ANNUAL REVIEWS **Further**

Click here to view this article's online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

Multiscale Materials Modeling in an Industrial Environment

Horst Weiß, Peter Deglmann, Pieter J. in 't Veld, Murat Cetinkaya, and Eduard Schreiner

BASF SE – Materials and Systems Research, Materials Modeling Group, 67056 Ludwigshafen, Germany; email: horst.weiss@basf.com, peter.deglmann@basf.com, pieter.intveld@basf.com, murat.cetinkaya@basf.com, eduard.schreiner@basf.com

Annu. Rev. Chem. Biomol. Eng. 2016. 7:65–86

First published online as a Review in Advance on February 24, 2016

The *Annual Review of Chemical and Biomolecular Engineering* is online at chembioeng.annualreviews.org

This article's doi: 10.1146/annurev-chembioeng-080615-033615

Copyright © 2016 by Annual Reviews. All rights reserved

Keywords

predictive modeling, dissipative particle dynamics, soft matter modeling, formulation polymers, COSMO-RS, Materials Genome Initiative

Abstract

In this review, we sketch the materials modeling process in industry. We show that predictive and fast modeling is a prerequisite for successful participation in research and development processes in the chemical industry. Stable and highly automated workflows suitable for handling complex systems are a must. In particular, we review approaches to build and parameterize soft matter systems. By satisfying these prerequisites, efficiency for the development of new materials can be significantly improved, as exemplified here for formulation polymer development. This is in fact in line with recent Materials Genome Initiative efforts sponsored by the US government. Valuable contributions to product development are possible today by combining existing modeling techniques in an intelligent fashion, provided modeling and experiment work hand in hand.

INTRODUCTION

What Are the Tasks and Challenges of Modeling in Industry?

Materials modeling in industry is a very versatile discipline. Depending on the industry, materials modeling is applied in areas such as properties of organic and inorganic materials (e.g., solid-state properties, melt properties during processing, chemical stability), formulations (e.g., pharmaceuticals, home care, personal care, coatings of all kinds), energy (e.g., batteries, solar cells), manufacturing (e.g., process modeling), and transport/mobility (automotive, fuel additives). In the chemical industry, materials modeling is involved in virtually all of those application areas. It deals with different types of polymers (functional polymers as well as structural materials) but also with, for example, surfactants (e.g., 1), dispersants (e.g., 2), colloids of different sorts [e.g., cement slurries (3), polymer colloids (4)], composites [e.g., thermal conductivity (5)], materials related to energy storage [e.g., advanced battery concepts (6, 7)], and electronics [e.g., organic electronics (8)].

Predictive power and cost efficiency are the most crucial issues related to modeling in industrial applications. Modeling is supposed to deliver insight and provide experimentalists with information that would not be as easily available otherwise. Thus, the experimental results could be interpreted and later predicted. As its main benefits, industry expects that modeling

- helps to elucidate the correlation between chemical and physical properties (i.e., to bridge chemistry to application performance);
- helps to save cost and time by avoiding trial-and-error experimentation;
- builds up formalized knowledge in terms of usable models, thus avoiding repetition of mistakes;
- supports screening of properties of new materials;
- avoids unnecessary handling of expensive or dangerous materials;
- helps with production scale up; and
- all in all, reduces time to market.

It is in fact our experience that modeling can be used to predict the properties of materials that have not yet been developed/produced. For example, the modeling-guided design of formulation polymers is discussed later on.

Aside from materials development, modeling can support process optimization, although this is not the focus of this article. In particular, polymer reaction engineering can support the design of defined polymer architectures for a given application if the relevant kinetic model is at hand. Some of the required input parameters, like copolymerization parameters, can also be calculated via molecular modeling (9).

However, currently only a limited number of products and systems are completely understood. New materials are often developed not based on rational considerations but rather by trial and error. Why is that?

An obvious reason is that some modeling technologies are still the subject of ongoing research. Dramatic improvement with respect to either calculation accuracy or computational cost is needed to be able to predict properties of real-world systems. As an example, one may think of the required effort to develop an accurate force field and the cost of large-scale atomistic simulations based on these force fields.

The second reason, which is more severe, is that a mechanistic or materials design approach requires a framework (competencies, workforce, and mindset) that is rarely found in either industry or academia. As an example of this mindset, modelers, physicists, and engineers likely are not part of the team in the first (and crucial) phase of the project. Even analytics will be reduced to a

minimum, because to the product developer, the route to solution seems very clear, i.e., assuming that the target will be reached with only a few experiments. If something does not work as expected according to this simplistic research design, the usual approach is to go back to the lab and to hope that with chemical intuition, some analytics, and a few new experiments, the hurdles will be overcome. But what if nothing helps? Then there are two options. Most likely, the idea is abandoned owing to financial and time constraints. The second option is to finally involve physics and modeling to obtain a better understanding. At this stage of the project it is—at least in an industrial environment—often too late to be successful, because resources are gone, and new ideas have come up.

The third reason is that product recipes are, after a long history of optimization steps, very complex; thus, in many applications, targets of basic research are only precursors for ingredients of final market products. Owing to time and cost constraints, the overall performance of the complex mixture/formulation/system is optimized within a narrow search space, because fundamental mechanistic understanding will not immediately lead to improved products. One consequence is that development of new products usually is an optimization of existing complex recipes, which typically even further increases their complexity. Detailed analysis of the ingredients of recipes is often missing, and chemical structures are not always known in materials science. Thus, it is, for instance, hardly possible to understand the correlation between emulsion polymerization conditions in which a latex is formed and the performance of this latex in a paint with respect to, e.g., pigment leaching stability. In such systems, many fundamental aspects are not yet fully understood (e.g., the interaction of a polymer latex with a solid pigment like rutile).

The lack of detailed understanding of a product makes it difficult for rational approaches to step in, and consequently, most of the existing products and processes have been developed empirically. The consequence is that modelers often have no experience with the systems they are all of a sudden supposed to improve using a rational approach. This lack of system expertise on the side of the modeler is very critical, because it determines the acceptance of the modeler by his experimental colleagues.

Can We Do Better? Yes, We Can!

Despite their decades-long persistence, the challenges outlined above have only recently come into the focus of think tanks and politics. The most press-covered political event is the Materials Genome Initiative, started in 2011 by the US government with the aim of doubling the speed and reducing the cost of discovering, developing, and deploying new advanced materials (10, 11). One of the key aspects addressed was to avoid trial-and-error experimentation and to transfer experiments from lab benches to computers. It was also made clear that the route to solution is composed of more aspects than just avoiding experiments. It should consist of an integrated framework for research providing accurate and fast tools for computer modeling; data management and analysis; a trained, open-minded workforce; and multidisciplinary teams. Parts of this approach are already implemented in industry. In this article, we review and discuss the demands and the present status, mainly within BASF, of this new, rational, multidisciplinary approach within materials research. Even more specifically, because materials modeling is extremely diverse, the discussion is focused on polymer solutions for formulations.

Combining scales and soft matter modeling toolbox. Molecular modeling involves a mathematical description of systems in chemical detail, usually resulting in a set of equations describing the situation of interest. The goal is to solve these complex equations, such as the Schrödinger equation for electronic and nuclear motion or Newton's equations of motion, for interactions in

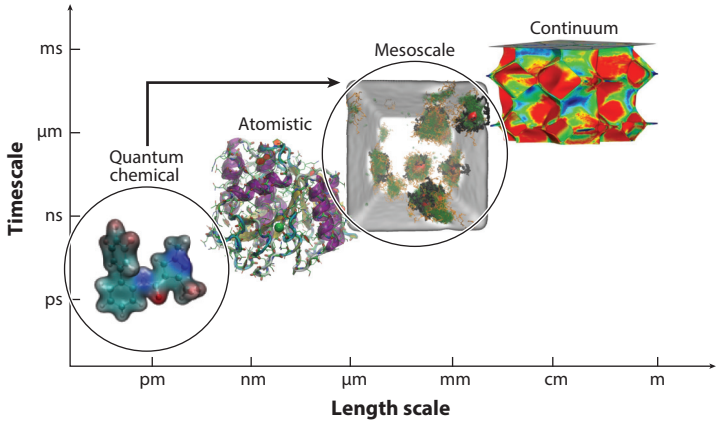

Figure 1

Time- and length scales relevant for materials and process development. The focus of this review is linking the mesoscale to specific chemistries by employing quantum chemistry for a parameterization of the mesoscale simulation method.

many particle systems, which describe natural phenomena in detail. In a practical application of molecular modeling, mathematical equations or algorithms are devised to quantitatively describe the physical and chemical phenomena (e.g., energy states, structures, reactivity, positions and moments of atoms, or order parameters) of a particular system. Molecular modeling can be used to describe a diversity of chemical systems with a wide range of complexity. Applications range from the high-accuracy prediction of spectra or thermodynamic properties of small molecules in the gas phase to polymers, which consist of hundreds to millions of atoms; the timescales of interest range from picoseconds, for molecular vibrations, to seconds, for stress-relaxation processes. Consequently, a hierarchy of methods is available to address particular molecular systems and properties of interest. It should be noted here that the communities working on the various levels are pretty diverse. In **Figure 1**, time- and length scales relevant for materials and process development are sketched. The focus of this review is linking the mesoscale to specific chemistries by employing quantum chemistry for a parameterization of the mesoscale simulation method.

At the quantum molecular level (relevant to deliver reaction mechanisms and kinetic rate coefficients), chemical systems of hundreds of atoms can be modeled today, and highly accurate calculations are possible for up to 20 atoms. For applications in reaction engineering, high-level methods usually are required.

Ab initio methods refer to a collection of approximations to solve the Schrödinger equation, as already envisioned by Dirac (12). They form a hierarchy of methods with increasing accuracy. The single-determinant Hartree-Fock (HF) wave function method is computationally cheap. However, it has the drawback that it neglects the correlated movement of the electrons—each electron experiences only the averaged potential of all other electrons (13). This leads to systematic errors that are not acceptable for calculations with chemical accuracy. Although HF energies are not accurate enough by themselves, almost all higher correlated methods are based on the HF reference wave function via an inclusion of further (excited-state) determinants into the ground-state wave function. Coupled cluster methods are currently the gold standard for validation (14). Benchmark calculations have clearly proven the reliability of coupled cluster methods, in particular of CCSD(T).

Density functional methods (DFT for short) (15, 16) are based on Hohenberg & Kohn's (16) theorem, which states that the exact ground-state energy of any molecular system can be

computed from the one-particle electron density. They introduce a one-particle functional that contains all many-body effects. With the knowledge of this functional, it would be possible to compute the exact ground-state energy. Although it has been shown that such a functional exists, its form is unknown. All current variants of functionals are based on the Kohn-Sham formalism (17), which states that functionals within the local density approximation or the generalized gradient approximation (GGA), meta-GGAs, or hybrid functionals are guesses or approximations of this functional. BP86 and B3LYP, the two long-popular DFT methods, represent a GGA- and a hybrid-GGA-functional, respectively. More recent functional developments, e.g., the M06 and M11 families of functionals from the Truhlar group (18), have started to become standard methods as well. Particularly useful are approaches to add empirical dispersion corrections to the functional (19). Practically, DFT is the workhorse of quantum chemistry nowadays. It is a good compromise between accuracy and computational effort. All systems of interest up to approximately 1,000 atoms can be calculated on the DFT level. The accuracy is usually insufficient for a quantitative prediction of properties that exhibit an exponential dependence on energy or entropy-like equilibrium constants or reaction kinetics (9). Error bars are difficult to estimate because there is no strict hierarchy of methods, and systematic improvements like perturbation theory, as in the case of HF, are not possible because the Hamiltonian in DFT is itself nonvariational (i.e., approximated). Of the above-mentioned methods, only DFT is affordable for real chemical applications. In addition to the study of chemistry in the gas phase, quantum chemical modeling can also be applied to condensed-phase systems by either an explicit or (much more popularly) an implicit solvation treatment via solvation models like PCM (polarizable continuum model) or COSMO (conductor-like screening model) (20), which are also typically based on DFT calculations.

More approximate classical atomistic methods can handle systems of up to millions of atoms, depending on the timescale. The atomistic or molecular scale encompasses a wide variety of computations, which are usually performed by molecular dynamics (MD) or Monte Carlo (MC) (21) methods using fully atomistic or united atom force fields. Properties described at this scale might range from thermodynamic properties (critical points, pressures) to transport properties (mass and heat transfer) and phase equilibria. By using statistical mechanics, the results of atomistic or molecular-scale calculations can then be applied to describe behavior at the mesoscopic and macroscopic scale (e.g., process or bulk properties). Gubbins & Moore (22), Maginn & Elliott (23), and Theodorou (24) have recently published excellent reviews on MD and MC methods in the context of engineering applications.

At the upper end, the mesoscale involves systems of billions of atoms, which still manifest molecular effects. There are many potential applications of molecular modeling in chemical processes in which predicting the characteristics and behavior of a system may be beneficial. By predicting system behavior, molecular modeling can potentially be used to improve the efficiency of existing systems (e.g., adhesives), as well as the design of new systems. It can help to shorten product and process development cycles, optimize processes to improve energy efficiency and environmental performance, and solve problems as they arise in plant operations.

At the macroscale, material behavior is described with continuum mechanical methods. Properties are fully linked to their morphologies on the micrometer level. Finally, bridging techniques (e.g., coarse graining) attempt to provide continuity and interface between the various scales, allowing the results of calculations at one scale to be used as input parameters in calculations at another scale. These techniques are particularly important for polymers because of the many time- and length scales involved, but they are far from mature. Nevertheless, in an industrial setting they must be applied to day-to-day problems.

Model Setup

How to choose a model? In soft matter simulations, relevant length scales can be as large as micrometers, and timescales are often in the order of milliseconds. Any atomic-level description is prohibitively expensive for complex formulations. Consequently, coarse graining—i.e., grouping several heavy atoms together to form a single bead and replacing the atomic-level interactions by simpler bead–bead interactions—represents an important approach to actually enable soft matter simulations of real-world systems.

The literature on coarse graining and related methods is abundant; it has been one of the big topics in soft matter simulation method development for more than 15 years. Excellent reviews (25–27) are available, and it is not the purpose of this article to give a comprehensive overview. From an application point of view, the choice of a suitable degree of coarse graining, the proper selection of beads, and their parameterization is time consuming when starting a set of simulations on new systems. However, a fast response time is one of the most crucial points in industry. If preliminary work takes weeks, modeling will most likely not be seen as a cost-efficient route to solution within a project. If a model is not applicable broadly enough for screening purposes, it will be of very limited use, e.g., only for understanding a benchmark system. Thus, only methods that can be parameterized automatically are the methods of choice. It is therefore no surprise that dissipative particle dynamics (DPD), one of the most prominent coarse-grained molecular modeling techniques, at least in industry, was largely developed at Unilever (28–36) and Shell (37, 38).

In the following, we sketch how to use DPD as a powerful semiautomated tool. First, we give a brief review of DPD basics. After describing BASF's parameterization scheme, we discuss some practical software requirements. Finally, we address the validation of the presented parameterization scheme.

How does dissipative particle dynamics work? Hoogerbrugge & Koelman (37) introduced DPD in the early 1990s. The method performs dynamic simulations of soft spherical particles (beads) interacting through a simple pairwise potential and ensures thermal equilibration through hydrodynamics. The time evolution of an interacting bead i is governed by Newton's equations of motion, given by

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i, \; \frac{d\vec{v}_i}{dt} = \vec{f}_i.$$

Here $\vec{r}_i$ and $\vec{v}_i$ represent position and velocity of particles. In the form given by Español & Warren (28), the forces $\vec{f}_{ij}$ acting on these particle consist of three pairwise additive terms:

$$\vec{f}_{ij} = \sum_{i \neq j} \left(\vec{f}_{ij}^{C} + \vec{f}_{ij}^{D} + \vec{f}_{ij}^{R} \right).$$

In the above formula, the sum runs over all beads other than i, which are centered within a certain cutoff radius r_c. As this is the only length scale in the system, the cutoff radius is used as the unit of length, i.e., $r_c = 1$. The three constituents of $\vec{f}_i$ from neighboring beads j are the conservative force $\vec{f}_{ij}^{C}$, the dissipative (or drag) force $\vec{f}_{ij}^{D}$, and the random force $\vec{f}_{ij}^{R}$.

The conservative force is a soft repulsion acting along the connecting line between the two centers of beads. It is given by

$$\left| \vec{f}_{ij}^{C} \right| = \begin{cases} a_{ij} \left(1 - r_{ij}/r_c \right), & r_{ij} < r_c \\ 0, & r_{ij} \geq r_c \end{cases},$$

where a_{ij} represents the maximum repulsion between particle i and particle j; r_{ij} is the distance between the centers of beads i and j positioned at $\vec{r}_i$ and $\vec{r}_j$, respectively; and

$$\hat{r}_{ij} = \frac{\vec{r}_i - \vec{r}_j}{r_{ij}}$$

defines the unit vector in the direction of the movement. Dissipative and random forces also act along the connecting lines of bead centers, and thus also depend on $\hat{r}_{ij}$ and are given by

$$\vec{f}_{ij}^{D} = -\gamma w^{D}(r_{ij})(\hat{r}_{ij} \cdot \vec{v}_{ij})\,\hat{r}_{ij}$$

and

$$\vec{f}_{ij}^{R} = \sigma w^{R}(r_{ij})\theta_{ij}(t)\,\hat{r}_{ij}.$$

Here, γ and σ are the friction parameter and the amplitude of the thermal noise, respectively. w^D and w^R are r-dependent weight functions vanishing for $r > r_c$.

$\vec{v}_{ij}$ is defined as $\vec{v}_{ij} = \vec{v}_i - \vec{v}_j$, with $\vec{v}_i$ and $\vec{v}_j$ being the velocities of particles i and j, and $\theta_{ij}(t)$ is a randomly fluctuating variable with Gaussian statistics. As suggested by the names, these two forces describe additional kinetic energy transfer between beads in a condensed phase and are chosen to conserve linear and angular momentum. Thus, the dissipative and random pairwise forces act as a thermostat for the simulation method.

The conservative soft repulsive forces a_{ij} between pairs of beads represent the essential chemistry of the system. Ideally, one would like to derive the conservative force for a given pair of beads from first principles simulations. In 1997, Groot & Warren (30) made an important contribution on this front by establishing a relationship between a simple functional form of the conservative repulsion in DPD and the Flory-Huggins χ-parameter theory. There are established theories in polymer science that link the χ-parameter to solubility and mixing energies of the polymeric components (39–42). The latter can be obtained either from atomistic simulations; from experiments; or, as will be shown in this work, by first principles calculations in connection with statistical thermodynamics, as implemented in the conductor-like screening model for real solvents (COSMO-RS) (43–45) framework. Thus, Groot & Warren provide a sound basis on which the conservative repulsion used in DPD can be derived starting from first principles simulations. Depending on the number density of particles, ρ, the following relationships between χ_{ij}^* and a_{ij} were derived:

$$\chi_{ij}^* = \left(a_{ij} - 25\right) \times (0.286 \pm 0.002), \quad \rho = 3$$

and

$$\chi_{ij}^* = \left(a_{ij} - 15\right) \times (0.689 \pm 0.002), \quad \rho = 5.$$

Here, 25 ($\rho = 3$) and 15 ($\rho = 5$) are the DPD interaction parameters between particles with the same identity. With this choice for parameters and densities, the compressibility of the DPD fluid was shown to be correct (for water). A detailed derivation of these quantities can be found in Groot & Warren's (30) paper.

The length and mass scales are set by specifying that the beads have mass $m = 1$ and, as mentioned before, also a cutoff distance for the interactions of one. Rather than to specify the units of time, simulations are performed in units such that $k_B T = 1$ (k_B representing the Boltzmann constant), which effectively defines a unit of time, because the root mean square velocity of the particles is $\sqrt{3}$, according to the Maxwell-Boltzmann distribution. Working in these units is useful because the conservative interaction potentials are automatically in the units of $k_B T$ and need not be rescaled.

It is noteworthy that the just-described standard DPD approach is constantly expanded to include electrostatic interactions, beads of different size, and species-dependent a_{ii} parameters (46). Many researchers use only the DPD thermostat without its pairwise companion with, e.g., Lennard-Jones interactions (47).

How to obtain parameters? As pointed out below, the χ_{ij}^{*}-parameter is connected to the free energy of mixing of the components *i* and *j*. This connection means that one could estimate the χ_{ij}^{*}-parameter from experiments or calculate it through various methods, e.g., MD or MC, or equations of state. Some of these options are discussed below. It should be noted up front that all of these methods suffer from the need to somehow obtain parameters. Thus, a scheme based on quantum-chemical calculations and COSMO-RS theory was developed.

Mapping the chemistry to beads: atomistic force fields versus COSMO-RS. The beading of a chemical system and the workflow to compute its properties are schematically outlined in **Figure 2**. In this particular example, a mixture of a C-28 alcohol and water is simulated, and thermodynamic properties like solubility and phase behavior are derived. First, the modeler must decide which partitioning of the molecules is appropriate. To have beads of a similar size (which is a prerequisite in DPD), groups of four nonhydrogen atoms are lumped into one bead. In such a beading scheme, one water bead will consist of four water molecules. In the present example, the beading is very crude to keep it simple. All C-4 units are assumed to be identical, even the end group. What is needed now is a set of binary interaction parameters describing the interaction between beads representing the specified chemistry. To finally start the simulation, the input structure and input files must be created. Thus, the typing must be done, and the 3D simulation box must be built. This is—particularly in academia—usually a time-consuming step if the simulation box is complicated and large. Often, a lot of interactive work is needed for a single system, which is of course prohibitive in industry. Manual labor is an expensive resource; thus, the construction of the system should be automated and also suitable for mass screening. We knew of no professional solution for our kind of simulations, involving hundreds of systems with up to 5 million particles and allowing for automated set-up of simulations, including pre-equilibration of the simulation

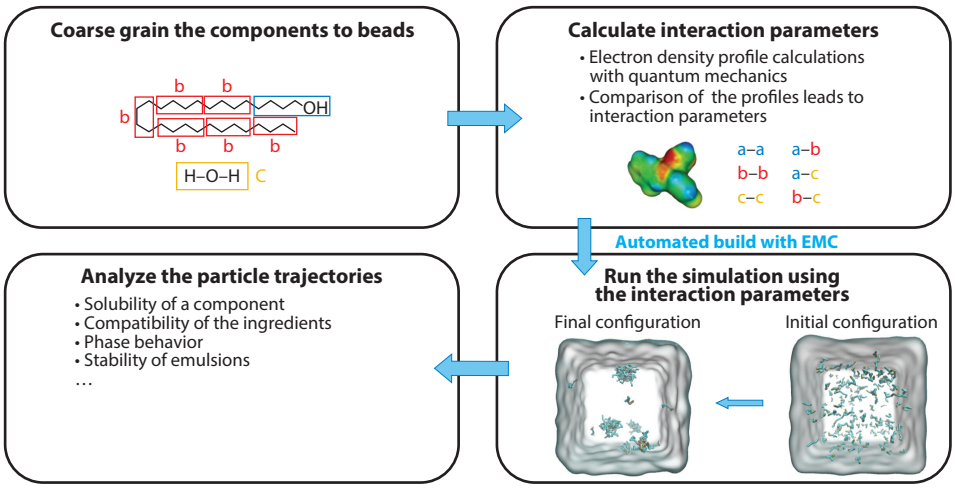

Figure 2

Beading process and workflow of a coarse-grained simulation based on pair potentials (here dissipative particle dynamics parameterized with Cosmo RS). Abbreviation: EMC, enhanced Monte Carlo.

box. Therefore, we at BASF created our own solution, named EMC (enhanced Monte Carlo), which is described in more detail below and has since been made open source. For the evaluation of interaction parameters, a BASF solution was developed for the same reasons. Established theories in polymer science, which link the χ-parameter to solubility and mixing energies of the polymeric components, serve as a starting point. Those properties can be computed from atomistic simulations, from equations of state derived from experiments or—as shown here—by first-principles calculations in connection with the statistical thermodynamics as implemented in the COSMO-RS framework.

A clear distinction exists between modeling of chemicals (small molecules) and modeling of polymers, especially when one intends to model thermodynamics with engineering fidelity (in other words, predictive numbers for, e.g., process design or, in our case, parameterization of interaction parameters). Several well-developed methods are available to compute data necessary for process engineering for small molecules. Eckl et al. (48, 49) and others have documented their success in the prediction of thermophysical data in many publications and contests (**http://fluidproperties.org/simulation-challenge-publications#4th**) showing that molecular modeling methods are a convenient and reliable tool for providing accurate data. Other approaches, like the aforementioned COSMO-RS, achieve similar accuracy (50). The main advantage of the latter method from an industrial perspective is that, with little effort and time, it is possible to provide high-quality data without the need for costly reparameterization. A variety of physical properties of chemical compounds can be modeled with these approaches (43).

With polymers, however, the situation is different, although there are very promising attempts to modify COSMO-RS for polymers (51). The reason is the need for a proper treatment of entropy for molecules with a large number of conformations like polymers. From the viewpoint of statistical mechanics, the proper approach for this type of molecule is MD and MC. Beautiful and remarkably accurate results have been published by Theodorou & Suter (52) and Banaszak et al. (53), among others, proving that deep insight into complex systems can be obtained provided a suitable force field is given. Siepmann et al. (54–57) have put effort into developing a transferable force field suitable for a wide range of conditions; however, it is incomplete compared with requirements in polymer research.

For practical purposes, however, the key point is the nonavailability of accurate force fields. Compared with the situation in biophysics, where force fields such as OPLS (58, 59), CHARMM (60), AMBER (61), and others are available and well parameterized for physiological conditions (298.15 K, normal pressure), there is a remarkable lack of general-purpose force fields for polymer modeling with sufficient accuracy, although recent developments of generalized force fields like gaFF (62) and CGenFF (63, 64) are trying to fill that gap. One of the reasons is the large parameter space of polymer simulations. Polymers must be simulated in solid state, melt, and solution. The relevant temperature range is from 200 K (T_g acrylates) to more than 500 K (processing engineering plastics). Relevant pressures range from normal pressure to a few hundred bar (e.g., extrusion). No available force field can treat these different conditions with acceptable accuracy; thus, it is very likely that a project using simulation as a tool for enhanced product development will start with force field (atomistic and/or coarse-grained) development and validation efforts. Another reason for the lack of a general-purpose force field is the nature of the force fields themselves. Typically, the nonbonded part is described by Lennard-Jones types of interaction, although many others have been used and are available, e.g., in solvers like LAMMPS (65), where many styles are implemented. We cannot expect a single set of parameters to describe the situation in a melt with similar accuracy as, for example, in solution or with solids. To complicate matters further, the validation of force fields requires accurate data from experiment, which are usually not at hand.

Equations of state are a convenient source of thermodynamic data. Particularly popular within BASF is PC-SAFT (66), a variant of the statistical associating fluid theory (SAFT) equation of state. SAFT in its various flavors (67) is based on Wertheim's ideas and has been gaining ground as an equation-of-state approach for a wide variety of fluid systems. In Wertheim's theory (68–71) of associating fluids, a multidensity formalism is introduced in which the expansion of the free energy is expressed in terms of the densities of the various bonding states of the monomers. Wertheim developed the theory for the simplest model of associating fluids, namely, hard spheres with one, two, or more bonding sites. Owing to their simplicity, accuracy (if properly parameterized), and low computational demands, equations of state are at present the method of choice when thermodynamic information is required for polymer process modeling. Consequently, there are numerous ongoing activities to broaden the scope of applicability [initially developed for hard spheres, then parameterized for alkanes, and now even applied to polyelectrolytes] and to facilitate parameterization [e.g., group-contribution SAFT (36, 72)]. Some groups also use molecular simulation to provide parameters for equations of state. In principle, this is a further parameterization option, although presently not used, for mesoscale simulations. In the following, we discuss an approach based on ab initio calculations and COSMO-RS to derive meaningful pair-interaction parameters for any chemical system of interest without the need for chemistry-dependent parameters.

COSMO-RS: basics and computation of Gibbs free energies of mixing. The COSMO-RS theory was established a decade ago. Excellent reviews on COSMO-RS have been published (45), and thus the underlying concepts are sketched only briefly here. Replacement of the ensemble of interacting molecules with the corresponding ensemble of independent, pairwise interacting surface pieces forms the central concept of COSMO-RS. Hence, this approximation results in a lack of molecular neighborhood information about these surface pieces and therefore a loss of structural or steric information, which is similar to the basic assumptions of group-contribution methods (GCMs) like UNIQUAC or UNIFAC (e.g., 73, 74). The great advantage of such an approximation is the extreme reduction in the complexity of the problem, allowing for a fast and yet accurate solution.

The screening charge density σ is the only descriptor determining the residual part of interaction free energies. Thus, the ensemble of surface pieces characterizing a fluid mixture S is sufficiently described by the distribution function $p^S(\sigma)$ (σ-profile), which gathers the amount of surface in the ensemble that exhibits a screening charge density between σ and $\sigma + d\sigma$. The surface charge distribution then contributes to both an electrostatic misfit term, which covers more or less dipolar interactions, and a hydrogen bonding term, which requires highly positive or negative surface charges to adopt values other than zero.

In case of a pure compound i, the σ-profile of the system S would simply be the σ-profile $p^i(\sigma)$ of this species. σ-Profiles of single compounds are obtained from quantum chemical calculations; for this reason, almost any chemistry can be studied, which makes COSMO-RS quite attractive for industrial applications (75).

In COSMO-RS theory, the chemical potential μ of the solute i in its solution environment, which represents a partial molar Gibbs free energy, is composed of the following contributions:

$$\mu^i = \mu^i_{res} + \mu^i_{perm} + \mu^i_{comb}.$$

Here, μ_{res} is the residual part that results from the electrostatic interactions discussed above. The permutational contribution μ_{perm} is $RT\ln(x_i)$, which reflects the ideal concentration dependence of the chemical potential. The last constituent is the combinatorial contribution μ_{comb}, which arises from the different shapes and sizes of species in a mixture. In the present parameterization of COSMO-RS, similar to in GCMs, a Staverman-Guggenheim-like term (76) is added. This is

not discussed further here, as the authors' recommendation for the computation of interaction parameters is to base it only on μ_{res}; it is assumed that the statistical, and thus entropic, features of a considered system are dealt with explicitly in the DPD simulation.

With any predictive thermodynamic method that yields chemical potentials (or activity coefficients), thermodynamic functions of mixing can be computed for a fluid binary system (containing species *i* and *j*); e.g., the Gibbs free energy of mixing is obtained via

$$\Delta G^{mix,M} = \{x_i \mu^i(x_i) + x_j \mu^j(x_j)\} - \{x_i \mu^i(1) + x_j \mu^j(1)\}.$$

The above formula refers to one mole of species *i* and *j* (altogether) in the mixture. x_i is the mole fraction of species *i*. $\mu^i(1)$ is the chemical potential of species *i* in pure species *i*. Such thermodynamic properties of mixing can be expected to yield unbiased and complete information about intermolecular interactions, which is required for molecular simulations: Thermodynamic functions of mixing contain both response of the solute to the solvent and response of the solvent to the presence of a solute. The ability to compute μ^i for all compounds in a mixture is a very valuable advantage compared with interaction parameter determination from experiments, which often do not yield information on chemical potentials (or activity coefficients) of both solvent and solute. The following section outlines the derivation of interaction parameters from mixing thermodynamics within the COSMO-RS framework using the software COSMOtherm.

From mixing thermodynamics to Flory-Huggins-like interaction parameters. A correct description of mixing thermodynamics requires the knowledge of effects from nonideality, e.g., observed for a binary system consisting of species or beads *i* and *j*, if interactions *i-i* and *j-j* are not equal to *i-j*. A quantification of these differences in interaction thermodynamics leads to (binary) parameters.

There is no unique, adequate choice of interaction parameters for coarse-grained simulations; rather, every simulation method will require its special description of nonideal behavior to yield an optimal agreement with the experiment. It has to be noted that here it is assumed that the coarse-grained simulation takes care of all permutational and combinatorial entropy within the system. This means that interaction parameters are dominated by enthalpic contributions and additionally can include entropic effects beyond permutation (e.g., destabilization of hydrogen bond–forming species in hydrophobic media is taken into account by the interaction parameter as well, although this effect is understood rather as a lowering of entropy) (77).

Because the mesoscale simulation takes care of combinatorial effects arising from all size and shape effects of components in a mixture, the combinatorial contribution within COSMO-RS should not be included for computation of interaction parameters. Hence, the thermodynamic functions of mixing to be used in the calculations of binary interaction parameters for DPD simulations contain only residual contributions. The corresponding purely residual Gibbs free energies of mixing are then

$$\Delta G_{res}^{mix,M} = \{x_i \mu_{res}^i(x_i) + x_j \mu_{res}^j(x_j)\} - \{x_i \mu_{res}^i(1) + x_j \mu_{res}^j(1)\}$$

for one mol species (*i* and *j* together). In our convention of interaction parameters, consistent with the choice of many others (30, 74, 78), these Gibbs free energies are divided by *RT*, the latter representing the product of gas constant and temperature. An advantage of this definition is that temperature thus enters into the simulation only via the underlying interaction parameters, which are dimensionless (and, of course, temperature dependent). Additionally, interaction parameters between identical beads are by definition zero. The binary interaction parameter χ_{ij}^* is then computed as

$$\chi_{ij}^* = \frac{\Delta G_{res}^{mix,M}}{RT\,\varphi_i \varphi_j} \frac{V_{ref} N_A}{V_M}.$$

Here, i and j denote species of any size, ϕ_i represents the volume fraction of species i, and the reference volume V_{ref} is the volume of one bead in the mesoscopic simulation. V_{ref} will typically be very small, i.e., below or in the order of 1 nm^3. The transition from the above molar expressions to these microscopic dimensions is performed via dividing V_{ref} by the average volume of one species in the mixture of i and j, which is V_M/N_A (with the Avogadro constant N_A). The volume of a bead can be related to the size of solvent molecules or chain segments of a polymer, which, however, does not represent a necessary requirement.

As a consequence of the convention for Gibbs free energies, positive χ^*_{ij} indicates unfavorable binary interactions (compared with the pure compounds and omitting the ideal entropy of mixing), whereas negative χ^*_{ij} indicates that even without the (ideal) gain of entropy there is a driving force for mixing. As is clear from the above derivation, for any pair of species i and j, χ^*_{ij} depends on system composition (e.g., χ^*_{ij} for a 50:50 mixture is different from that for 90:10) and, of course, temperature. If the mesoscale simulation method allows the use of concentration-dependent interaction parameters, they can be computed and tabulated right away. Methods like standard DPD, however, work with one fixed interaction parameter per pair. In these cases, an obvious choice is the computation of χ^*_{ij} on a 50:50 mixture of species (by volume) as a standard.

An often-published alternative approach is the estimation of χ-parameters between polymer–polymer, polymer–solvent, or solvent–solvent systems from bulk atomistic simulations of blends and single-component systems via the relevant cohesive energies. In this approach, the cohesive energies are calculated for pure components i and j and their mixture by using a suitable atomistic method (mainly MD). The Flory-Huggins χ^*_{ij} parameter is defined in terms of energy of mixing as

$$\chi^*_{ij} = \frac{\Delta E^{mix}}{RT}\frac{V_{ref}N_A}{V_M},$$

with

$$\Delta E^{mix} = \phi_i\left(\frac{E^{cob}}{V_M}\right)_i + \phi_j\left(\frac{E^{cob}}{V_M}\right)_j - \left(\frac{E^{cob}}{V_M}\right)_{mix}.$$

E^{cob} is connected to the Hildebrand (79) solubility parameter δ_i via

$$\delta_i = \sqrt{\frac{E_i^{cob}}{V_M}}.$$

The quality of this parameterization strategy depends on the adopted force field. To ensure reliable calculations, the use of high-quality force fields and validation of the cohesive energy by comparison with experimental values is recommended. The advantage is that several commercial simulation packages (e.g., Materials Studio; **http://accelrys.com/products/collaborative-science/biovia-materials-studio/**) offer automated procedures. For a massive screening with diverse chemistries, it is very likely that the force field [e.g., COMPASS (80), a quite-flexible Class 2 force field, which is often used in industry] is incomplete, and that several compounds are insufficiently parameterized. Also, the procedure is time consuming. Details on the scheme, together with an application to nanocomposites, can be found in Reference 81.

Computational methods.

Turbomole. Within all modeling groups of BASF, TURBOMOLE (82, 83) is the quantum chemistry package of choice. COSMO-RS calculations were performed with the COSMOtherm software, using release C2.1 (Rev. 01.05). If not mentioned otherwise, interaction parameters χ^* were computed according to the above formula based on the most stable conformer of the corresponding molecular models.

LAMMPS. The large-scale atomistic/molecular parallel simulator, or LAMMPS (65), is the main particle simulator in BASF research, but it is also heavily used within 3M, P&G, and Corning (84), among others. Its modularity, flexibility, and large user base make it particularly suited for industrial use. Other open-source alternatives are the nanoscale molecular dynamics program (NAMD) (85) and Groningen machine for chemical simulations (GROMACS) (86), which have their strength in biophysics but lack variety in polymer-related force fields when compared with LAMMPS.

Charged systems are widespread in industrial practice (e.g., ionic surfactants or polyelectrolytes) and treated by an extension of the standard DPD treatment. Electrostatics are long ranged, but charges are nonlocalized, when coarse-grained systems are concerned. These types of interactions are treated by a pairwise solution of the Poisson-Boltzmann equation (33). Groot suggested using a polynomial fit to capture the functional form of this solution. However, this fit is not easily extended to represent the long-range aspect of electrostatics. To this end, González-Melchor and coworkers (87) introduced a solution, which is based on Slater exponentials, as used in quantum mechanics. Unfortunately, closer evaluation discloses an unsatisfactory fit. In the BASF implementation, an error function is implemented, which follows the polynomial fits from Groot much more closely. This form has the added benefit that it seamlessly fits within the framework of long-range electrostatic treatment by means of Ewald summation and its discretized counterpart (particle-particle-particle mesh using fast Fourier transforms), which can be found in most MD codes. For our purposes we introduce a long-range electrostatic treatment that ties into particle-particle-particle mesh Ewald summations without the need to alter existing programming within LAMMPS. This implementation follows an error function and is given by

$$E_{elec}(r) = \mathrm{erf}(c\ r)/r,$$

where r represents the interparticle distance. The constant c in this expression relates to the Bjerum length. The error function acts as a switching function to transition between the short-range diffuse nature and the standard tail of falling off electrostatics, which coincides with pairwise Poisson-Boltzmann behavior.

Building workflow. As already stressed above, to enable simulations in high throughput, one of the most crucial issues of soft matter modeling in an industrial setting is automation. Commercial builders from Biovia (formerly Accelrys) (**http://accelrys.com/products/collaborative-science/biovia-materials-studio/**), Scienomics (**http://www.scienomics.com/**), Schrödinger (**http://www.schrodinger.com/**), Culgi (**https://www.culgi.com/**), and Materials Design (**http://www.materialsdesign.com/**), to name a few, offer GUIs, scripting options, and predefined force fields. The most important tools for setup of condensed-phase simulation boxes are amorphous builders, which are based on MC techniques (52, 88).

The requirement of automatization of processes poses restrictions on the software used. Apart from the ability to work in parallel environments, the most important point is to be able to run every step in preparing, running, and analyzing the simulation results from scripts. Once each step is encapsulated in a script, it is possible to set up a workflow containing all of the steps. Subsequently, the scientist can then focus on scientific—rather than technical—aspects of the project.

To ensure optimal flexibility (e.g., force fields) and performance in a highly parallelized environment, BASF created its own building workflows, owing to the fact that no suitable commercial solution existed for the type of systems it intended to simulate. Additionally, available commercial solutions tend not to provide the scripting flexibility needed for efficiently scanning through parameter space. The BASF EMC solution is divided in two parts: (*a*) a wrapper script, which generates EMC and LAMMPS input scripts with minimized effort, and (*b*) a highly optimized

initial structure generator, based on MC principles and coded in C. The latter program carries the name EMC and forms the heart of the building package. The EMC program stems from long-standing development involving cavity-size work on small molecules and studies of semicrystalline polymers (89–91). The incorporation of LAMMPS into a workflow illustrates the need for consistent solutions in the complex environment of industrial modeling. The bulk of simulations performed in our environment focus on interfacial phenomena, hence the need for creation of multiphase or grafted systems.

Building methodology. The applied MC scheme in EMC resembles configurational bias methods, with the exception that no effort is taken to maintain detailed balance (92, 93). Configuration construction begins with a box representing a volume commensurate with the final density. This empty volume is filled up atom by atom, using the energy landscape stemming from the already-present atoms in the decision of where to place the next atom. Subsequently, the surrounding is relaxed through a local equilibration by means of standard MC displacement moves within a preset spherical volume. This scheme repeats itself until all the particles are inserted.

Flexibility in the choice of force fields forms an integral part of industrial modeling. EMC provides several standard type-able atomistic force fields, which include the all-atom polymer consistent force field (PCFF) (94), which in form is identical to its derivative COMPASS (80); the all-atom and united-atom optimized potential for liquid simulations (OPLS) (58, 59); the all-atom Chemistry at Harvard macromolecular mechanics force field (CHARMM) (60); and the united atom transferable potentials for phase equilibria (TraPPE) force field (54–57). Furthermore, EMC provides coarse-grained force fields, which include Marrink and coworkers' (95) MARTINI and Shinoda-DeVane-Klein (SDK) (96, 97), for both of which parameters are provided as reported in the literature. Of interest to industry is the Mainz theory group's recent paper on the automated parameterization of MARTINI (98)—EMC provides output ports to LAMMPS, PDB, and XYZ formats. The builder has been made open source; it can be downloaded from **http://montecarlo.sourceforge.net/**.

Validation. Spenley (99) already showed that DPD performs well concerning scaling relations for polymers in solvents and in polymeric melts in the Rouse regime. What remains to be seen is whether the DPD parameters evaluated with the approach outlined above allow for quantitative predictions for specific chemistries.

An energetic case: prediction of interfacial tensions. Interfacial tensions are a useful property to benchmark a set of interaction parameters for their suitability in a DPD simulation. Groot & Warren (30) presented an application to interfacial tensions between two different molten polymers. Also, Maiti & McGrother (100) investigated bead-bead interaction parameters in DPD and their relation to bead size, solubility parameter, and surface tension.

To illustrate the success of the method to compute binary interaction parameters, interfacial tensions between *n*-dodecane and three hydrophilic compounds with varying degrees of hydrophilicity (water, formamide, and glycerol) were calculated and compared with the experimental data available in the literature.

The reference volume used in the simulation is 0.054 nm^3. This reference volume corresponds roughly to 2 water molecules, 1 formamide, 1/2 glycerol, and 1/5 dodecane; i.e., each dodecane molecule is made up of 5 beads, 2 water molecules form 1 bead, and each formamide and glycerol molecule is represented by 1 and 2 beads, respectively. Binary interaction parameters between a dodecane bead and a bead from the hydrophilic compounds were computed at 20°C and are tabulated in **Table 1**.

Table 1 Volume V and computed χ^* (at $T = 20°C$ and with a reference volume of 0.054 nm^3) for the interaction of n-dodecane ($V = 0.278$ nm^3) with several hydrophilic media

	$V_{hydrophilic\ species}$ (nm^3)	computed χ^*
Water	0.026	+5.06
Formamide	0.058	+3.63
Glycerol[a]	0.116	+2.86

[a]For glycerol, the most stable conformer with no intramolecular hydrogen bridging has been used.

DPD simulations were performed using the above binary interaction parameters for 1:1 (v:v) mixtures of n-dodecane and the hydrophilic compounds. Owing to the chosen dimensions of the DPD box (25:10:10), two clear interfaces formed along the longest direction in all cases. The primary method used to compute the interfacial tension σ using molecular simulations, developed by Tolman (101) and refined by Kirkwood & Buff (102), computes the surface tension as an integral of the difference between the normal and tangential pressures $p_\perp(x)$ and $p_\parallel(x)$,

$$\sigma = \frac{1}{2}\int_0^{Lx} \left(p_\perp(x) - p_\parallel(x)\right) d_Z,$$

where, in our geometry, $p_\perp(x) = p_x(x)$ and $p_\parallel(x) = \frac{1}{2}\left(p_y(x) + p_z(x)\right)$, and L_x represents the simulation box length perpendicular to the interface. Away from an interface, $p_\perp = p_\parallel$, and the integrand vanishes. Therefore, the nonzero contributions to the above integral equation come from the interfacial region. In the case of an interface between two fluid phases, the integral can be replaced with an ensemble average of the difference between the normal and tangential pressures,

$$\sigma_{DPD} = \frac{1}{2}L_x\langle p_\perp - p_\parallel\rangle = \frac{1}{2}L_x\left[\langle p_{xx}\rangle - \frac{\langle p_{yy}\rangle + \langle p_{zz}\rangle}{2}\right],$$

with interfacial tension σ_{DPD}, which is computed from the anisotropy of the pressure tensors p. This translates into, e.g., SI units via

$$\sigma_{real-units} = \frac{k_B T}{r_c^2}\sigma_{DPD}.$$

In this formula, k_B represents the Boltzmann constant, and the cutoff radius r_c depends on the bead-number density ρ and the reference bead volume V_{ref}, as

$$r_c = (\rho V_{ref})^{\frac{1}{3}}.$$

Table 2 shows the resulting calculated interfacial tensions and compares to experiment.

Computed interfacial tensions compare well to experiment. This suggests that the definition of χ^* is very well suited for use in DPD simulations and will also lead to reasonable results if

Table 2 Computed and experimental interfacial tensions σ for n-dodecane with several hydrophilic compounds at $T = 20°C$

	$\sigma_{computed}$ (mN/m)		$\sigma_{experimental}$ (mN/m)
	Averaged result	Numerical error	
Water	54.7	±1.1	52.87 (78)
Formamide	39.3	±1.0	38.90 (103)
Glycerol	41.3	±3.2	43.34 (103)

properties are studied where experimental benchmark data do not exist. Oviedo-Roa et al. (78) recently published a very similar approach. It should be mentioned here that for such simple molecular species, as in the above benchmark case, a prediction of surface tensions has recently become directly accessible via a COSMO-RS-based QSPR model (103); i.e., the property of interest is already defined by the interaction parameters.

Application Example: Tailored Polymer Design for an Agro Formulation

Challenge. In this section, the tailored design of a polymer for a crop protection formulation is described. There are many types of crop protection formulations, depending on the active ingredient, target, and application itself (104, pp. 219–48). In particular, the challenge was to design a polymer for an emulsion concentrate, where the polymer would prevent the crystallization of a water-insoluble active ingredient dissolved in a given organic solvent. Because the concentrate is later emulsified with water by the farmer, the whole formulation still must stay at least kinetically stable during application. This means that a tailored polymer must convey compatibility with media of very different polarity.

A brute-force approach based on intuition and trial and error is unfeasible owing to millions of options in this case. Thus, synthesis guided by modeling is necessary. In the literature, the GCMs of Hansen (see 105) and Abraham (see 104, pp. 273–305) enable fast screening of components for higher efficacy and wetting. However, design of a new formulation with many components requires a particle-based approach in which enthalpy and entropy are properly treated. To our knowledge, there is no systematic method reported in the literature.

Prescreening of copolymeric constituents with quantum chemistry. In the first stage, the desired comonomer behavior is established. According to the presented challenge, the copolymer must prevent the crystallization of the active ingredient in an organic solvent at low temperatures. This necessitates the copolymer's compatibility not only with the active ingredient but also with the organic solvent that makes up the concentrated formulation. Considering the facts that the active ingredient must be compatible with the organic solvent and the organic solvent must be emulsified in water, i.e., it should be hydrophobic, only a handful of possible organic solvents remain.

In the next step, all of the possible comonomers are screened according to their interactions with the three key components of the formulation. As an approximation for interactions, the free energies of mixing are calculated using COSMO-RS theory (50:50 mixture at 298 K), as described above for the computation of DPD interaction parameters. The comonomers are then grouped and ranked according to their affinities toward the active ingredient, organic solvents, or water. By taking into account a certain quantitative threshold for these affinities, the number of suitable comonomers is kept at a reasonable quantity. A good starting point can be reverse engineering of the benchmark polymer. By calculating the affinities between the comonomers of the benchmark polymer and the formulation components, new comonomers with affinities falling into these ranges can be chosen. Whereas some comonomers can show affinity toward only one component, e.g., the organic solvent, others may show affinity toward more than one component simultaneously. Therefore, merely choosing the comonomer chemistry is not sufficient for designing the desired polymer. The copolymer ratios that will come into the picture later play a critical role in achieving the desired polymer behavior.

Screening of concentrates with dissipative particle dynamics. After the prescreening step, the number of possible comonomer combinations is reduced to the order of several tens per organic

solvent. For each solvent candidate, the same procedure must be repeated. This number is then multiplied with the copolymer ratios that must be tested against the copolymer chemistry, because the amount of each building block will strongly affect the final copolymer behavior. Only at this stage are DPD-type particle simulations run to quantitatively explore the overall formulation behavior. In our experience, it is not necessary to spend too much time and effort considering minute differences in copolymer ratios, because a handful of combinations based on the possible boundary conditions already provide insight into the copolymer chemistry under investigation. Furthermore, at this stage, polymer synthesis experts should be consulted to clarify if the synthesis of the predicted polymer is feasible, because not all copolymers can be synthesized at every ratio. These restrictions can also significantly cut down the number of candidates to simulate.

The formulations to simulate contain not only the prescreened copolymers but also the active ingredient, the organic solvent (later the polar solvent as well), and many other additives. To obtain the binary interaction parameters to be used in DPD simulations of such complex mixtures, the same workflow as for the prescreening step is used, namely, the quantum chemical calculations running in an automatized manner. Only with such a fast and accurate computational workflow can one simulate complex mixtures of varying chemistry in a limited time frame. The structures for the quantum-chemical calculations are built with EMC and automatically submitted for geometry optimization and the calculation of the sigma surface necessary for the COSMO-RS calculations. The parameterization of solid active ingredients requires extra attention owing to their crystalline nature (104, pp. 250–72). This necessitates a refined parameterization based on experimental information, namely, differential calorimetry scanning measurements and solubility in various solvents.

The building of condensed phase systems and running of the DPD simulations are also automatized and rest on EMC for building and on LAMMPS as the simulation engine. The compositions of the simulation boxes are kept as close to reality as possible to provide quantitative predictions prior to experiments. A cubic box containing at least 100,000 particles is required to include all the chemical species in a formulation with a statistical significance. Boxes that are too large, however, are computationally intensive and do not provide significantly better statistics. The largest chemical component in size is the copolymer chain itself. Hence, one must ensure that the chains are neither too short, which leads to increased solubility, nor too long, which might lead to artefacts owing to finite box size. The descriptors used to predict the formulation behavior are based on statistical analysis of the trajectories but are not discussed further here; neither is the simulation of the dilution step with water, which mimics the application of the formulation in the field.

After simulating the emulsified formulations, the number of copolymer candidates is expected to come down drastically. In this example, copolymer lead structures were found that led to a new formulation polymer family and contributed to patent applications. A total of >10,000 polymers were screened in approximately 12 months, demonstrating the suitability for predictive screening of the described machinery, i.e., the coupling of quantum chemistry, Cosmo-RS, and DPD.

SUMMARY POINTS

1. Materials modeling is an established and rapidly growing discipline in the chemical industry.
2. Modeling is expected to deliver fast and predictive models, which contribute to more efficient product development.

3. Smart modeling workflows can be constructed using the current state of the art, but they are far from being black-box solutions. Improvement is highly welcome.
4. With carefully adjusted and validated methods, significant contributions to real-time product development are possible.
5. From an industrial perspective, fast generation of parameters properly representing a specific chemistry is of utmost importance.

FUTURE ISSUES

1. The combination of mechanistic understanding based on experimentation and modeling and complementary data-driven optimization based on (historical) data will accelerate the development of new materials and processes considerably.
2. In particular, emerging big data approaches (which were not discussed here) will be main drivers for rapid development and will most likely bring modeling (mechanistic, stochastic) much closer to production and customer needs.
3. Method development in multiscale modeling is a must. We are still far from the modeling of real systems, although the progress made in the past 10 years is fantastic.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Goh GB, Eike DM, Murch BP, Brooks CL. 2015. Accurate modeling of ionic surfactants at high concentration. *J. Phys. Chem. B* 119:6217–24
2. Toivakka M, Alam P, Touaiti F, Oravilahti A, Oravilahti T, et al. 2015. Understanding coating strength at the molecular and microscopic level. *Tappi J.* 14:373–84
3. Papatzani S, Paine K, Calabria-Holley J. 2015. A comprehensive review of the models on the nanostructure of calcium silicate hydrates. *Constr. Build. Mater.* 74:219–34
4. Jain S, Ginzburg VV, Jog P, Weinhold J, Srivastava R, Chapman WG. 2009. Modeling polymer-induced interactions between two grafted surfaces: comparison between interfacial statistical associating fluid theory and self-consistent field theory. *J. Chem. Phys.* 131:044908
5. Ma LK, Srivastava R, Barpanda D, Fowler T, Theophanous T, Verghese N. 2013. An inverse approach to characterize anisotropic thermal conductivities of a dry fibrous preform composite. *J. Reinf. Plast. Compos.* 32:1916–27
6. Christensen J, Albertus P, Sanchez-Carrera RS, Lohmann T, Kozinsky B, et al. 2012. A critical review of Li/air batteries. *J. Electrochem. Soc.* 159:R1–R30
7. Albertus P, Girishkumar G, McCloskey B, Sanchez-Carrera RS, Kozinsky B, et al. 2011. Identifying capacity limitations in the Li/oxygen battery using experiments and modeling. *J. Electrochem. Soc.* 158:A343–51
8. Sasabe H, Takamatsu J, Motoyama T, Watanabe S, Wagenblast G, et al. 2010. High-efficiency blue and white organic light-emitting devices incorporating a blue iridium carbene complex. *Adv. Mater.* 22:5003–7

9. Deglmann P, Müller I, Becker F, Schäfer A, Hungenberg KD, Weiß H. 2009. Prediction of propagation rate coefficients in free radical solution polymerization based on accurate quantum chemical methods: vinylic and related monomers, including acrylates and acrylic acid. *Macromol. React. Eng.* 3:496–515
10. Jain A, Ong SP, Hautier G, Chen W, Richards WD, et al. 2013. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater.* 1:011002
11. Natl. Sci. Technol. Counc. 2011. *Materials Genome Initiative for Global Competitiveness 2011*. Washington, DC: Natl. Sci. Technol. Counc. **https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf**
12. Dirac PAM. 1929. The basis of statistical quantum mechanics. *Proc. Camb. Philos. Soc.* 25:62–66
13. Szabo A, Ostlund NS. 1996. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. New York: Dover
14. Gauss J. 1998. Coupled-cluster theory. In *Encyclopedia of Computational Chemistry*, ed. P. von Ragué Schleyer. New York: Wiley
15. Koch W, Holthausen C. 2001. *A Chemist's Guide to Density Functional Theory*. Weinheim: Wiley-VCH. 2nd ed.
16. Hohenberg P, Kohn W. 1964. Inhomogeneous electron gas. *Phys. Rev. B* 136:B864
17. Kohn W, Sham LJ. 1965. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140:A1133
18. Zhao Y, Truhlar DG. 2008. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* 41:157–67
19. Grimme S, Antony J, Ehrlich S, Krieg H. 2010. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* 132:154104
20. Tomasi J, Mennucci B, Cammi R. 2005. Quantum mechanical continuum solvation models. *Chem. Rev.* 105:2999–3093
21. Frenkel D, Smit B. 2001. *Understanding Molecular Simulation: From Algorithms to Applications*. San Diego: Academic
22. Gubbins KE, Moore JD. 2010. Molecular modeling of matter: impact and prospects in engineering. *Ind. Eng. Chem. Res.* 49:3026–46
23. Maginn E, Elliott J. 2010. Historical perspective and current outlook for molecular dynamics as a chemical engineering tool. *Ind. Eng. Chem. Res.* 49:3059–78
24. Theodorou DN. 2010. Progress and outlook in Monte Carlo simulations. *Ind. Eng. Chem. Res.* 49:3047–58
25. Langner KM, Sevink GJA. 2012. Mesoscale modeling of block copolymer nanocomposites. *Soft Matter* 8:5102–18
26. Moeendarbary E, Ng TY, Zangeneh M. 2010. Dissipative particle dynamics in soft matter and polymeric applications—a review. *Int. J. Appl. Mech.* 2:161–90
27. Voth G, ed. 2009. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. Boca Raton, FL: CRC Press, Taylor & Francis Group
28. Español P, Warren P. 1995. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* 30:191–96
29. Goujon F, Malfreyt P, Tildesley DJ. 2004. Dissipative particle dynamics simulations in the grand canonical ensemble: applications to polymer brushes. *ChemPhysChem* 5:457–64
30. Groot RD, Warren PB. 1997. Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* 107:4423–35
31. Groot RD, Madden TJ. 1998. Dynamic simulation of diblock copolymer microphase separation. *J. Chem. Phys.* 108:8713–24
32. Groot RD, Madden TJ, Tildesley DJ. 1999. On the role of hydrodynamic interactions in block copolymer microphase separation. *J. Chem. Phys.* 110:9739–49
33. Groot RD. 2003. Electrostatic interactions in dissipative particle dynamics—simulation of polyelectrolytes and anionic surfactants. *J. Chem. Phys.* 118:11265–77
34. Malfreyt P, Tildesley DJ. 2000. Dissipative particle dynamics simulations of grafted polymer chains between two walls. *Langmuir* 16:4732–40

35. Noro MG, Meneghini F, Warren PB. 2003. Application of dissipative particle dynamics to materials physics problems in polymer and surfactant science. *Mesoscale Phenom. Fluid Syst.* 861:242–57
36. Jones JL, Lal M, Ruddock JN, Spenley NA. 1999. Dynamics of a drop at a liquid/solid interface in simple shear fields: a mesoscopic simulation study. *Faraday Discuss.* 112:129–42
37. Hoogerbrugge PJ, Koelman JMVA. 1992. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.* 19:155–60
38. Koelman JMVA, Hoogerbrugge PJ. 1993. Dynamic simulations of hard-sphere suspensions under steady shear. *Europhys. Lett.* 21:363–68
39. Abrams DS, Prausnitz JM. 1975. Statistical thermodynamics of liquid-mixtures—new expression for excess Gibbs energy of partly or completely miscible systems. *AIChE J.* 21:116–28
40. Flory PJ. 1944. Thermodynamics of heterogeneous polymers and their solutions. *J. Chem. Phys.* 12:425–38
41. Flory PJ. 1953. *Principles of Polymer Chemistry*. Ithaca, NY: Cornell Univ.
42. Flory PJ. 1942. Thermodynamics of high polymer solutions. *J. Chem. Phys.* 10:51–61
43. Klamt A. 2005. *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*. Amsterdam: Elsevier
44. Klamt A, Eckert F. 2000. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib.* 172:43–72
45. Klamt A, Eckert F, Arlt W. 2010. COSMO-RS: an alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annu. Rev. Chem. Biomol. Eng.* 1:101–22
46. Kacar G, Peters EAJF, de With G. 2013. A generalized method for parameterization of dissipative particle dynamics for variable bead volumes. *EPL* 102:40009
47. Soddemann T, Dünweg B, Kremer K. 2003. Dissipative particle dynamics: a useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E* 68:046702
48. Eckl B, Vrabec J, Hasse H. 2008. On the application of force fields for predicting a wide variety of properties: ethylene oxide as an example. *Fluid Phase Equilib.* 274:16–26
49. Eckl B, Vrabec J, Hasse H. 2008. Molecular modelling and simulation for the process design. *Chem. Ing. Tech.* 80:25–33
50. Eckert F, Klamt A. 2002. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* 48:369–85
51. Loschen C, Klamt A. 2014. Prediction of solubilities and partition coefficients in polymers using COSMO-RS. *Ind. Eng. Chem. Res.* 53:11478–87
52. Theodorou DN, Suter UW. 1985. Detailed molecular structure of a vinyl polymer glass. *Macromolecules* 18:1467–78
53. Banaszak BJ, Faller R, de Pablo JJ. 2004. Simulation of the effects of chain architecture on the sorption of ethylene in polyethylene. *J. Chem. Phys.* 120:11304–15
54. Chen B, Potoff JJ, Siepmann JI. 2001. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *J. Phys. Chem. B* 105:3093–104
55. Martin MG, Siepmann JI. 1998. Transferable potentials for phase equilibria. 1. United-atom description of *n*-alkanes. *J. Phys. Chem. B* 102:2569–77
56. Martin MG, Siepmann JI. 1999. Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes. *J. Phys. Chem. B* 103:4508–17
57. Wick CD, Martin MG, Siepmann JI. 2000. Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes. *J. Phys. Chem. B* 104:8008–16
58. Jorgensen WL, Tirado-Rives J. 1988. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–66
59. Jorgensen WL, Maxwell DS, Tirado-Rives J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–36
60. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, et al. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586–616

61. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, et al. 1995. A 2nd generation force field for the simulation of proteins, nucleic-acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–97
62. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25:1157–74
63. Vanommeslaeghe K, Pamidighantam S, Sheetz RM, Connolly JWD, Roitberg AE, MacKerell AD. 2009. Toward an automatic force field parametrization engine: assignment of parameters by analogy for the CHARMM General Force Field (CGenFF). *Abstr. Pap. Am. Chem. Soc.* 238:352–COMP
64. Vanommeslaeghe K, MacKerell AD. 2012. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J. Chem. Inform. Model.* 52:3144–54
65. Plimpton S. 1995. Fast parallel algorithms for short-range molecular-dynamics. *J. Comput. Phys.* 117:1–19
66. Gross J, Sadowski G. 2002. Modeling polymer systems using the perturbed-chain statistical associating fluid theory equation of state. *Ind. Eng. Chem. Res.* 41:1084–93
67. Muller EA, Gubbins KE. 2001. Molecular-based equations of state for associating fluids: a review of SAFT and related approaches. *Ind. Eng. Chem. Res.* 40:2193–211
68. Wertheim MS. 1984. Fluids with highly directional attractive forces. 2. Thermodynamic perturbation theory and integral equations. *J. Stat. Phys.* 35:35–47
69. Wertheim MS. 1984. Fluids with highly directional attractive forces. 1. Statistical thermodynamics. *J. Stat. Phys.* 35:19–34
70. Wertheim MS. 1986. Fluids with highly directional attractive forces. 3. Multiple attraction sites. *J. Stat. Phys.* 42:459–76
71. Wertheim MS. 1986. Fluids with highly directional attractive forces. 4. Equilibrium polymerization. *J. Stat. Phys.* 42:477–92
72. Lymperiadis A, Adjiman CS, Galindo A, Jackson G. 2007. A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ). *J. Chem. Phys.* 127:234903
73. Cunico LP, Hukkerikar AS, Ceriani R, Sarup B, Gani R. 2013. Molecular structure-based methods of property prediction in application to lipids: a review and refinement. *Fluid Phase Equilib.* 357:2–18
74. Schuld N, Wolf BA. 1999. Polymer-solvent interaction parameters. In *Polymer Handbook*, Vol. 7, ed. J Brandrup, EH Immergut, EA Grulke, pp. 247–64. New York: Wiley
75. Deglmann P, Schafer A, Lennartz C. 2015. Application of quantum calculations in the chemical industry—an overview. *Int. J. Quantum Chem.* 115:107–36
76. Staverman AJ. 1950. The entropy of high polymer solutions—generalization of formulae. *Recl. Trav. Chim. Pays-Bas* 69:163–74
77. Horvath C, Melander W, Molnar I. 1976. Solvophobic interactions in liquid-chromatography with nonpolar stationary phases. *J. Chromatogr.* 125:129–56
78. Oviedo-Roa R, Martínez-Magadán JM, Muñoz-Colunga A, Gómez-Balderas R, Pons-Jiménez M, Zamudio-Rivera LS. 2013. Critical micelle concentration of an ammonium salt through DPD simulations using COSMO-RS-based interaction parameters. *AIChE J.* 59:4413–23
79. Hildebrand JH, Scott RL. 1949. *The Solubility of Non-Electrolytes*. New York: Reinhold
80. Sun H. 1998. COMPASS: an ab initio force-field optimized for condensed-phase applications—overview with details on alkane and benzene compounds. *J. Phys. Chem. B* 102:7338–64
81. Fermeglia M, Ferrone M, Pricl S. 2003. Computer simulation of nylon-6/organoclay nanocomposites: prediction of the binding energy. *Fluid Phase Equilib.* 212:315–29
82. Ahlrichs R, Bar M, Häser M, Horn H, Kölmel C. 1989. Electronic structure calculations on workstation computers: the program system turbomole. *Chem. Phys. Lett.* 162:165–69
83. Furche F, Ahlrichs R, Hattig C, Klopper W, Sierka M, Weigend F. 2014. Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4:91–100
84. Schunk P, Pierce F, Lechman J, Grillet A, in 't Veld PJ, et al. 2012. Performance of mesoscale modeling methods for predicting rheological properties of charged polystyrene/water suspensions. *J. Rheol.* 56:353–84
85. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–802

86. Berendsen HJC, Vanderspoel D, Vandrunen R. 1995. GROMACS: a message-passing parallel molecular-dynamics implementation. *Comput. Phys. Commun.* 91:43–56
87. González-Melchor M, Mayoral E, Velázquez ME, Alejandre J. 2006. Electrostatic interactions in dissipative particle dynamics using the Ewald sums. *J. Chem. Phys.* 125:224107
88. Siepmann JI, Frenkel D. 1992. Configurational bias Monte Carlo: a new sampling scheme for flexible chains. *Mol. Phys.* 75:59–70
89. in 't Veld PJ, Hütter M, Rutledge GC. 2006. Temperature-dependent thermal and elastic properties of the interlamellar phase of semicrystalline polyethylene by molecular simulation. *Macromolecules* 39:439–47
90. in 't Veld PJ, Stone MT, Truskett TM, Sanchez IC. 2000. Liquid structure via cavity size distributions. *J. Phys. Chem. B* 104:12028–34
91. Kuppa VK, in 't Veld PJ, Rutledge GC. 2007. Monte Carlo simulation of interlamellar isotactic polypropylene. *Macromolecules* 40:5187–95
92. de Pablo JJ, Laso M, Siepmann JI, Suter UW. 1993. Continuum-configurational-bias Monte-Carlo simulations of long-chain alkanes. *Mol. Phys.* 80:55–63
93. Escobedo FA, de Pablo JJ. 1995. Extended continuum configurational bias Monte Carlo methods for simulation of flexible molecules. *J. Chem. Phys.* 102:2636–52
94. Sun H. 1994. Force field for computation of conformational energies, structures, and vibrational frequencies of aromatic polyesters. *J. Comput. Chem.* 15:752–68
95. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH. 2007. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* 111:7812–24
96. Shinoda W, DeVane R, Klein ML. 2007. Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Mol. Simul.* 33:27–36
97. Shinoda W, DeVane R, Klein ML. 2008. Coarse-grained molecular modeling of non-ionic surfactant self-assembly. *Soft Matter* 4:2454–62
98. Bereau T, Kremer K. 2015. Automated parametrization of the coarse-grained Martini force field for small organic molecules. *J. Chem. Theory Comput.* 11:2783–91
99. Spenley NA. 2000. Scaling laws for polymers in dissipative particle dynamics. *Europhys. Lett.* 49:534–40
100. Maiti A, McGrother S. 2004. Bead-bead interaction parameters in dissipative particle dynamics: relation to bead-size, solubility parameter, and surface tension. *J. Chem. Phys.* 120:1594–601
101. Tolman RC. 1948. Consideration of the Gibbs theory of surface tension. *J. Chem. Phys.* 16:758–74
102. Kirkwood JG, Buff FP. 1949. The statistical mechanical theory of surface tension. *J. Chem. Phys.* 17:338–43
103. Andersson MP, Bennetzen MV, Klamt A, Stipp SLS. 2014. First-principles prediction of liquid/liquid interfacial tension. *J. Chem. Theory Comput.* 10:3401–8
104. Jeschke P, Krämer W, Schirmer U, Witschel M. 2012. *Modern Methods in Crop Protection*. Weinheim: Wiley-VCH
105. Khayet M, Fernández V. 2012. Estimation of the solubility parameters of model plant surfaces and agrochemicals: a valuable tool for understanding plant surface interactions. *Theor. Biol. Med. Model.* 9:45