain boundaries in organic field effect transistors and of the morphology of bulk hetero junctions. Various methods have been proposed

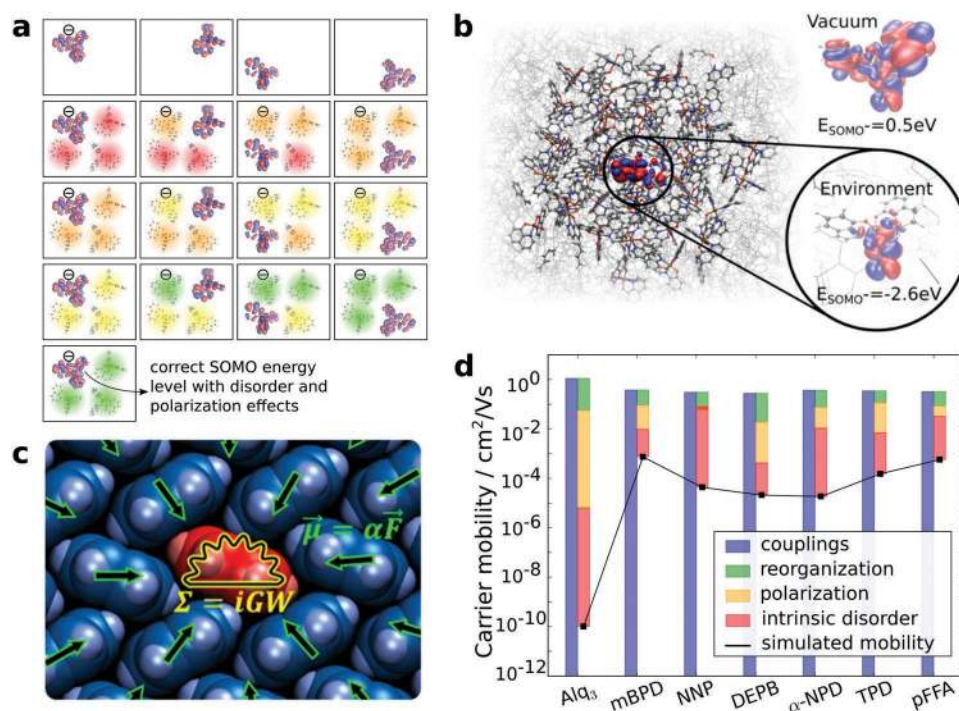


Figure 3. a) Illustration of the quantum embedding approach to compute the electronic structure of polarons and excitons in a disordered environment. b) Effect of embedding or orbital shape and energy on the polaron electronic structure. c) Illustration of an embedding scheme that couples GW calculations of a central molecule with a classical polarizable embedding method. d) Influence of electronic couplings, reorganization energy, electrostatic disorder and conformational disorder on the charge carrier mobility of amorphous organic semiconductors. a, b) Reproduced with permission.^[56] Copyright 2016, American Chemical Society. c) Reproduced with permission.^[104] Copyright 2016, American Chemical Society. d) Reproduced with permission.^[53] Copyright 2016, Wiley-VCH.

to address these issues, which often start by a discretization of the representation of the system or by its representation by continuous fields. Here we have to differentiate between amorphous systems, which are the main focus of this work and crystalline systems, which exhibit a wide range of transport regimes. In the end of this section, we will give a brief overview about models for crystalline organic semiconductors. Due to the weak intermolecular coupling in amorphous systems, charge and exciton transport in organic semiconductors is in many cases adequately described by hopping processes between molecules or polymers.^[117,118] This approach results in coupled master equations for the particles (charges, excitons), which need to be solved at the mesoscale and device level. To solve these equations, kinetic Monte Carlo (kMC) and Master Equation (ME) models have been used with various levels of approximation.^[50,117,119] The level of accuracy is highest in kMC, where all hopping and interaction processes are treated explicitly, which makes it the computationally most demanding model. In ME models, the explicit interaction between particles is simplified with a mean field model but the notion of molecular sites and their occupation probability is retained. Drift-diffusion models, which are widely used to model inorganic semiconductors and devices, provide limited predictive power of material properties because charge and exciton transport is computed using empirically parameterized models rather than modeled on a microscopic scale.^[50]

Charge transport in kMC and ME simulations is modeled as a sequence of hopping processes with phenomenological

rates, such as semiclassical Miller-Abrahams/Marcus rates, or quantum-mechanical rate expressions. Recently, the validity of these rate equation has been reevaluated: a superexchange transport mechanism between guest molecules is responsible for a significant fraction of hopping processes in mixed emitter–host systems.^[114a] This is illustrated in **Figure 4a**, which shows the distribution of transfer integrals as a function of the intermolecular distance with and without superexchange. At large distances, superexchange starts to play an important role. Furthermore, a fully quantum mechanical hopping rate equation beyond the semiclassical approximation of the electron–phonon coupling was derived. In this approach, all phonon modes are treated quantum mechanically,^[120] in contrast to the semiclassical Marcus theory. When a sufficient number of low energy phonon modes (compared to $k_B T$) is present, the full quantum mechanical rate approaches Marcus theory, while in the absence of such modes, Marcus theory significantly deviates from the quantum mechanical rates. It was shown that these effects are relevant for some small molecule organic semiconductors.

Further progress was made in elucidating differences between ME and kMC models^[119] and in addressing computational challenges of kMC methods such as parallelization,^[121] explicit long-range Coulomb interaction,^[122] and finite system sizes.^[123] In combination, these developments resulted in a large gain in speed for modern kMC implementations.^[124] The description of processes that involve multiple types of particles such as polaron–exciton interaction and their inclusion in kMC and ME models are part of current research efforts.

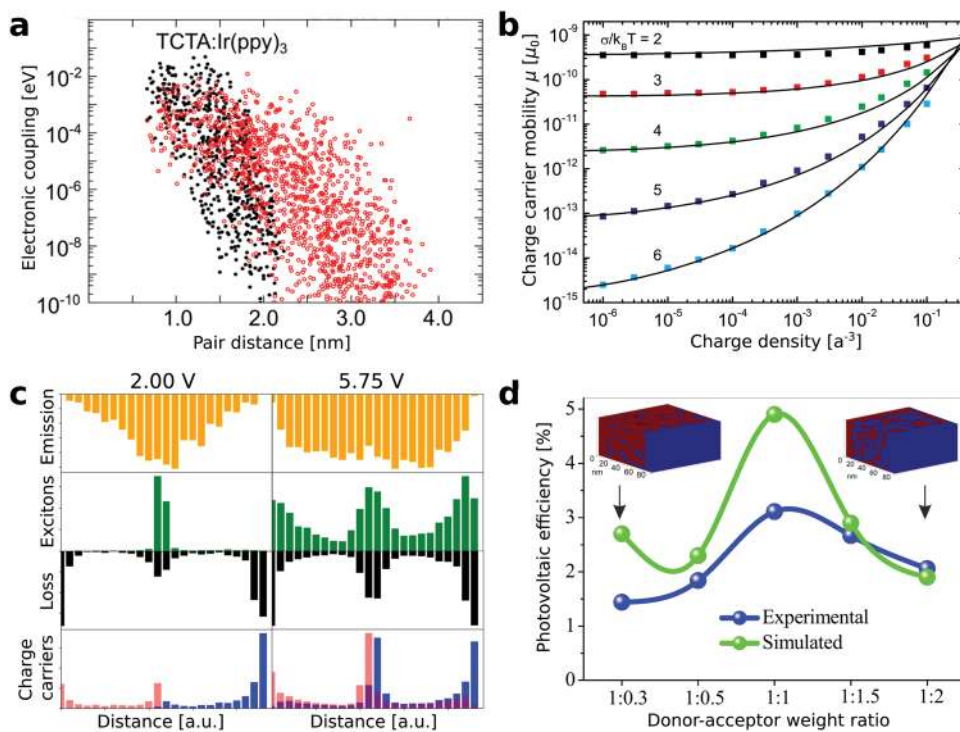


Figure 4. a) Electronic couplings between pairs of molecules as a function of the intermolecular distance with (red) and without (black) the superexchange effect. b) Charge carrier mobility as a function of the charge carrier density for different energy disorder strengths. c) Charge and exciton distributions as well as emission profiles of a bilayer OLED. d) Power conversion efficiency obtained in kMC simulations of organic solar cells. a) Reproduced with permission.^[114a] Copyright 2016, American Physical Society. b) Reproduced with permission.^[108] Copyright 2005, American Physical Society. c) Reproduced with permission.^[114b] Copyright 2018, The Society for Information Display. d) Reproduced with permission.^[87] Copyright 2005, American Chemical Society.

kMC methods are demanding in terms of computational cost and parallelization of the kMC algorithm is challenging^[121] and usually only possible in a tradeoff between speed and accuracy. With growing device size and complexity (doped/multi-component systems with realistic morphologies and explicit electrodes), numerical challenges and convergence problems arise. ME simulations are considered as a computationally less demanding alternative to kMC simulations. However, there are several limitations of ME methods: For charge-carrier concentration above 10^{-3} charges per molecule and electric field strengths smaller than 10^8 V m⁻¹, the mean-field based ME method significantly deviates from fully correlated kMC simulations due to strong polaron correlations.^[119] The origin of these correlations is the long-range Coulomb interaction between charges. A cutoff-based evaluation of the electrostatic interaction between particles as used in many kMC implementations can lead to an underestimation of the performance of organic solar cells in terms of internal quantum efficiency and current density.^[122] Excluding finite-size effects and explicitly considering long-range interactions in charged systems helps to remedy numerical challenges in kMC simulations.^[43]

Following on the pioneering work of Bäessler, kMC simulations are now routinely used to analyze the charge carrier mobility as a function of the charge carrier density, temperature and other parameters (see Figure 4b).^[108] In many simulations, the distribution of the hopping sites in the morphology is simplified to a (cubic) lattice with hopping rates fitted to

experimental data of a single-component material. Recent advances in kMC and ME models included the development of off-lattice kMC codes and their parameterization with ab initio calculations, which was shown to influence the field dependence of the charge carrier mobility.^[55,114a]

In addition to the motion of charges, exciton dynamics must be addressed in OLED and organic photovoltaic (OPV) models. The internal quantum efficiency of organic solar cells is limited by the fraction of excitons that diffuse to an interface and separate before decaying or quenching, making the exciton diffusion length an important material parameter for OPV applications. In OLEDs, the internal quantum efficiency decreases with increasing current (“efficiency roll-off”) caused by exciton–exciton and exciton–polaron quenching processes, which become relevant at high exciton and polaron densities. This is illustrated in Figure 4c, where the spatial distribution of charges, excitons, generated photons, and quenching processes is analyzed in detail on the basis of kMC calculations. Quenching processes mainly occur in regions with high polaron and exciton density and are considered as one of the potential causes of degradation of OLEDs.^[125] Excitonic processes have been incorporated into kMC models of organic materials and devices,^[126–129] which were deployed to investigate loss processes and degradation due to exciton quenching^[54,130] or the impact of dopant densities on exciton dynamics.^[131]

Recently, kMC simulations have been extended from single materials to complex multilayer device architectures. In a

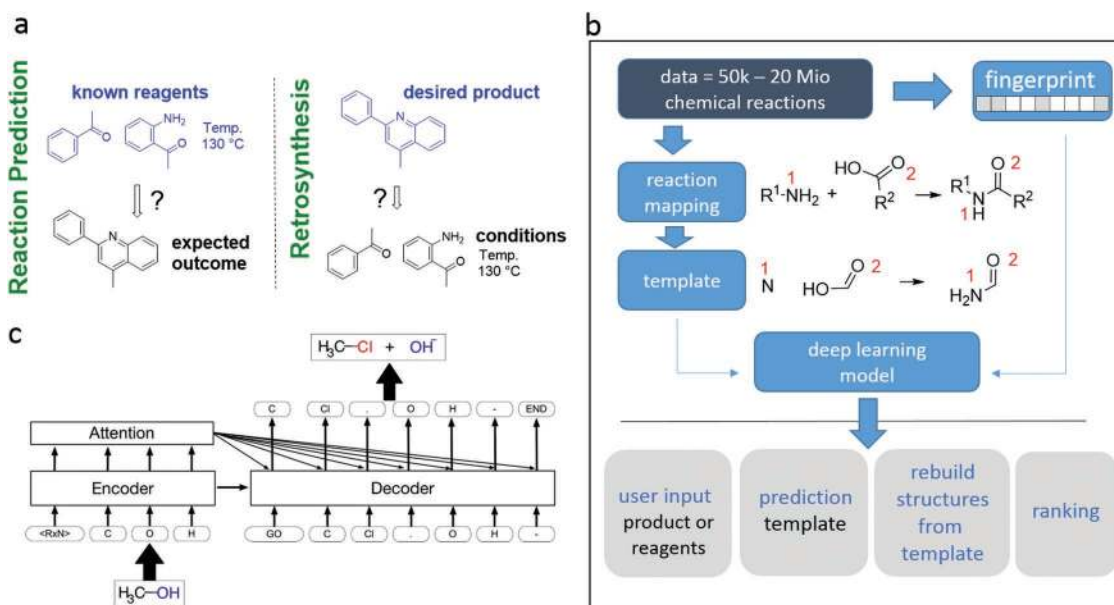


Figure 5. a) Schematic explanation of differences between forward reaction prediction and retrosynthesis as two similar approaches solved by machine learning. b) Summary of the most often seen main components of machine learning projects for supervised, template-based reaction prediction or retrosynthesis. c) Basic model of a seq2seq approach as example for an unsupervised, template-free approach to retrosynthesis. c) Reproduced with permission.^[151] Copyright 2017, American Chemical Society.

landmark study, a white OLED device with multiple emissive layers was modeled,^[132] which required the simulation of both charge and exciton dynamics.^[130,133] Figure 1c shows the device architecture of the white OLED as well as the experimental and calculated J - V characteristics of the device. In order to perform these calculations, the hopping parameters of the kMC model were chosen to fit experimental mobility.

Crystalline organic materials are also heavily investigated due to their high mobility and applications in single-crystal devices.^[134] However, they pose unique challenges that differ significantly from those of amorphous systems, which are the main focus of this review. The charge transport in organic crystals cannot be adequately described by either purely bandlike/coherent transport, nor by hopping transport. The most general approach are polaron models.^[135,136] In contrast to amorphous semiconductor materials, in high-quality crystalline organic solids, the structural (static) disorder is small and the fluctuation of intermolecular transfer integrals (dynamic disorder) remains the factor that intrinsically limits the mobility. Recently, a model based on the transient localization theory^[137] has shown good agreement with experimental data for several classes of organic crystals.^[138,139]

2.4. Synthetic Accessibility

Moving toward molecular design, the synthetic accessibility of the proposed molecules is a major challenge. Screening existing molecular databases severely restricts the chemical space that is accessible for novel applications, while proposing databases of molecules that are either impossible to make or very costly limits the impact of computational design. For these reasons, it would be very beneficial to integrate the synthetic

accessibility of novel molecules or design rules for the accessible chemical space into molecular design efforts. In the past, automatic retrosynthesis or reaction prediction (see Figure 5a) required information from databases and/or the manual encoding of chemical rules. In particular vendors of chemical information systems like the Chemical Abstracts Service (CAS), Wiley, Elsevier, and ChemicalInfo^[140] but also academic initiatives developed expert systems that still serve as instruments to evaluate possible reaction pathways. Some of the most widely used systems are ChemPlanner,^[141] PathFinder, ICSynth, LHASA, CAMEO, SOPHIA, and EROS.^[142] Rule-based chemical expert systems can be very helpful for specific chemical questions but suffer from disadvantages that limit their use in particular for virtual materials design: 1) many predictions or proposed reaction pathways are not precise enough or at least not match the decisions human experts would make, 2) due to performance reasons, the expert systems are not designed to handle high throughput prediction, 3) the underlying rules of the expert systems have to be encoded at least in parts manually which means that any logic that was missed by human beings is lacking in the information system, and 4) most of the reliable systems are only commercially available and very expensive. More recent approaches to the prediction of reactions and retrosynthesis may solve several problems that limit the systems used so far. Newer developments include the automated extraction of reaction rules,^[143] new models for chemical reasoning,^[144] heuristics aided methods,^[145] and the use of machine learning. In particular, the application of machine learning is very promising with respect to high throughput reaction prediction and retrosynthesis. The combination of modern cheminformatic data preprocessing tools with training of neural networks was shown to be very successful during the past years. Till today, two general approaches com-

pete, namely template-based and template-free approaches. Template-based approaches consist of two steps including the automatic generation of rules from the available reaction data set and the training of a neural network, which is then used for the reaction decision making via ranking of the most likely occurring reaction template. Examples for such models have been presented for reaction prediction as well as retrosynthesis (for a schematic summary see Figure 5b).^[146–148] A recently published method for reaction prediction by neural networks in combination with symbolic artificial intelligence (AI) highlights the impact of the current developments and the scope of machine learning methods: A model to retrosynthetic analysis via training of a neural network with a dataset of more than 12.5 million reactions was developed to select the most promising retrosynthesis step.^[149] Monte Carlo tree search with an expansion policy network guides in a second step the search for the most suitable chemical route. Furthermore, template-free machine learning models have been developed and successfully applied to retrosynthesis.^[150–152] These methods, consisting in most of the cases of seq2seq approaches (Figure 5c) that are well known for language translation, don't need any prior rule extraction. However, at the moment, the template-based and the template-free methods need additional improvements to serve as a reliable recommender tool that can be used without the additional approval by a chemistry expert. Key challenges of the current procedures are to improve accuracy of the most likely reaction outcome or preferred retrosynthesis and to increase reproducibility which suffers from the probabilistic components of the AI methods used. This currently complicates a transfer into a productive mode for materials design. As a result, the current developments serve as promising models but they do not provide a ready to use or easy to rebuild solution to a research problem. The major bottlenecks that have to be targeted are:

- 1) the generation of Open Access, high-quality, curated reaction databases that can be used as solid basis for all training projects;
- 2) the development of better and more transparent data curation methods for information on chemical reactions;
- 3) the development of machine learning models that include to-date unsolved/challenging topics such as to differentiate stereoisomers or to include/predict reaction conditions, and
- 4) the disclosure of source code and detailed description of investigated mechanisms to other scientists, allowing the transparency of claimed developments and the comparison of available algorithms.

For the implementation of predictive systems supporting in silico materials design, the challenges 1–4 are even more important than other claimed improvements of the current procedures which is, e.g., the missing innovation character of neural network based machine learning models.^[147] The generation of perpetuating machine learning models should be the most important aim to allow the implementation of robust systems that provide useful information for nonchemists or machines. Only very few initiatives tackle the described problems. One important work describes the parsing of patents granted by the United States Patent and Trademark Office between 1976 and 2013 and generated an open database for the described chemical

reactions.^[148,153] Currently, this work combined with some small manually curated data sets,^[154,155] builds the basis of all available, reproducible training sets for machine learning based reaction prediction and retrosynthesis projects in chemistry. Solutions to overcome the currently existing obstacles are also established at the Karlsruhe Institute of Technology, where several groups work on materials design projects that require fast and efficient identification of possible synthesis pathways. Besides the implementation and testing of published machine learning algorithms for reaction prediction and retrosynthesis, novel methods for automatic data curation and preprocessing of reaction data as well as improvements of existing models are part of two ongoing research projects. The group uses diverse synergies in particular originating from activities in cheminformatics. The development of an own infrastructure for the recording and processing of chemical reaction data via modern electronic lab journals^[156–158] fosters not only systematic digitization of research data but also the in situ generation of reaction databases that are curated by the responsible researcher. In parallel, an Open Access database was launched.^[158] This allows the retrieval of information on successfully synthesized molecules to serve as reliable source for comprehensive information on chemical reactions for the training of neural networks. The data capture activities are combined with collaborative projects on data mining, e.g., with the Beilstein Institute^[159] providing software for the automatic extraction of chemical information from Word documents or CDX-Files and its import to, e.g., an electronic lab notebook.^[158]

3. Multiscale Modeling and Material Design

3.1. Multiscale Modeling

Material properties of organic semiconductors, such as charge carrier mobility, exciton distribution, thermoelectric properties, not only depend on single molecule properties, but are influenced by the microscale arrangement of molecules and the mesoscale morphology, including phase separation, domain formation, and the formation of percolation paths. Computational models for materials properties relevant for devices thus require treatment of all relevant effects on the different length scales, but most studies presently focus on one particular aspect of the problem. To date, kMC stack simulations have been based on experimentally measured parameters. To enable device design, it is desirable to compute all of the required parameters from first principles. Materials parameters in the context of applications are typically queried at the scale of several hundreds of nanometers or beyond, but increasingly modulated by changes at the molecular, i.e., subnanometer scale. Because these changes often entail complex modifications of the material properties, which presently cannot be addressed by a single physical or chemical model, linkage of several models into robust and reproducible workflows is a prerequisite for materials design.

For these and other reasons, multiscale modeling is of increasing relevance in many scientific disciplines. These efforts include large community-wide efforts, such as the work on QM/MM integration, which was awarded the 2013 Nobel prize in chemistry,^[160,161] as well as molecular dynamics based methods for the prediction of structure formation of polymeric

and/or crystalline materials.^[68,162–164] Coarse graining is a widely used technique to overcome the length and timescale limitations of atomistic simulations.^[163] Multiscale modeling of organic electronic devices is furthermore used to analyze relations between the molecular structure of single molecules and properties of disordered materials on a mesoscopic length scale.^[43,50]

In Sections 2.1–2.3, we discussed individual simulation methods on different length scales, in the following we will focus on the combination of these simulation methods to multiscale simulation workflows, which enables the description of complex organic devices in a fully computational way. In one application, aiming at the description of the carrier mobility in small molecule organic semiconductors,^[53] single molecule electronic structure calculations are used to parameterize atomistic or molecular force fields for the simulation of morphology formation. The resulting morphology is then passed in a third step to calculate the electronic structure of the bulk material. Parameters are extracted to parameterize mesoscale models such as kMC simulations to calculate properties of mesoscale materials. One of the challenges of multiscale models is the transfer of results between different levels of complexity. Molecule specific DFT-parameterized force fields act as links between the first two steps of the multiscale workflow described above. The link from morphology generation to the evaluation of the electronic structure is the positions of all atoms of a molecular system. Energy disorder, distributions of coupling matrix elements, and reorganization energies are parameters, which are extracted from electronic structure calculations that are used as input to parameterize mesoscale simulations of charge and exciton transport or full device simulations.

Mesosopic device models typically require electronic couplings, reorganization energies and energy differences between pairs of molecules. Until recently, these parameters were fitted to experimental data, which yields insight into the function of the materials but precludes screening of new compounds. In more recent studies the kMC and ME models were fully parameterized using ab initio calculations without need for experimental or empirical input parameters.^[50,165–170] The transfer of these microscopic calculations to mesoscopic systems is nontrivial. OLED or OPV devices consist of multiple layers with layer thicknesses of up to several hundred nanometers, while morphologies created with current MM/MC methods are limited to system sizes of ≈ 10 –20 nm. To include percolation and fluctuation effects on a larger length scale, device-level simulations therefore require a coarse-grained model of the atomistic structure.^[123] A widely used method is the representation of molecular systems as spatial distributions of points, which follow the same pair distance distributions as atomistic systems. This can be achieved with iterative Boltzmann-Inversion^[50] or stochastic methods.^[171] The energy disorder of the coarse grained samples can be modeled with correlated disorder models (CDM),^[171] which fit spatial correlations in site energies to atomistic samples. Distance-dependent electronic couplings are stochastically generated using Gaussian distributions fitted to atomistic calculations.^[114a,171]

The simulation of bulk heterojunctions in organic solar cells intrinsically requires methods capable of describing inhomogeneous material mixtures. One of the first kMC simulations of such systems was reported in 2005^[87] where phase-separated

morphologies on cubic lattices were generated to calculate the dependence of the internal quantum efficiency of an organic solar cell on the domain size of the bulk heterojunction. More recently, the relation between morphology and solar cell efficiency was quantified in detail using a multiscale simulation method including kMC simulations.^[172] The efficiency of organic solar cells as function of the domain-size, donor-acceptor ratio, and thickness has been simulated with a novel multiscale method.^[173] Figure 4d, as an example, compares experimental data and kMC results on the relation between power conversion efficiency and the ratio of donor and acceptor in the active layer of the solar cell.

Master equation simulations that were fully parameterized using the multiscale workflow presented above have been performed for electron-only^[168] as well as hole-only^[170] devices and extended to include superexchange and disorder correlations.^[169,174] kMC and Master Equation methods have also been applied to organic thin films comprising mixtures of materials, e.g., doped organic semiconductors. This enabled the simulation of electron hole pair dissociation in crystalline pentacene: F4TCNQ systems^[175] as well as the analysis of the doping dependent mobility and density of states.^[176] kMC simulations also helped to gain insights into the polaron distribution in doped organic semiconductors, such as ZnPc:F₆-TCNNQ,^[177] and disordered polymers, where polaron delocalization plays an important role.^[178,179]

3.2. Materials Design

There is an increasing number of examples for computational design of organic semiconductors.^[180,181] These include crystal structure prediction,^[68] screening and design of organic crystals with high and isotropic charge carrier mobility.^[182,183] Another application addresses the emissive layers of fluorescent and phosphorescent OLEDs comprise of mixtures of emissive dye molecules and host materials with slightly larger optical gap. To prevent transfer of triplet excitons to hosts, the triplet-singlet gap of host materials has to be small. Host materials that fulfill this requirement can be systematically designed in silico.^[184] TADF-based materials led to a boost in quantum efficiency in OLEDs by efficiently harvesting triplet excitons. Design rules for small singlet triplet energy gaps and low fluorescence lifetimes help to computationally screen and design TADF emitters.^[41,185]

In the following, we illustrate the state of the art of computational design in detail for three recent, representative studies, which is in no way exhaustive. Specifically, we will discuss the optimization of the electron mobility,^[39] the band structure engineering of organic crystals^[40] and the search for promising materials for organic solar cells.^[36]

The availability of suitable predictive simulation methods increasingly enables de novo design of new organic materials with improved properties. One prototypical application of the multiscale simulation protocol described in Section 3.1 is the systematic computational improvement of the charge carrier mobility of electron transport materials for OLEDs.^[39] Here, the electron mobility of an aluminum complex (Alq₃) and various close derivatives (see **Figure 6a**) was computed using

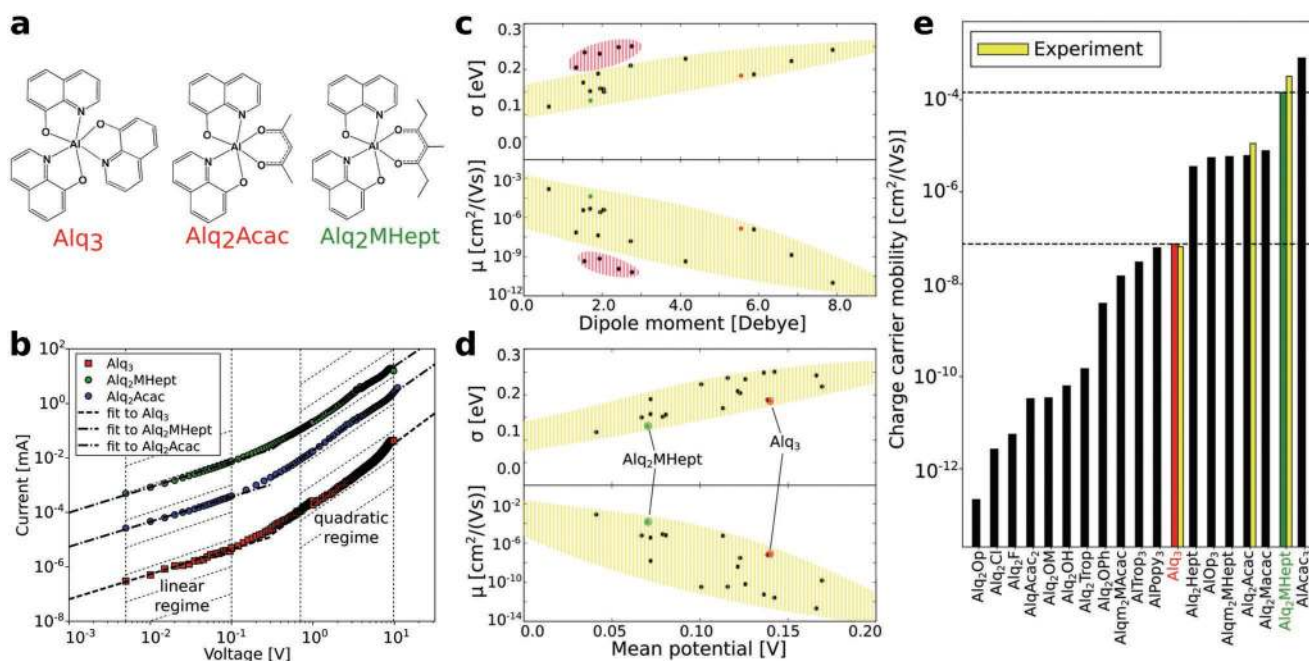


Figure 6. a) Chemical structure of Alq₃ and two derivatives. b) SCLC measurements of the electron mobility of the compounds. c,d) correlation between dipole moments and mean potentials with the energy disorder and the electron mobility of amorphous thin films. e) Predicted electron mobility of a series of Alq₃ derivatives and comparison with SCLC measurements shown in (b). a–e) Reproduced with permission.^[39] Copyright 2017, Wiley-VCH.

a combination of force field parameterization, morphology generation, electronic structure calculation, and the estimation of the electron transport level as well as the electron mobility of the amorphous materials. The authors showed not only that they can reliably predict the charge carrier mobility of different classes of materials^[53] but that they can additionally define design rules, which connect the chemical structure of single molecules to mesoscale material properties like the charge carrier mobility. One example of such a design rule is that minimization of the mean electrostatic potential in the near and far field of a single molecule reduces the energy disorder of the material and thus maximizes its charge carrier mobility (see Figure 6c,d). This enabled the authors to design a derivative of Alq₃ with low mean electrostatic potential which showed both in simulation and in experimental validation a boost in electron mobility by three orders of magnitude compared to the reference material Alq₃ (see Figure 6b,e).

A second example of successful application of multiscale models to tune materials properties by systematic molecular design was performed on ordered blends of organic semiconductors.^[40] Here, the quadrupole moment of zinc/boron (sub)phthalocyanines influences the HOMO–LUMO gap and permits band-structure engineering (see Figure 7a,b). Halogenation of the molecules reverses the sign of the quadrupole moment which allows for fine-tuning of the energy levels by systematic modification of the molecular structure of the materials. This provides additional degrees of freedom for the optimization of devices made of organic semiconductors and in particular organic solar cells containing (sub)phthalocyanines:C₆₀ blends.

A third example of computational efforts to find new materials for organic electronics applications was performed in the

context of the Clean Energy Project (CEP),^[36,186–188] which aims to systematically improve the efficiency of organic solar cells by finding new materials with better suited electronic properties. Of 2.3 million organic molecules were designed and computationally characterized using DFT calculations.^[186] Characteristic energy levels were calculated for each of the compounds (see Figure 7c). These were used to estimate the maximally possible power conversion efficiency of a solar cell with the respective material as one of its constituents (see Figure 7d). The materials properties were saved as in a publicly available database to serve as a compound database for future organic electronics applications and to be used as training data for machine learning models.

4. Outlook

The discussion of recent efforts to complement experimental materials discovery by theoretical methods demonstrates that there have been significant improvements in the theoretical methods spectrum to enable characterization and increasingly prediction of materials properties and the design of new organic semiconductors. The methods discussed so far were based on physical models of the electronic structure, the interaction between atoms, and molecules and/or the physical laws that govern charge transport in organic materials. Recent years have shown the emergence of a novel class of methods that are not based on physical models but on learning systematic relations in large datasets and on methods of statistical inference. While originally developed for applications such as speech and image recognition, there is now an intense interest in machine learning models, which are now combined with methods from

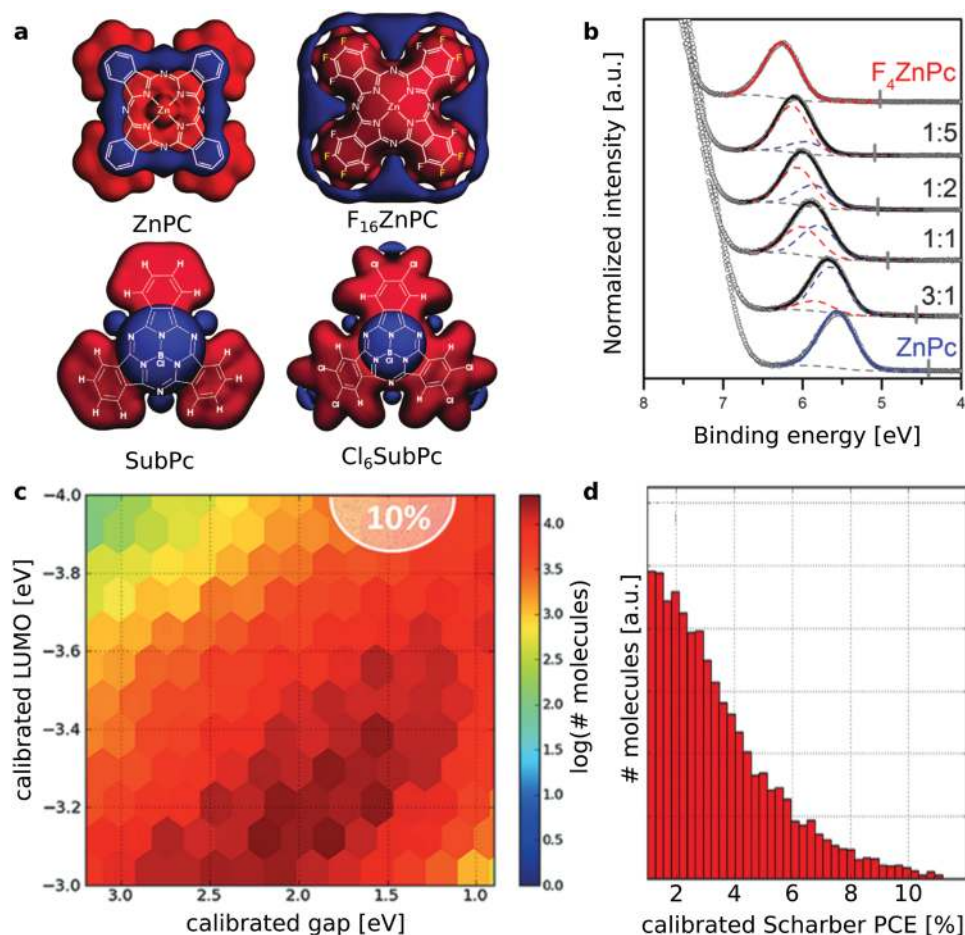


Figure 7. a–d) Derivatives of ZnPC and SubPc with modified quadrupole moments (a) allow for continuous tuning of the bandgap of organic semiconductors as shown in the LUMO and HOMO–LUMO gap distribution (b,c) and the power conversion efficiency distribution (d) for a subset of the materials screened in the Harvard Clean Energy project. The region that allows for power conversion efficiency (with respect to a PCBM acceptor according to the Scharber model) of more than 10% is highlighted in (c). a,b) Reproduced with permission.^[40] Copyright 2016, The American Association for the Advancement of Science. c,d) Reproduced with permission.^[36] Copyright 2011, American Chemical Society.

cheminformatics to be applied on chemistry and materials science.^[189–193] Examples include the use of regression and classification models such as neural networks for the prediction of molecular or materials properties^[194,195] and for synthesis planning^[146,149,150] as well as the use of generative models such as variational autoencoders and generative adversarial networks for inverse molecular design.^[192,196–198] However, the ultimate impact of these approaches on the development of new materials and devices has yet to be evaluated.

Challenges of the machine learning based approaches include the collection or generation of sufficient training data. This either requires large reliable databases^[199] of reactions as frequently used for reaction prediction,^[200] automated experiments combined with reinforcement or Bayesian inference based approaches^[201,202] or computational generation of training data by screening large databases of molecules.^[41,102,203] A second challenge is the representation of molecular structures (or more general of any material) in a machine learning approach. Recent developments include fingerprint based representations of the molecular graph^[204,205] or the 3D structure.^[206]

Current achievements in reaction prediction and retrosynthesis demonstrate the potential of machine learning to solve one of the bottlenecks of state-of-the-art materials design, which is the planning of efficient reaction routes of new molecules.^[144,149,151,152] To date, machine learning based systems already outperform traditional methods in chemistry including examination by human experts and database search in several fields of work. Examples can be given for, e.g., feature detection,^[207] bioactivity prediction,^[208] or drug target prediction,^[209] and others.^[210]

Potential benefits of a reliable and fast computational method for chemical synthesis planning are manifold: The expertise to design reactions and desired products could become a publicly available tool without need of long-lasting database research. This would be a big step towards the democratization of chemistry.^[211,212] Automated computational reaction planning can be combined with in silico materials design (be it with physical models and design rules,^[39,40] computational screening approaches,^[186] or machine learning methods^[41]) and automated in situ synthesis and characterization,^[201,213] which has the potential to open up new

pathways toward fully autonomous high-throughput materials discovery.^[190]

Acknowledgements

The authors received funding from the European Union Horizon 2020 research and innovation programme under grant agreement No 646176 (EXTMOS) and under the Marie Skłodowska-Curie grant agreement No 795206, the High Performance Computing 2 program of the Baden-Württemberg Stiftung (Project MSME), the ERA-NET project MODIGLIANI, the Helmholtz programs BIFTM/STN, and the Cooperative Research Centre (SFB) 1176 'Molecular Structuring of Soft Matter'.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

computational materials design, multiscale modelling, organic semiconductors

Received: December 22, 2018

Published online:

-
- [1] C. W. Tang, S. A. VanSlyke, *Appl. Phys. Lett.* **1987**, *51*, 913.
- [2] M. Fröbel, F. Fries, T. Schwab, S. Lenk, K. Leo, M. C. Gather, S. Reineke, *Sci. Rep.* **2018**, *8*, 9684.
- [3] D.-H. Kim, A. D'Aléo, X.-K. Chen, A. D. S. Sandanayaka, D. Yao, L. Zhao, T. Komino, E. Zaborova, G. Canard, Y. Tsuchiya, E. Choi, J. W. Wu, F. Fages, J.-L. Brédas, J.-C. Ribierre, C. Adachi, *Nat. Photonics* **2018**, *12*, 98.
- [4] K. T. Ly, R.-W. Chen-Cheng, H.-W. Lin, Y.-J. Shiau, S.-H. Liu, P.-T. Chou, C.-S. Tsao, Y.-C. Huang, Y. Chi, *Nat. Photonics* **2017**, *11*, 63.
- [5] D. Baran, N. Gasparini, A. Wadsworth, C. H. Tan, N. Wehbe, X. Song, Z. Hamid, W. Zhang, M. Neophytou, T. Kirchartz, C. J. Brabec, J. R. Durrant, I. McCulloch, *Nat. Commun.* **2018**, *9*, 2059.
- [6] M. L. Petrus, J. Schlipf, C. Li, T. P. Gujar, N. Giesbrecht, P. Müller-Buschbaum, M. Thelakkat, T. Bein, S. Hüttner, P. Docampo, *Adv. Energy Mater.* **2017**, *7*, 1700264.
- [7] C. Wang, X. Ren, C. Xu, B. Fu, R. Wang, X. Zhang, R. Li, H. Li, H. Dong, Y. Zhen, S. Lei, L. Jiang, W. Hu, *Adv. Mater.* **2018**, *30*, 1706260.
- [8] T. Tsujimura, *OLED Display Fundamentals and Applications*, John Wiley & Sons, Hoboken, NJ, USA **2017**.
- [9] T. Setzer, P. Friederich, V. Meded, W. Wenzel, C. Lennartz, A. Dreuw, *ChemPhysChem* **2018**, *19*, 2961.
- [10] U. Würfel, D. Neher, A. Spies, S. Albrecht, *Nat. Commun.* **2015**, *6*, 6951.
- [11] J. Benduhn, K. Tvingstedt, F. Piersimoni, S. Ullbrich, Y. Fan, M. Tropiano, K. A. McGarry, O. Zeika, M. K. Riede, C. J. Douglas, S. Barlow, S. R. Marder, D. Neher, D. Spoltore, K. Vandewal, *Nat. Energy* **2017**, *2*, 17053.
- [12] B. Xu, D. Bi, Y. Hua, P. Liu, M. Cheng, M. Grätzel, L. Kloo, A. Hagfeldt, L. Sun, *Energy Environ. Sci.* **2016**, *9*, 873.
- [13] T. P. Osedach, T. L. Andrew, V. Bulović, *Energy Environ. Sci.* **2013**, *6*, 711.
- [14] S. Park, G. Giri, L. Shaw, G. Pitner, J. Ha, J. H. Koo, X. Gu, J. Park, T. H. Lee, J. H. Nam, Y. Hong, Z. Bao, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5561.
- [15] B. Lüssem, M. Riede, K. Leo, *Phys. Status Solidi B* **2013**, *210*, 9.
- [16] X. Lin, B. Wegner, K. M. Lee, M. A. Fusella, F. Zhang, K. Moudgil, B. P. Rand, S. Barlow, S. R. Marder, N. Koch, A. Kahn, *Nat. Mater.* **2017**, *16*, 1209.
- [17] J. Li, I. Duchemin, O. Maria Roscioni, P. Friederich, M. Anderson, E. D. Como, G. Kociok-Köhn, W. Wenzel, C. Zannoni, D. Beljonne, X. Blase, G. D'Avino, *Mater. Horiz.* **2019**, *6*, 107.
- [18] L. Ye, H. Hu, M. Ghasemi, T. Wang, B. A. Collins, J.-H. Kim, K. Jiang, J. H. Carpenter, H. Li, Z. Li, T. McAfee, J. Zhao, X. Chen, J. L. Y. Lai, T. Ma, J.-L. Bredas, H. Yan, H. Ade, *Nat. Mater.* **2018**, *17*, 253.
- [19] N. B. Kotadiya, H. Lu, A. Mondal, Y. Ie, D. Andrienko, P. W. M. Blom, G.-J. A. H. Wetzelaer, *Nat. Mater.* **2018**, *17*, 329.
- [20] H. Uoyama, K. Goushi, K. Shizu, H. Nomura, C. Adachi, *Nature* **2012**, *492*, 234.
- [21] J. E. Anthony, A. Facchetti, M. Heeney, S. R. Marder, X. Zhan, *Adv. Mater.* **2010**, *22*, 3876.
- [22] A. Mishra, P. Bäuerle, *Angew. Chem., Int. Ed.* **2012**, *51*, 2020.
- [23] B. J. Jung, N. J. Tremblay, M.-L. Yeh, H. E. Katz, *Chem. Mater.* **2011**, *23*, 568.
- [24] J. Mei, Y. Diao, A. L. Appleton, L. Fang, Z. Bao, *J. Am. Chem. Soc.* **2013**, *135*, 6724.
- [25] Z. B. Henson, K. Müllen, G. C. Bazan, *Nat. Chem.* **2012**, *4*, 699.
- [26] A. Endo, M. Ogasawara, A. Takahashi, D. Yokoyama, Y. Kato, C. Adachi, *Adv. Mater.* **2009**, *21*, 4802.
- [27] Q. Zhang, B. Li, S. Huang, H. Nomura, H. Tanaka, C. Adachi, *Nat. Photonics* **2014**, *8*, 326.
- [28] M. J. Leil, V. A. Krylova, P. I. Djurovich, M. E. Thompson, H. Yersin, *J. Am. Chem. Soc.* **2014**, *136*, 16032.
- [29] J. Lee, H.-F. Chen, T. Batagoda, C. Coburn, P. I. Djurovich, M. E. Thompson, S. R. Forrest, *Nat. Mater.* **2015**, *15*, 92.
- [30] Y. Im, M. Kim, Y. J. Cho, J.-A. Seo, K. S. Yook, J. Y. Lee, *Chem. Mater.* **2017**, *29*, 1946.
- [31] Z. Yang, Z. Mao, Z. Xie, Y. Zhang, S. Liu, J. Zhao, J. Xu, Z. Chi, M. P. Aldred, *Chem. Soc. Rev.* **2017**, *46*, 915.
- [32] Y. Liu, C. Li, Z. Ren, S. Yan, M. R. Bryce, *Nat. Rev. Mater.* **2018**, *3*, 18020.
- [33] D. Di, A. S. Romanov, L. Yang, J. M. Richter, J. P. H. Rivett, S. Jones, T. H. Thomas, M. Abdi Jalebi, R. H. Friend, M. Linnolahti, M. Bochmann, D. Credginton, *Science* **2017**, *356*, 159.
- [34] X. Ai, E. W. Evans, S. Dong, A. J. Gillett, H. Guo, Y. Chen, T. J. H. Hele, R. H. Friend, F. Li, *Nature* **2018**, *563*, 536.
- [35] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. A. Persson, *APL Mater.* **2013**, *1*, 011002.
- [36] J. Hachmann, R. Olivares-Amaya, S. Atahan-Evrenk, C. Amador-Bedolla, R. S. Sánchez-Carrera, A. Gold-Parker, L. Vogt, A. M. Brockway, A. Aspuru-Guzik, *J. Phys. Chem. Lett.* **2011**, *2*, 2241.
- [37] N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, N. Marzari, *Nat. Nanotechnol.* **2018**, *13*, 246.
- [38] C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, *Nat. Chem.* **2012**, *4*, 83.
- [39] P. Friederich, V. Gómez, C. Sprau, V. Meded, T. Strunk, M. Jenne, A. Magri, F. Symalla, A. Colsmann, M. Ruben, W. Wenzel, *Adv. Mater.* **2017**, *29*, 1703505.
- [40] M. Schwarze, W. Tress, B. Beyer, F. Gao, R. Scholz, C. Poelking, K. Ortstein, A. A. Günther, D. Kasemann, D. Andrienko, K. Leo, *Science* **2016**, *352*, 1446.

- [41] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, A. Aspuru-Guzik, *Nat. Mater.* **2016**, *15*, 1120.
- [42] *Multiscale Modelling of Organic and Hybrid Photovoltaics*, (Eds: D. Beljonne, J. Cornil), Springer, Berlin, Germany **2014**.
- [43] D. Andrienko, in *Handbook of Materials Modeling* (Eds.: W. Andreoni, S. Yip), Springer International Publishing, Cham, Switzerland **2018**, pp. 1–12.
- [44] A. B. Walker, *Proc. IEEE* **2009**, *97*, 1587.
- [45] a) J. J. Kwiatkowski, J. Nelson, H. Li, J. L. Bredas, W. Wenzel, C. Lennartz, *Phys. Chem. Chem. Phys.* **2008**, *10*, 1852;
b) D. Beljonne, J. Cornil, L. Muccioli, C. Zannoni, J.-L. Brédas, F. Castet, *Chem. Mater.* **2011**, *23*, 591.
- [46] L. Muccioli, G. D'Avino, C. Zannoni, *Adv. Mater.* **2011**, *23*, 4532.
- [47] Y. Youn, D. Yoo, H. Song, Y. Kang, K. Yeop Kim, S. Ho Jeon, Y. Cho, K. Chae, S. Han, *J. Mater. Chem. C* **2018**, *6*, 1015.
- [48] T. Neumann, D. Danilov, C. Lennartz, W. Wenzel, *J. Comput. Chem.* **2013**, *34*, 2716.
- [49] P. Friederich, R. Coehoorn, W. Wenzel, *Chem. Mater.* **2017**, *29*, 9528.
- [50] P. Kordt, J. J. M. van der Holst, M. Al Helwi, W. Kowalsky, F. May, A. Badinski, C. Lennartz, D. Andrienko, *Adv. Funct. Mater.* **2015**, *25*, 1955.
- [51] I. Yavuz, B. N. Martin, J. Park, K. N. Houk, *J. Am. Chem. Soc.* **2015**, *137*, 2856.
- [52] I. Yavuz, S. A. Lopez, J. B. Lin, K. N. Houk, *J. Mater. Chem. C* **2016**, *4*, 11238.
- [53] P. Friederich, V. Meded, A. Poschlad, T. Neumann, V. Rodin, V. Stehr, F. Symalla, D. Danilov, G. Lüdemann, R. F. Fink, I. Kondov, F. von Wrochem, W. Wenzel, *Adv. Funct. Mater.* **2016**, *26*, 5757.
- [54] M. Mesta, H. van Eersel, R. Coehoorn, P. A. Bobbert, *Appl. Phys. Lett.* **2016**, *108*, 133301.
- [55] V. Rühle, A. Lukyanov, F. May, M. Schrader, T. Vehoff, J. Kirkpatrick, B. Baumeier, D. Andrienko, *J. Chem. Theory Comput.* **2011**, *7*, 3335.
- [56] P. Friederich, F. Symalla, V. Meded, T. Neumann, W. Wenzel, *J. Chem. Theory Comput.* **2014**, *10*, 3720.
- [57] P. Friederich, V. Meded, F. Symalla, M. Elstner, W. Wenzel, *J. Chem. Theory Comput.* **2015**, *11*, 560.
- [58] T. Moot, O. Isayev, R. W. Call, S. M. McCullough, M. Zemaitis, R. Lopez, J. F. Cahoon, A. Tropsha, *Mater. Discovery* **2016**, *6*, 9.
- [59] A. Aspuru-Guzik, K. Persson, Materials Acceleration Platform: Accelerating Advanced Energy Materials Discovery by Integrating High-Throughput Methods and Artificial Intelligence. Mission Innovation: Innovation Challenge **6**, **2018**, <http://nrs.harvard.edu/urn-3:HUL.InstRepos:35164974>, (accessed: January 2019).
- [60] K. Alberi, M. B. Nardelli, A. Zakutayev, L. Mitas, S. Curtarolo, A. Jain, M. Fornari, N. Marzari, I. Takeuchi, M. L. Green, M. G. Kanatzidis, M. F. Toney, S. Butenko, B. Meredig, S. Lany, U. Kattner, A. Davydov, E. Toberer, V. Stevanovic, A. Walsh, N.-G. Park, A. Aspuru-Guzik, D. Tabor, J. Nelson, J. Murphy, A. Setlur, J. Gregoire, H. Li, R. Xiao, A. Ludwig, L. W. Martin, A. Rappe, S.-H. Wei, J. Perkins, *J. Phys. D: Appl. Phys.* **2018**, *52*, 013001.
- [61] J. M. Goodman, in *Applied Chemoinformatics*, Wiley-VCH, Weinheim, Germany **2018**, pp. 86–105.
- [62] S. Szymkuć, E. P. Gajewska, T. Klucznik, K. Molga, P. Dittwald, M. Startek, M. Bajczyk, B. A. Grzybowski, *Angew. Chem., Int. Ed.* **2016**, *55*, 5904.
- [63] T. Klucznik, B. Mikulak-Klucznik, M. P. McCormack, H. Lima, S. Szymkuć, M. Bhowmick, K. Molga, Y. Zhou, L. Rickershauser, E. P. Gajewska, A. Toutchkine, P. Dittwald, M. P. Startek, G. J. Kirkovits, R. Roszak, A. Adamski, B. Sieredzińska, M. Mrksich, S. L. J. Trice, B. A. Grzybowski, *Chem* **2018**, *4*, 522.
- [64] “IBM Launches Free AI Tool in the Cloud for Predicting Chemical Reactions,” can be found under <http://www.ibm.comhttps://www.ibm.com/blogs/research/2018/08/neural-networks-organic-chemistry/> (accessed: October 2018).
- [65] A. Facchetti, *Chem. Mater.* **2010**, *23*, 733.
- [66] S. R. Forrest, *Nature* **2004**, *428*, 911.
- [67] D. M. Huang, R. Faller, K. Do, A. J. Moulé, *J. Chem. Theory Comput.* **2010**, *6*, 526.
- [68] J. Yang, S. De, J. E. Campbell, S. Li, M. Ceriotti, G. M. Day, *Chem. Mater.* **2018**, *30*, 4361.
- [69] C. Sprau, F. Buss, M. Wagner, D. Landerer, M. Koppitz, A. Schulz, D. Bahro, W. Schabel, P. Scharfer, A. Colsmann, *Energy Environ. Sci.* **2015**, *8*, 2744.
- [70] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- [71] C. Oostenbrink, A. Villa, A. E. Mark, W. F. V. Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656.
- [72] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
- [73] Y. Nagata, C. Lennartz, *J. Chem. Phys.* **2008**, *129*, 034709.
- [74] H. Sun, *J. Phys. Chem. B* **1998**, *102*, 7338.
- [75] A. C. T. van Duin, S. Dasgupta, F. Lorant, W. A. Goddard, *J. Phys. Chem. A* **2001**, *105*, 9396.
- [76] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, T. Verstraelen, A. Grama, A. C. T. van Duin, *npj Comput. Mater.* **2016**, *2*, 15011.
- [77] J. S. Smith, O. Isayev, A. E. Roitberg, *Chem. Sci.* **2017**, *8*, 3192.
- [78] P. Friederich, M. Konrad, T. Strunk, W. Wenzel, *Sci. Rep.* **2018**, *8*, 2559.
- [79] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, D. Andrienko, *J. Chem. Theory Comput.* **2009**, *5*, 3211.
- [80] C. Scherer, D. Andrienko, *Eur. Phys. J.: Spec. Top.* **2016**, *225*, 1441.
- [81] R. Alessandri, J. J. Uusitalo, A. H. de Vries, R. W. A. Havenith, S. J. Marrink, *J. Am. Chem. Soc.* **2017**, *139*, 3697.
- [82] C. Du, Y. Ji, J. Xue, T. Hou, J. Tang, S.-T. Lee, Y. Li, *Sci. Rep.* **2015**, *5*, 16854.
- [83] Z. Li, X. Bian, B. Caswell, G. Em Karniadakis, *Soft Matter* **2014**, *10*, 8659.
- [84] J. F. Rudzinski, T. Bereau, *Eur. Phys. J.: Spec. Top.* **2016**, *225*, 1373.
- [85] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. de Vries, *J. Phys. Chem. B* **2007**, *111*, 7812.
- [86] T. Bereau, K. Kremer, *J. Chem. Theory Comput.* **2015**, *11*, 2783.
- [87] P. K. Watkins, A. B. Walker, G. L. B. Verschoor, *Nano Lett.* **2005**, *5*, 1814.
- [88] T. Moench, P. Friederich, F. Holzmueller, B. Rutkowski, J. Benduhn, T. Strunk, C. Koerner, K. Vandewal, A. Czyrska-Filemonowicz, W. Wenzel, K. Leo, *Adv. Energy Mater.* **2016**, *6*, 1501280.
- [89] E. Ito, Y. Washizu, N. Hayashi, H. Ishii, N. Matsuie, K. Tsuboi, Y. Ouchi, Y. Harima, K. Yamashita, K. Seki, *J. Appl. Phys.* **2002**, *92*, 7306.
- [90] M. J. Jurow, C. Mayr, T. D. Schmidt, T. Lampe, P. I. Djurovich, W. Brütting, M. E. Thompson, *Nat. Mater.* **2015**, *15*, 85.
- [91] P. Friederich, V. Rodin, F. von Wrochem, W. Wenzel, *ACS Appl. Mater. Interfaces* **2018**, *10*, 1881.
- [92] J. J. P. Stewart, *J. Mol. Model.* **2013**, *19*, 1.
- [93] T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. D. Carlo, S. Suhai, *J. Phys.: Condens. Matter* **2002**, *14*, 3015.
- [94] C. Möller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.
- [95] R. J. Bartlett, M. Musiał, *Rev. Mod. Phys.* **2007**, *79*, 291.
- [96] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, W. Thiel, *J. Chem. Phys.* **2008**, *128*, 134110.
- [97] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, W. Thiel, *J. Chem. Phys.* **2008**, *129*, 104103.

- [98] K. Andersson, P. Malmqvist, B. O. Roos, *J. Chem. Phys.* **1992**, *96*, 1218.
- [99] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, K. A. Lyssenko, *Science* **2017**, *355*, 49.
- [100] S. Hammes-Schiffer, *Science* **2017**, *355*, 28.
- [101] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, J.-F. Méhaut, *C. R. Méc.* **2011**, *339*, 149.
- [102] M. Nakata, T. Shimazaki, *J. Chem. Inf. Model.* **2017**, *57*, 1300.
- [103] C. Faber, P. Boulanger, C. Attaccalite, I. Duchemin, X. Blase, *Philos. Trans. R. Soc., A* **2014**, *372*, 20130271.
- [104] J. Li, G. D'Avino, I. Duchemin, D. Beljonne, X. Blase, *J. Phys. Chem. Lett.* **2016**, *7*, 2814.
- [105] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, S. G. Louie, *Comput. Phys. Commun.* **2012**, *183*, 1269.
- [106] X. Blase, C. Attaccalite, V. Olevano, *Phys. Rev. B* **2011**, *83*, 115103.
- [107] P. T. van Duijnen, M. Swart, *J. Phys. Chem. A* **1998**, *102*, 2399.
- [108] W. F. Pasveer, *Phys. Rev. Lett.* **2005**, *94*, 206601.
- [109] V. Rodin, F. Symalla, V. Meded, P. Friederich, D. Danilov, A. Poschlad, G. Nelles, F. von Wrochem, W. Wenzel, *Phys. Rev. B* **2015**, *91*, 155203.
- [110] N. R. Tummala, Z. Zheng, S. G. Aziz, V. Coropceanu, J.-L. Brédas, *J. Phys. Chem. Lett.* **2015**, *6*, 3657.
- [111] P. de Silva, T. Van Voorhis, *J. Phys. Chem. Lett.* **2018**, *9*, 1329.
- [112] H. Sun, S. Ryno, C. Zhong, M. K. Ravva, Z. Sun, T. Körzdörfer, J.-L. Brédas, *J. Chem. Theory Comput.* **2016**, *12*, 2906.
- [113] T. Schneider, F. Limberg, K. Yao, A. Armin, N. Jürgensen, J. Czolk, B. Ebenhoch, P. Friederich, W. Wenzel, J. Behrends, H. Krüger, A. Colsmann, *J. Mater. Chem. C* **2017**, *5*, 770.
- [114] a) F. Symalla, P. Friederich, A. Massé, V. Meded, R. Coehoorn, P. Bobbert, W. Wenzel, *Phys. Rev. Lett.* **2016**, *117*, 276803; b) F. Symalla, P. Friederich, S. Kaiser, T. Strunk, T. Neumann, W. Wenzel, *Dig. Tech. Pap. – Soc. Inf. Disp. Int. Symp.* **2018**, *49*, 340; <https://doi.org/10.1002/sdtp.12556>.
- [115] M. Sulpizi, P. Carloni, J. Hutter, U. Rothlisberger, *Phys. Chem. Chem. Phys.* **2003**, *5*, 4798.
- [116] C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, B. Mennucci, *J. Chem. Theory Comput.* **2009**, *5*, 1838.
- [117] H. Bässler, *Phys. Status Solidi B* **1981**, *107*, 9.
- [118] H. Bässler, *Phys. Status Solidi* **1993**, *175*, 15.
- [119] F. Liu, H. van Eersel, B. Xu, J. G. E. Wilbers, M. P. de Jong, W. G. van der Wiel, P. A. Bobbert, R. Coehoorn, *Phys. Rev. B* **2017**, *96*, 205203.
- [120] X. de Vries, P. Friederich, W. Wenzel, R. Coehoorn, P. A. Bobbert, *Phys. Rev. B* **2018**, *97*, 075203.
- [121] N. J. van der Kaap, L. J. A. Koster, *J. Comput. Phys.* **2016**, *307*, 321.
- [122] M. Casalegno, G. Raos, R. Po, *J. Chem. Phys.* **2010**, *132*, 094705.
- [123] A. Massé, R. Coehoorn, P. A. Bobbert, *Phys. Rev. Lett.* **2014**, *113*, 116604.
- [124] T. Brereton, O. Stenzel, B. Baumeier, D. Andrienko, V. Schmidt, D. Kroese, *Methodol. Comput. Appl. Probab.* **2014**, *16*, 465.
- [125] S. Schmidbauer, A. Hohenleutner, B. König, *Adv. Mater.* **2013**, *25*, 2114.
- [126] S. T. Hoffmann, S. Athanasopoulos, D. Beljonne, H. Bässler, A. Köhler, *J. Phys. Chem. C* **2012**, *116*, 16371.
- [127] S. Athanasopoulos, S. T. Hoffmann, H. Bässler, A. Köhler, D. Beljonne, *J. Phys. Chem. Lett.* **2013**, *4*, 1694.
- [128] S. M. Menke, R. J. Holmes, *J. Mater. Chem. C* **2016**, *4*, 3437.
- [129] I. Maqsood, L. D. Cundy, M. Biesecker, J.-H. Kimn, D. Johnson, R. Williams, V. BommiSETTY, *J. Phys. Chem. C* **2013**, *117*, 21086.
- [130] R. Coehoorn, H. van Eersel, P. Bobbert, R. Janssen, *Adv. Funct. Mater.* **2015**, *25*, 2024.
- [131] W. Kaiser, J. Popp, M. Rinderle, T. Albes, A. Gagliardi, W. Kaiser, J. Popp, M. Rinderle, T. Albes, A. Gagliardi, *Algorithms* **2018**, *11*, 37.
- [132] M. Mesta, M. Carvelli, R. J. de Vries, H. van Eersel, J. J. M. van der Holst, M. Schober, M. Furno, B. Lüssem, K. Leo, P. Loebel, R. Coehoorn, P. A. Bobbert, *Nat. Mater.* **2013**, *12*, 652.
- [133] H. van Eersel, P. A. Bobbert, R. A. J. Janssen, R. Coehoorn, *Appl. Phys. Lett.* **2014**, *105*, 143303.
- [134] H. Dong, X. Fu, J. Liu, Z. Wang, W. Hu, *Adv. Mater.* **2013**, *25*, 6158.
- [135] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, J.-L. Brédas, *Chem. Rev.* **2007**, *107*, 926.
- [136] A. Troisi, *Chem. Soc. Rev.* **2011**, *40*, 2347.
- [137] S. Fratini, D. Mayou, S. Ciuchi, *Adv. Funct. Mater.* **2016**, *26*, 2292.
- [138] S. Illig, A. S. Eggeman, A. Troisi, L. Jiang, C. Warwick, M. Nikolka, G. Schweicher, S. G. Yeates, Y. Henri Geerts, J. E. Anthony, H. Siringhaus, *Nat. Commun.* **2016**, *7*, 10736.
- [139] S. Fratini, S. Ciuchi, D. Mayou, G. T. de Laissardière, A. Troisi, *Nat. Mater.* **2017**, *16*, 998.
- [140] *Pathfinder*, <https://chemicalinfo.com/pathfinder/> (accessed: September 2018).
- [141] "ChemPlanner" <https://www.cas.org/products/scifinder-n/chemplanner> (accessed: September 2018).
- [142] F. Feng, L. Lai, J. Pei, *Front. Chem.* **2018**, *6*, 199.
- [143] C. D. Christ, M. Zentgraf, J. M. Kriegl, *J. Chem. Inf. Model.* **2012**, *52*, 1745.
- [144] M. H. S. Segler, M. P. Waller, *Chem. - Eur. J.* **2017**, *23*, 6118.
- [145] D. Rappoport, C. J. Galvin, D. Y. Zubarev, A. Aspuru-Guzik, *J. Chem. Theory Comput.* **2014**, *10*, 897.
- [146] J. N. Wei, D. Duvenaud, A. Aspuru-Guzik, *ACS Cent. Sci.* **2016**, *2*, 725.
- [147] M. H. S. Segler, M. P. Waller, *Chem. Eur. J.* **2017**, *23*, 5966.
- [148] C. W. Coley, R. Barzilay, T. S. Jaakkola, W. H. Green, K. F. Jensen, *ACS Cent. Sci.* **2017**, *3*, 434.
- [149] M. H. S. Segler, M. Preuss, M. P. Waller, *Nature* **2018**, *555*, 604.
- [150] P. Schwaller, T. Gaudin, D. Lányi, C. Bekas, T. Laino, *Chem. Sci.* **2018**, *9*, 6091.
- [151] B. Liu, B. Ramsundar, P. Kawthekar, J. Shi, J. Gomes, Q. L. Nguyen, S. Ho, J. Sloane, P. Wender, V. Pande, *ACS Cent. Sci.* **2017**, *3*, 1103.
- [152] I. Sutskever, O. Vinyals, Q. V. Le, *arXiv:1409.3215 [cs]* **2014**.
- [153] Daniel Lowe - Patent Reaction Extraction, <https://bitbucket.org/dan2097/patent-reaction-extraction/downloads/> (accessed: September 2018).
- [154] InfoChem - test set, http://www.infochem.de/downloads/testset_classify.shtml (accessed: September 2018).
- [155] N. Schneider, N. Stiefl, G. A. Landrum, *J. Chem. Inf. Model.* **2016**, *56*, 2336.
- [156] P. Tremouilhac, A. Nguyen, Y.-C. Huang, S. Kotov, D. S. Lütjohann, F. Hübsch, N. Jung, S. Bräse, *J. Cheminf.* **2017**, *9*, 54.
- [157] S. Kotov, P. Tremouilhac, N. Jung, S. Bräse, *J. Cheminf.* **2018**, *10*, 38.
- [158] Chemotion, <https://chemotion.net/> (accessed: September 2018).
- [159] Beilstein-Institut zur Förderung der Chemischen Wissenschaften, <http://www.beilstein-institut.de/en/home> (accessed: September 2018).
- [160] R. Van Noorden, *Nat. News* **2013**, <https://doi.org/10.1038/nature.2013.13903>.
- [161] R. Van Noorden, *Nat. News* **2013**, *502*, 280.
- [162] G. Maurel, F. Goujon, B. Schnell, P. Malfreyt, *RSC Adv.* **2015**, *5*, 14065.
- [163] A. Karatrantos, N. Clarke, M. Kröger, *Polym. Rev.* **2016**, *56*, 385.
- [164] V. Dantanarayana, J. Fuzell, D. Nai, I. E. Jacobs, H. Yan, R. Faller, D. Larsen, A. J. Moule, *J. Phys. Chem. C* **2018**, *122*, 7020.
- [165] J. Nelson, J. J. Kwiatkowski, J. Kirkpatrick, J. M. Frost, *Acc. Chem. Res.* **2009**, *42*, 1768.
- [166] D. Kipp, V. Ganesan, *J. Appl. Phys.* **2013**, *113*, 234502.
- [167] H. Li, Y. Qiu, L. Duan, *Org. Electron.* **2016**, *33*, 164.
- [168] A. Massé, P. Friederich, F. Symalla, F. Liu, R. Nitsche, R. Coehoorn, W. Wenzel, P. A. Bobbert, *Phys. Rev. B* **2016**, *93*, 195209.
- [169] A. Massé, P. Friederich, F. Symalla, F. Liu, V. Meded, R. Coehoorn, W. Wenzel, P. A. Bobbert, *Phys. Rev. B* **2017**, *95*, 115204.
- [170] F. Liu, A. Massé, P. Friederich, F. Symalla, R. Nitsche, W. Wenzel, R. Coehoorn, P. A. Bobbert, *Appl. Phys. Lett.* **2016**, *109*, 243301.

- [171] B. Baumeier, O. Stenzel, C. Poelking, D. Andrienko, V. Schmidt, *Phys. Rev. B* **2012**, *86*, 184202.
- [172] B. Lei, Y. Yao, A. Kumar, Y. Yang, V. Ozolins, *J. Appl. Phys.* **2008**, *104*, 024504.
- [173] U. Neupane, B. Bahrami, M. Biesecker, M. F. Baroughi, Q. Qiao, *Nano Energy* **2017**, *35*, 128.
- [174] T. Neumann, J. Liu, T. Wächter, P. Friederich, F. Symalla, A. Welle, V. Mugnaini, V. Meded, M. Zharnikov, C. Wöll, W. Wenzel, *ACS Nano* **2016**, *10*, 7085.
- [175] A. Mityashin, Y. Olivier, T. Van Regemorter, C. Rolin, S. Verlaak, N. G. Martinelli, D. Beljonne, J. Cornil, J. Genoe, P. Heremans, *Adv. Mater.* **2012**, *24*, 1535.
- [176] G. Zuo, H. Abdalla, M. Kemerink, *Phys. Rev. B* **2016**, *93*, 235203.
- [177] M. L. Tietze, J. Benduhn, P. Pahner, B. Nell, M. Schwarze, H. Kleemann, M. Krammer, K. Zojer, K. Vandewal, K. Leo, *Nat. Commun.* **2018**, *9*, 1182.
- [178] N. Vukmirović, L.-W. Wang, *Nano Lett.* **2009**, *9*, 3996.
- [179] M. Mladenović, N. Vukmirović, *Adv. Funct. Mater.* **2015**, *25*, 1915.
- [180] T. Bereau, D. Andrienko, K. Kremer, *APL Mater.* **2016**, *4*, 053101.
- [181] J.-L. Brédas, K. Persson, R. Seshadri, *Chem. Mater.* **2017**, *29*, 2399.
- [182] K. Takimiya, M. Nakano, H. Sugino, I. Osaka, *Synth. Met.* **2016**, *217*, 68.
- [183] Y. Zhu, X. Xu, X. Zhang, Y. He, X. Zeng, I. Murtaza, H. Meng, *Org. Electron.* **2018**, *61*, 87.
- [184] B. Milián-Medina, J. Gierschner, *Org. Electron.* **2012**, *13*, 985.
- [185] Q. Zhang, H. Kuwabara, W. J. Potscavage, S. Huang, Y. Hatae, T. Shibata, C. Adachi, *J. Am. Chem. Soc.* **2014**, *136*, 18070.
- [186] J. Hachmann, R. Olivares-Amaya, A. Jinich, A. L. Appleton, M. A. Blood-Forsythe, L. R. Seress, C. Román-Salgado, K. Treppe, S. Atahan-Evrenk, S. Er, S. Shrestha, R. Mondal, A. Sokolov, Z. Bao, A. Aspuru-Guzik, *Energy Environ. Sci.* **2014**, *7*, 698.
- [187] S. A. Lopez, B. Sanchez-Lengeling, J. de Goes Soares, A. Aspuru-Guzik, *Joule* **2017**, *1*, 857.
- [188] S. A. Lopez, E. O. Pyzer-Knapp, G. N. Simm, T. Lutzow, K. Li, L. R. Seress, J. Hachmann, A. Aspuru-Guzik, *Sci. Data* **2016**, *3*, 160086.
- [189] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, *New J. Phys.* **2013**, *15*, 095003.
- [190] D. P. Tabor, L. M. Roch, S. K. Saikin, C. Kreisbeck, D. Sheberla, J. H. Montoya, S. Dwaraknath, M. Aykol, C. Ortiz, H. Tribukait, C. Amador-Bedolla, C. J. Brabec, B. Maruyama, K. A. Persson, A. Aspuru-Guzik, *Nat. Rev. Mater.* **2018**, *3*, 5.
- [191] M. Rupp, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, *Phys. Rev. Lett.* **2012**, *108*, 058301.
- [192] B. Sanchez-Lengeling, A. Aspuru-Guzik, *Science* **2018**, *361*, 360.
- [193] R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *J. Chem. Theory Comput.* **2015**, *11*, 2087.
- [194] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, O. A. von Lilienfeld, *J. Chem. Theory Comput.* **2017**, *13*, 5255.
- [195] B. Sanchez-Lengeling, L. M. Roch, J. D. Perea, S. Langner, C. J. Brabec, A. Aspuru-Guzik, *Adv. Theory Simul.* **2018**, *2*, 1800069.
- [196] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, A. Aspuru-Guzik, *ACS Cent. Sci.* **2018**, *4*, 268.
- [197] B. Sanchez-Lengeling, C. Outeiral, G. L. Guimaraes, A. Aspuru-Guzik, ChemRxiv.org 2017, 5309668, <https://doi.org/10.26434/chemrxiv.5309668.v3>.
- [198] G. L. Guimaraes, B. Sanchez-Lengeling, C. Outeiral, P. L. C. Farias, A. Aspuru-Guzik, *arXiv:1705.10843 [cs, stat]* **2017**.
- [199] J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad, R. G. Coleman, *J. Chem. Inf. Model.* **2012**, *52*, 1757.
- [200] Reaxys, <https://www.reaxys.com/#/search/synthesis-planner> (accessed: September 2018).
- [201] L. M. Roch, F. Häse, C. Kreisbeck, T. Tamayo-Mendoza, L. P. E. Yunker, J. E. Hein, A. Aspuru-Guzik, *Sci. Rob.* **2018**, *3*, eaat5559.
- [202] F. Häse, L. M. Roch, C. Kreisbeck, A. Aspuru-Guzik, *arXiv:1801.01469 [physics, stat]* **2018**.
- [203] R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *Sci. Data* **2014**, *1*, 140022.
- [204] D. Rogers, M. Hahn, *J. Chem. Inf. Model.* **2010**, *50*, 742.
- [205] D. Duvenaud, D. Maclaurin, J. Aguilera-Iparraguirre, R. Gómez-Bombarelli, T. Hirzel, A. Aspuru-Guzik, R. P. Adams, *arXiv:1509.09292 [cs, stat]* **2015**.
- [206] S. D. Axen, X.-P. Huang, E. L. Cáceres, L. Gendele, B. L. Roth, M. J. Keiser, *J. Med. Chem.* **2017**, *60*, 7393.
- [207] P. Hop, B. Allgood, J. Yu, *Mol. Pharmaceutics* **2018**, *15*, 4371.
- [208] E. B. Lenselink, N. ten Dijke, B. Bongers, G. Papadatos, H. W. T. van Vlijmen, W. Kowalczyk, A. P. IJzerman, G. J. P. van Westen, *J. Cheminform.* **2017**, *9*, 45.
- [209] A. Mayr, G. Klambauer, T. Unterthiner, M. Steijaert, J. K. Wegner, H. Ceulemans, D.-A. Clevert, S. Hochreiter, *Chem. Sci.* **2018**, *9*, 5441.
- [210] G. B. Goh, N. O. Hodas, A. Vishnu, *J. Comput. Chem.* **2017**, *38*, 1291.
- [211] J. W. Lehmann, D. J. Blair, M. D. Burke, *Nat. Rev. Chem.* **2018**, *2*, 0115.
- [212] M. Trobe, M. D. Burke, *Angew. Chem., Int. Ed.* **2018**, *57*, 4192.
- [213] A. B. Henson, P. S. Gromski, L. Cronin, *ACS Cent. Sci.* **2018**, *4*, 793.