

Review

# Recent Progress towards Chemically-Specific Coarse-Grained Simulation Models with Consistent Dynamical Properties

Joseph F. Rudzinski 

Max Planck Institute for Polymer Research, 55128 Mainz, Germany; rudzinski@mpip-mainz.mpg.de;  
Tel.: +49-6131-379-267

Received: 26 June 2019; Accepted: 16 August 2019; Published: 20 August 2019



**Abstract:** Coarse-grained (CG) models can provide computationally efficient and conceptually simple characterizations of soft matter systems. While generic models probe the underlying physics governing an entire family of free-energy landscapes, bottom-up CG models are systematically constructed from a higher-resolution model to retain a high level of chemical specificity. The removal of degrees of freedom from the system modifies the relationship between the relative time scales of distinct dynamical processes through both a loss of friction and a “smoothing” of the free-energy landscape. While these effects typically result in faster dynamics, decreasing the computational expense of the model, they also obscure the connection to the true dynamics of the system. The lack of consistent dynamics is a serious limitation for CG models, which not only prevents quantitatively accurate predictions of dynamical observables but can also lead to qualitatively incorrect descriptions of the characteristic dynamical processes. With many methods available for optimizing the structural and thermodynamic properties of chemically-specific CG models, recent years have seen a stark increase in investigations addressing the accurate description of dynamical properties generated from CG simulations. In this review, we present an overview of these efforts, ranging from bottom-up parameterizations of generalized Langevin equations to refinements of the CG force field based on a Markov state modeling framework. We aim to make connections between seemingly disparate approaches, while laying out some of the major challenges as well as potential directions for future efforts.

**Keywords:** coarse-grained dynamics; bottom-up coarse-graining; Mori–Zwanzig; generalized Langevin equation; structural–kinetic relationships; Markov state models; time rescaling

## 1. Introduction

Particle-based, low-resolution molecular simulation models, which represent groups of atoms with a single coarse-grained (CG) site, have been instrumental in shaping our perspective of soft matter systems. From generic bead–spring polymer models [1] to experimentally-informed Gō-type protein models [2,3], representing a system with reduced resolution not only enables access to larger length and longer time scales, but also—and arguably more importantly—can help to elucidate the essential microscopic driving forces of emergent phenomena. The coarse-graining procedure is rigorously grounded by statistical mechanics, which can be leveraged to develop systematic routines for parameterizing the potential energy function governing interactions at the CG level of resolution as well as the corresponding equations of motion. As early as the 1930s, both Onsager and Kirkwood discussed the concept of the many-body potential of mean force (MB-PMF)—a state-point-dependent free-energy function which acts as the proper CG interaction potential for reproducing all structural and thermodynamic properties of the underlying system at the CG level of resolution [4,5]. A whole

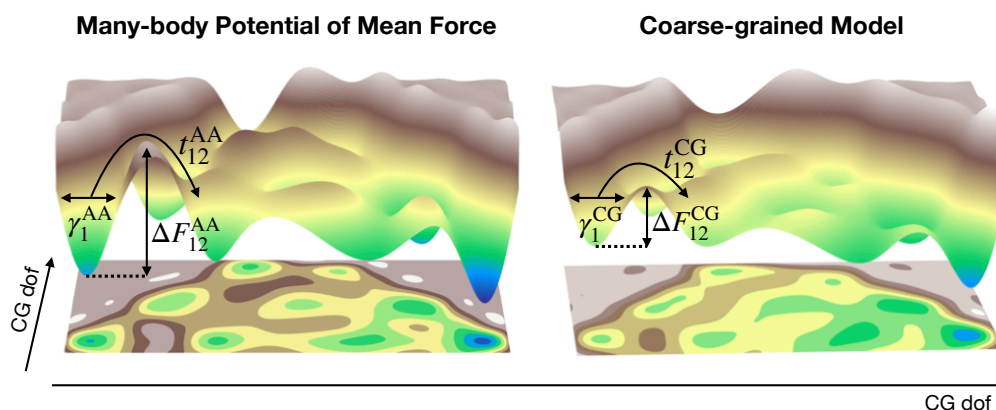
range of “bottom-up” techniques have been introduced to approximate this high-dimensional function using simulations of a higher-resolution model, normally with an atomically-detailed and classical representation [6]. These methods typically employ standard molecular mechanics potentials for this purpose, which in most cases correspond to a highly deficient basis set for approximating the MB-PMF. The reader is directed to [6–8] for instructive reviews on this topic.

A few decades following the work of Onsager and Kirkwood, Mori and Zwanzig independently developed a framework for mapping the dynamics of a microscopic system onto an arbitrary set of observables that are functions of the phase space variables of the underlying system [9,10]. The crux of this theory is a projection operator formalism which results in “projected dynamics” that evolve according to a generalized Langevin equation. In the context of particle-based CG models, this results in the time evolution of CG sites that is not only determined by the MB-PMF, but also by frictional and stochastic forces which account for the missing degrees of freedom. The latter forces are related through a fluctuation–dissipation theorem and depend in general not only on the coordinates and momenta of the CG sites but also on time. Similar to methods for approximating the MB-PMF, the Mori–Zwanzig (MZ) framework can be employed to parameterize the CG equations of motion, but only after significantly simplifying the form of the frictional and stochastic forces [11]. Despite the well-known result from Mori and Zwanzig, most bottom-up CG models are simulated according to a simple Langevin equation, i.e., with a single, uniform friction coefficient and the corresponding Gaussian distributed white noise for the stochastic force. This implementation assumes that the frictional forces are Markovian—not explicitly dependent on time—and independent of the coordinates and momenta of the CG sites. If this rather harsh assumption is valid, then the friction coefficient need not be explicitly calculated and, instead, any CG dynamical property may be interpreted via a uniform time-rescaling factor, which relates the sampled CG time scales with the true underlying time scales of the system. Remarkably, such a rescaling has been demonstrated in certain regimes for both configurational polymer dynamics and transport phenomena in liquids [12,13].

More generally, the time rescaling is not uniform and instead is a complicated function of the dynamical process of interest, as dictated by the nature of the MZ frictional force. As a consequence, not only is the direct connection to the true dynamical time scales of the system lost, but the distribution of pathways sampled by a CG model may be qualitatively incorrect. To better understand this concept, it is useful to consider the evolution of the system over the free-energy landscape determined by the CG potential and corresponding equations of motion (Figure 1). Even if we were able to perfectly reproduce the true underlying landscape (at the CG level of resolution), i.e., simulate the MB-PMF, a nontrivial MZ frictional force may be necessary to account for the reduced friction resulting from the removal of degrees of freedom. As already mentioned above, in practice, the landscape generated by a CG model is almost always a drastic approximation to the true landscape, since we typically employ a highly deficient basis set of interactions. This approximation is often interpreted as performing a smoothing of the MB-PMF, since CG models typically employ softer interaction potentials with respect to their all-atom (AA) counterparts. Smoothing generally results in nontrivial adjustments in the relative barrier heights along the landscape and almost certainly complicates the form of the MZ frictional force. Thus, both missing frictional forces as well as incorrect relative barrier heights may result in a qualitatively incorrect distribution of pathways sampled by a CG model.

Clearly, the challenge of accurately extracting dynamical properties from CG models is severely limiting for their predictive capabilities. At the same time, despite the generally daunting nature of the problem as laid out by the MZ framework, CG models have been extensively and successfully employed to gain insight into dynamical processes in soft matter. This is especially true for more generic CG models, which are often explicitly constructed using physical intuition about the driving forces for the phenomena of interest. This “built-in physics” restricts certain features of the free-energy landscape and, as a result, also the corresponding dominant pathways between states. For chemically-specific CG models, the situation is complicated by incorporating finer details into the model which, if not performed carefully, may lead to a misleading description of the characteristic

dynamical processes. In this case, researchers have approached the problem by: (i) explicitly correcting the dynamics following the MZ framework; (ii) employing empirical rescaling relations, valid for particular classes of systems, to match specific transport properties; or (iii) refining the CG interaction potential in an attempt to represent the hierarchy of long-time scale processes (i.e., kinetic properties) through the reproduction of the dominant free-energy barriers.



**Figure 1.** Illustration of the challenge of CG dynamics. Free-energy surfaces generated by the MB-PMF (left) and an approximate CG model (right) at the CG level of resolution (i.e., plotted along CG degrees of freedom (dof)). CG models lose connection with the true dynamics for two distinct reasons: (i) the friction,  $\gamma_1$ , within a free-energy basin is different from the AA model, even if the MB-PMF is employed, due to the missing degrees of freedom; and (ii) the approximate CG model effectively smooths over the MB-PMF, changing the relative barrier heights,  $\Delta F_{12}$ , between a pair of basins in a non-trivial way. As a result, not only is the time scale associated with hopping between pairs of basins, e.g.,  $t_{12}$ , not quantitatively reproduced, but the ratio of time scales with respect to two distinct transitions is changed. Furthermore, these errors propagate into the description of longer time scale processes, severely limiting the predictive capabilities of CG models.

In this review, we discuss recent work towards characterizing and improving the dynamical properties generated by CG simulation models, while attempting to build a conceptual connection between seemingly disparate approaches. We focus on relatively high-resolution CG representations, which aim to retain a high level of chemical specificity, thereby complicating the description of both static and dynamical properties of the system. At the same time, we occasionally look to representative examples of coarser or more generic models for intuition and inspiration. The review is organized as follows. Section 2 discusses bottom-up approaches following the MZ formalism, starting with a brief summary of the necessary theoretical foundations. Section 3 presents time-rescaling approaches, with a focus on polymers and liquids. Section 4 covers the free-energy-landscape perspective, including the use of structural–kinetic–thermodynamic relationships and Markov state models to assist in the interpretation and refinement of CG kinetic properties. Section 5 briefly highlights a few representative applications which provide significant challenges for the methods discussed in the preceding sections. Finally, Section 6 provides a brief overarching discussion and outlook for future work.

## 2. Bottom-Up Approaches Following the Mori–Zwanzig (MZ) Formalism

In this section, we review recent work that utilizes the MZ approach to address discrepancies in CG dynamics. In contrast to the original application of MZ to model the evolution of macroscopic observables [9,10,14–17], e.g., scattering functions [18], we focus here on modeling the time evolution of CG sites that represent groups of atoms.

### 2.1. Preliminaries

Consider an atomic system with coordinates  $\mathbf{r} = \{\mathbf{r}_1, \dots, \mathbf{r}_n\}$  and momenta  $\mathbf{p} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$  interacting according to a classical potential energy function  $u(\mathbf{r})$  in the canonical ensemble. The atomistic Hamiltonian is:

$$h(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^n \frac{1}{2m_i} \mathbf{p}_i^2 + u(\mathbf{r}), \tag{1}$$

where  $m_i$  represents the mass of atom  $i$ . The probability of finding the system in configuration  $\mathbf{r}$  is given by:

$$p_r(\mathbf{r}) = z(\mathbf{r})^{-1} \exp(-\beta u(\mathbf{r})), \tag{2}$$

where  $z(\mathbf{r}) = \int d\mathbf{r} \exp(-\beta u(\mathbf{r}))$  is the configuration integral and  $\beta = 1/k_B T$  is the inverse temperature. Let us define transformations from atomic coordinates and momenta to a lower resolution representation:  $\mathbf{R}_I = M_{\mathbf{R}_I}(\mathbf{r})$  and  $\mathbf{P}_I = M_{\mathbf{P}_I}(\mathbf{p})$  for each CG site  $I = 1, \dots, N$ . The mapping functions  $\{M_{\mathbf{R}_I}\}$  and  $\{M_{\mathbf{P}_I}\}$  are typically chosen as linear functions. We assume that the CG system also lies within the canonical ensemble with probability of configuration  $\mathbf{R}$  given by:

$$P_R(\mathbf{R}) = Z(\mathbf{R})^{-1} \exp(-\beta U(\mathbf{R})), \tag{3}$$

where  $Z(\mathbf{R}) = \int d\mathbf{r} \exp(-\beta U(\mathbf{r}))$  is the configuration integral and  $U(\mathbf{R})$  is a classical potential governing interactions at the CG level of resolution. One natural consistency criterion for the CG model is that the configurational probability distribution should match the distribution generated by the high-resolution model after being mapped to the CG representation [19]:

$$P_R(\mathbf{R}) = p_R(\mathbf{R}) \equiv \int d\mathbf{r} p_r(\mathbf{r}) \delta(M_{\mathbf{R}}(\mathbf{r}) - \mathbf{R}), \tag{4}$$

where  $\delta(M_{\mathbf{R}}(\mathbf{r}) - \mathbf{R}) \equiv \prod_{I=1}^N \delta(M_{\mathbf{R}_I}(\mathbf{r}) - \mathbf{R}_I)$ .  $p_R(\mathbf{R})$  is the probability that an atomic configuration is mapped to a particular CG configuration. By solving Equation (4) for the CG potential, we recover a special potential energy function,  $U^0(\mathbf{R}) \propto -k_B T \ln p_R(\mathbf{R})$ , which is known as the many-body potential of mean force (MB-PMF). An equivalent consistency criterion can be imposed for the distribution of CG momenta,  $\mathbf{P}$ , which is automatically satisfied for certain forms of the mapping functions  $\{M_{\mathbf{P}_I}\}$  [19].

The dynamical evolution of the high-resolution model is determined by Hamilton's equations of motion. Equivalently, the phase-space distribution,  $\rho(\mathbf{r}, \mathbf{p})$ , evolves according to the Liouville equation:

$$\frac{\partial \rho}{\partial t} + i\mathcal{L}\rho = 0, \tag{5}$$

where  $\mathcal{L} = \sum_{i=1}^n (\partial/\partial \mathbf{r}_i + \partial/\partial \mathbf{p}_i)$  is the Liouville operator. In the MZ approach,  $\mathcal{L}$  is decomposed into two parts using orthogonal operators  $P$  and  $Q$ :  $i\mathcal{L} = Pi\mathcal{L} + Qi\mathcal{L}$ , where the sum of  $P$  and  $Q$  is equal to the identity operator.  $P$  projects out the dynamics of the  $N$  CG degrees of freedom, which are generally considered to correspond to slowly relaxing degrees of freedom. Then, the "missing" or "removed" degrees of freedom (i.e., the degrees of freedom projected out by  $Q$ ) are meant to relax quickly and, thus, can be represented implicitly in the resulting CG model. One can then derive an expression for the equations of motion of CG sites, which evolve according to a generalized Langevin equation (GLE) [11,20–22]:

$$\frac{d\mathbf{P}}{dt}(t) = - \underbrace{\frac{d}{d\mathbf{R}} U^0(\mathbf{R}(t))}_{\text{conservative}} - \underbrace{\int_0^t dt' \Gamma(\mathbf{R}, \mathbf{P}, t-t') \mathbf{V}(t')}_{\text{frictional}} + \underbrace{\delta F^Q(t)}_{\text{random}}. \tag{6}$$

The chosen projection operator,  $P$ , determines the form of the dissipative forces [23]—i.e., the frictional and random forces.

In most applications, a Markovian assumption is made to remove the explicit time dependence from the friction kernel:  $\Gamma = \Gamma(\mathbf{R}, \mathbf{P})$ . Hydrodynamic-like interactions between CG sites are incorporated by retaining  $\Gamma$ 's dependence on  $\mathbf{P}$ , in contrast to traditional hydrodynamic approaches [24] that explicitly represent sites that mimic the effective behavior of the missing degrees of freedom, e.g., solvent. Most often, the friction kernel is assumed to depend only on the CG coordinates,  $\Gamma = \Gamma(\mathbf{R})$ , and can be further simplified by assuming a particular functional form. The most widely used approximation is to assume that the friction is pairwise additive and depends only on the distance between two particles. In this case, Equation (6) simplifies to a form which corresponds to the dissipative particle dynamics (DPD) approach [25]:

$$\frac{d\mathbf{P}_I}{dt} = - \sum_{J \neq I} \left[ \frac{d}{d\mathbf{R}_I} U_{IJ}(\mathbf{R}_{IJ}) + \gamma_{IJ}(\mathbf{R}_{IJ}) \mathbf{V}_{IJ} - \sigma_{IJ}(\mathbf{R}_{IJ}) \xi_{IJ} \right], \tag{7}$$

where  $\mathbf{V}_{IJ} = \mathbf{V}_I - \mathbf{V}_J$  is the relative velocity between CG sites  $I$  and  $J$ ,  $\gamma_{IJ}$  is the DPD pair friction, and  $\xi_{IJ}$  is a vector of Gaussian random variables with zero mean and unit variance.  $\gamma_{IJ}$  is often decomposed into parallel and transverse components [26].  $\sigma_{IJ}$  is related to  $\gamma_{IJ}$  through the fluctuation–dissipation relation:  $\sigma_{IJ} = \sqrt{2k_B T \gamma_{IJ}}$ . Here, we have also simplified the form of the conservative potential,  $U_{IJ}(\mathbf{R}_{IJ})$ , to be pairwise additive and depend only on the distance between two particles, as an approximation to the MB-PMF,  $U^0$ . Of course, by further simplifying the functional form of the friction kernel to be independent of particle identity and position, we will recover a simple Langevin equation which represents the “default” equations of motion for bottom-up CG models in practice.

We end this very brief introduction to the MZ formalism by pointing the interested reader to more detailed discussions of the approach: (i) Hijón et al. [11] provided a foundational discussion of the MZ approach for parameterizing CG equations of motion; (ii) Di Pasquale et al. [27] clearly demonstrated the link between MZ and other bottom-up coarse-graining approaches; and (iii) Izvekov [23] gave a thorough account of various approximations and derivations of GLEs used within the MZ framework.

### 2.2. Friction Kernel Parameterization Using Higher-Resolution Simulations

Assuming a reasonable approximation,  $U_{IJ}(\mathbf{R}_{IJ})$ , to the MB-PMF,  $U^0$ , the main challenge of the MZ approach is to determine an approximate friction kernel,  $\Gamma$ . A conceptually simple approach to this problem is to choose a functional form for the friction kernel and then tune the parameters to reproduce local dynamical properties, e.g., velocity–velocity time correlation functions (TCFs), of the CG sites. However, this parameterization can be practically difficult and may lead to unpredictable behavior due to uncontrolled approximations [28]. As a consequence, significant work has been done to develop more systematic and theoretically-founded methodologies for calculating the friction kernel.

Izvekov and Voth [29] proposed to determine the frictional forces directly from simulations of a higher-resolution model using a GLE formulation for TCFs. They approximated the lost friction due to missing degrees of freedom as the difference between the total atomistic force on each site for a given configuration and the force generated by the chosen conservative interactions for the CG model. According to this approximation, the frictional forces were then determined from force–velocity and velocity–velocity TCFs. Kauzaric et al. [28] expanded upon the link between TCFs and the MZ

frictional forces by proposing three distinct routes to the friction kernel based on Green–Kubo, Onsager, and Einstein–Helfand relations:

$$\Gamma_{IJ} = \frac{1}{k_B T} \int_0^t dt' \langle \delta \mathbf{F}_I(0) \delta \mathbf{F}_J(t') \rangle \quad (8)$$

$$= -\frac{M_I M_J}{k_B T} \frac{d}{dt'} \langle \mathbf{V}_I(0) \mathbf{V}_J(t') \rangle |_{t'=t} \quad (9)$$

$$= \frac{C_{IJ}^P(t) - C_{IJ}^P(0)}{k_B T t}, \quad (10)$$

where  $\langle \rangle$  denotes an ensemble average and  $C_{IJ}^P(t) = \langle \mathbf{P}_I(t) \mathbf{P}_J(t) \rangle$ . In Equation (8),  $\delta \mathbf{F}_I(t) = \mathbf{F}_I(t) - \mathbf{F}_I^C(t)$ , where  $\mathbf{F}_I(t)$  is the total force on CG site  $I$  from the higher-resolution simulation and  $\mathbf{F}_I^C(t)$  is the same quantity evaluated using the chosen CG conservative potential,  $U_{IJ}(\mathbf{R}_{IJ})$ . Each of these approaches assumes a time scale separation between the CG (slow) and missing (fast) degrees of freedom, requiring a careful choice of the time scale parameter,  $t$ . In particular, if the friction matrix is computed from unconstrained molecular dynamics simulations, the integral in Equation (8) vanishes at infinite times. This is commonly known as the “plateau problem”. More recently, Markutsya and Lamm [30] proposed to calculate the friction kernel from the fluctuating forces of the reference model via the fluctuation–dissipation theorem.

With any of these approaches, the question becomes how to evaluate the corresponding TCFs. According to the MZ formalism, the fluctuating force corresponds to dynamics in  $Q$  space, which we do not have direct access to. The simplest approach is to assume that the  $Q$ -space and real dynamics (i.e., those sampled by the higher-resolution reference model) are equivalent, leading to the above-mentioned plateau problem when applying the Green–Kubo relation (Equation (8)) to calculate the frictional forces. Hijón et al. [11] proposed to generate  $Q$ -space dynamics using simulations where the CG degrees of freedom, i.e., site positions, are constrained. The constrained dynamics can be obtained by appropriately modifying the equations of motion for the high-resolution model [11,31]. This approach avoids the plateau problem but can introduce artifacts when insufficient time scale separation exists between the  $P$  and  $Q$  degrees of freedom, as is often the case for high-resolution CG models. Izvekov [23,32,33] proposed an alternative approach which, by employing a distinct projection operator and representing the friction as a product of momentum-dependent Hermite polynomials, determines the frictional forces from unconstrained molecular dynamics simulations without the onset of the plateau problem. Furthermore, very recent work from Español and coworkers [34] derives corrected Green–Kubo relations that avoid the plateau problem in cases where the time scale separation is not extreme.

### 2.3. Dissipative Particle Dynamics (DPD)

As described above, if one assumes that  $\Gamma$  is only dependent upon CG site positions and is pairwise additive, the full GLE is transformed into the equations of motion employed in the DPD method. Although DPD models traditionally use soft-sphere potentials to describe mesoscale structure, conservative forces derived from, for example, structure-based CG methodologies can also be used within this framework [25]. Due to its practical simplicity and efficiency, this approach has emerged as a popular framework for constructing CG models that more accurately represent the dynamical properties of a higher-resolution reference model. The MZ-DPD approach has been applied to a range of molecular systems, especially for very coarse representations of star polymers and cluster-based mappings of fluids [11,21,35–39]. (We note that, although considerable effort has been made to derive the DPD equations from first principles for CG sites that represent clusters of nonbonded molecules [21,40–47], the precise connection to the microscopic dynamics remains unclear [25]. In fact, there appears to be no rigorous foundation for employing conservative forces to describe interactions between such supramolecular blobs [48].) As articulated very nicely in a recent paper by Deichmann and van der Vegt [22], these studies demonstrate success of the MZ-DPD approach in cases where there

is a relatively clear time scale separation between the CG ( $P$ ) and removed ( $Q$ ) degrees of freedom. Here, we focus on applications with higher-resolution CG representations, where this time scale separation does not clearly exist.

Izvekov and Voth [29] constructed one- and two-site models of liquid methanol, using force-velocity and velocity-velocity TCFs to determine scalar, but particle-dependent, friction constants. Conservative forces were obtained from the force-matching-based multiscale coarse-graining (MS-CG) method [19]. They demonstrated that the two-site representation results in accurate molecular center of mass velocity-velocity TCFs, although the diffusion constant was only reproduced within a factor of 2. Izvekov and Rice [21] later proposed an approach to match two-body force-velocity and three-body velocity-velocity TCFs for DPD models that represent a cluster of molecules with a single DPD site. Conservative forces were again obtained with the MS-CG method. The frictional forces were decomposed into parallel and transverse components and were not restricted to a particular functional form. The authors examined varying resolutions for liquid nitromethane, using a Voronoi tessellation to determine clusters of molecules that map to a single CG site. Of particular note, while coarser representations demonstrated Markovian behavior in the TCFs that could be accurately represented with the DPD models, the one-molecule-per-site mapping resulted in significant deviations in these observables due to a lack of time scale separation. It was also demonstrated that, although the parallel frictional contributions are dominant for all representations, the transverse friction contributes significantly to both viscosity and diffusion (as expected from previous investigations [26]). In a follow-up study from the same authors [49], the importance of the transverse frictional contributions was further established. Li et al. [37] also performed a detailed investigation into the effect of directional friction components in the context of coarser representations for star polymers.

Trément et al. [50] employed the approach of Hijón et al. (see Section 2.2) to derive DPD models with two distinct representations of 5 and 10 heavy atoms per CG site for liquids of  $n$ -pentane and  $n$ -decane. Using conservative potentials obtained from direct Boltzmann inversion, subsequently fitted to an analytic exponential form, they demonstrated qualitative agreement of both velocity-velocity TCFs and diffusion constants. In a subsequent study, Lemarchand et al. [51] applied the same approach to *cis*- and *trans*-1,4-polybutadiene and demonstrated reasonable agreement with the underlying self-diffusions and viscosities for a range of representations from one to six monomers per CG site. They demonstrated that dynamical accuracy increases with increasing level of coarse-graining, due to better fulfilment of the Markovianity assumption.

Deichmann et al. [52] applied the Green-Kubo approach (Equation (8)) to determine DPD friction matrices for three different molecular liquids with both one- and two-site representations, while using conservative potentials obtained with the conditional reversible work method [53]. While one-site models reproduced both the overall behavior of the velocity TCFs and also the diffusive dynamics reasonably well, the two-site representations failed to accurately capture dynamical properties, supposedly due to insufficient time scale separation between the CG and removed degrees of freedom. In a follow-up study, Deichmann and van der Vegt [22] investigated liquids, polymer solutions, and polymer melts based on repeat 2,2-dimethylpropane units. They demonstrated that the diffusion in single-component liquids and polymer solutions can be accurately modeled using the MZ-DPD formalism, despite a clear lack of time scale separation. In this case, there is not a one-to-one correspondence between reproducing TCFs and reproducing the properties of long time diffusion. However, discrepancies in the local free-energy barriers prevent accurate modeling of activated diffusion of small molecules in polymer melts.

Overall, these investigations have demonstrated that DPD is an efficient framework for adding relatively simple dissipative forces to the CG model, based on the MZ formalism. Recent work by Español and coworkers [54,55] has extended this approach to “entropy-based” CG models, which incorporate the internal energy as an additional CG variable, allowing simulations of thermal transport via DPD with energy conservation [56]. In general, in cases with clear time scale separation between the CG and removed degrees of freedom, the resulting DPD models can accurately represent

both local dynamics (e.g., velocity TCFs) as well as long time scale properties (e.g., diffusion constants). However, for higher-resolution CG representations, where this separation clearly does not exist, the resulting discrepancies can be rather large, depending on the precise CG mapping, functional form of the frictional forces, and the chosen conservative forces. As a result, researchers have pursued extensions of the DPD framework, which incorporate momentum- and time-dependent frictional forces to more accurately reproduce the underlying dynamics.

#### 2.4. Hydrodynamic-Like Momentum-Dependent Friction

Equation (6) demonstrates that the friction kernel is generally dependent not only on the coordinates but also the momenta of all CG sites. Although typically associated with movement of particles through a fluid, i.e., hydrodynamics, Soheilifard et al. [57] demonstrated that momentum-dependent frictional forces can emerge even when coarse-graining a linear model without solvent—in this case, an atomically-detailed elastic network protein model mapped to a C- $\alpha$  representation. To distinguish from the traditional case where effective interactions arise from correlations in solvent momentum, we refer to momentum-dependent friction that depends only on the CG degrees of freedom as “hydrodynamic-like”. In the realm of relatively coarse representations, Lei et al. [35] constructed DPD models with and without hydrodynamic-like contributions to the frictional force and demonstrated the necessity of momentum dependence for reproducing the flow about clusters of Lennard–Jones particles. They indicated that hydrodynamic-like forces will be required anytime the dissipative forces are on par with the conservative forces. They investigated two different pairwise conservative forces, one of which reproduces the radial distribution functions by construction, and demonstrated little difference in the resulting dynamics.

Markutsya and Lamm [30] proposed a method for computing hydrodynamic-like interactions directly from the distribution of random forces generated by higher-resolution simulations. They tested their method on both pure liquid water and a single glucose molecule solvated in water, using one-site representations for each molecule and the MS-CG method to determine the conservative potentials between CG sites. The viscosities, velocity–velocity TCFs, and self-diffusion constants were all reproduced with high accuracy relative to a CG model without dissipative forces, although there were notable structural discrepancies due to the employed conservative forces. However, because the model was not directly compared with a DPD model without hydrodynamic-like interactions, it is unclear what precise role these forces play. More recently, Izvekov [32,33] investigated the impact of momentum dependence on the resulting dynamical properties of a CG model by considering the approach to the Markovian limit and the regime of weak momentum dependence. He demonstrated that for a one-site model of liquid nitromethane the momentum dependence results in a much more accurate representation of the velocity TCFs, compared with the standard MZ-DPD model [21]. Moreover, the alternative projection operator [23] employed in this work may be better suited for high-resolution CG models that lack clear time scale separation, since it can deal with non-Markovian dynamics and avoids the plateau problem when using the Green–Kubo relation to determine the frictional forces from unconstrained molecular dynamics simulations.

At this point, we briefly cover more traditional modeling of hydrodynamic interactions. In the context of the GLE approaches described above, incorporating hydrodynamics amounts to reintroducing missing degrees of freedom (e.g., solvent) in an explicit but approximate way in order to more accurately model the movement of CG sites through an implicit medium. Zgorski and Lyman [58] employed stochastic rotational dynamics, also known as multi-particle collision dynamics, along with the implicit solvent Dry Martini model to more accurately model the thermodynamics and dynamics of a lipid bilayer. Stochastic rotational dynamics [59,60] is a particle-based Navier–Stokes fluid simulation technique that represents the solvent with tracer particles that conserve the hydrodynamic flow. The authors demonstrated that, with the appropriate choice of collision rules, the approach effectively corrects for known dynamical finite-size effects while generating the quasi-two-dimensional hydrodynamics of the bilayer. The lattice Boltzmann method [24]—an alternative hydrodynamics



approach which employs a lattice in lieu of tracer particles to represent the solvent—has been recently applied to study peptide self-assembly [61] and protein stability in crowded environments [62] using a chemically-specific CG protein model. These studies aim to probe the role of hydrodynamic interactions in these phenomena, rather than to directly match the dynamics of some underlying reference model. Researchers have also employed concurrent multiscale schemes that incorporate hydrodynamic interactions to model flow in liquids and at interfaces [63,64].

### 2.5. Friction Kernels with Memory

While the Markovian assumption is typically taken for practical convenience, researchers have also investigated memory effects, i.e. time dependence, in the friction kernel,  $\Gamma$ . Note that, through the fluctuation–dissipation theorem, this time dependence in  $\Gamma$  results in a time-dependent random force, which in practice requires implementation of a colored-noise generator. Cao et al. [65] demonstrated that, for modeling an ionic solution with implicit solvent, an exponential memory function is necessary for reproducing local dynamical properties, e.g., velocity TCFs, of the ions. Yoshimoto et al. [36] later showed that, due to a lack of time scale separation between CG and removed degrees of freedom, a non-Markovian friction kernel is necessary for reproducing viscosity in the high-density regime for a model that represents clusters of Lennard–Jones particles with a single CG site. Using an alternative projection operator, Izvekov [23] derived an expression for the memory as a function of force–force, force–position, and force–momentum TCFs. He showed that, if the projected force is decorrelated from both the CG positions and momenta, a position- and momentum-independent memory kernel is sufficient for representing the underlying dynamics. The work from Li et al. [38] demonstrated that, in the case of limited time scale separation, which can be quantified from the difference in relaxation time scales of velocity versus force TCFs, a time-correlated random force is necessary to reproduce short-time properties that are related to how the system responds to high-frequency disturbances. Additionally, the authors proposed two different methods for efficient implementation of a colored-noise generator. Very recently, Jung et al. [66,67] proposed a generalization of the approach of Li et al. [38], which can be interpreted as generalized Brownian dynamics with frequency-dependent friction kernels. In contrast to the previous approach, this method is not Galilean invariant and, thus, more applicable to the movement of CG particles in an implicit environment. The method was tested on nanocolloid suspensions, where it was demonstrated that the incorporation of time dependence is necessary for precisely reproducing the long-range hydrodynamic interactions between two nanocolloids. Moreover, the proposed methodology demonstrates an effective speed-up on the order of  $10^4$  for this particular application.

An alternative to the standard colored-noise generators was proposed by Andersen and coworkers [31,68], which uses a class of nonlinear systems interacting with special heat baths, originally introduced by Zwanzig. Within this scheme, fictitious particles are introduced which have no impact on the static properties of the system, but which couple directly to the CG degrees of freedom. The authors proposed the method as a simple scheme for reproducing dynamical features of an underlying model, without making a rigorous connection to the MZ framework. The approach was tested on several systems including a one-site model for ethanol dissolved in water and one- and two-site models for liquid methanol. The resulting models displayed excellent correspondence with the TCFs of the reference simulations. The discrepancies which did appear in these cases could be corrected by increasing the complexity of the fictitious particles. Lei et al. [69] subsequently demonstrated that a fictitious particle approach can be more rigorously connected to the GLE representation through a rational function approximation of the Laplace transform of the memory kernel. More recently, Karniadakis and coworkers [39,70] pursued a similar auxiliary variable approach, to simplify the implementation of non-Markovian DPD. Although these studies consider models with coarser representations (one site for a cluster of Lennard–Jones particles or an entire star polymer), the results contain potentially useful lessons for coarse-graining at a higher-resolution.

In comparison to the previous non-Markovian DPD approach from the same group [38], the auxiliary variable approach can reduce the computational cost by a factor of  $\sim 20$ – $30$ .

A different approach to simplifying the parameterization and implementation of the memory kernel was proposed by Lyubimov and Guenza [71], who employed an approximate but analytic expression for the time-dependent friction, derived from atomistic dynamics. This work is described further in Section 3.1.

## 2.6. Variational Approaches

Within the MZ approaches described above, the accuracy of the resulting dynamical properties of the model is dependent on how well the dissipative forces approximate those determined from MZ theory. The precise form of these dissipative forces is implicitly dependent on both the CG representation and the conservative forces, chosen to approximate the MB-PMF. Fu et al. [72] showed that the accurate representation of TCFs can be very sensitively dependent on fine features of the dissipative forces, e.g., the cut-off value employed. While the MZ formalism is rigorously founded, it does not provide in itself a systematic route for targeted optimization of particular properties, in contrast to variational approaches developed for determination of the conservative forces [19,73]. In the following, we describe two different variational approaches that have been proposed for constructing GLE models from higher-resolution reference simulations.

Español and Zúñiga [74] formulated a generalization of the relative entropy method—a variational approach for determining the conservative forces for a CG model [73]. The approach employs a Markovian approximation to the Fokker–Planck equation, which is expressed in terms of the first two Kramers–Moyal coefficients, i.e., a drift vector and a diffusion tensor. These terms can be estimated from sets of reference simulation trajectories with distinct initial states. In particular, the authors employed a new relative entropy functional which depends on two-time stationary joint probabilities, in contrast to the time-independent probabilities of the original relative entropy method. They also demonstrated that the same expression can be derived starting with the path probability and following a Bayesian approach, along with the appropriate (e.g., Markovian) approximations. The variational property of the method was demonstrated by showing that the Kramers–Moyal coefficients are recovered in the limit of a full basis set—i.e., arbitrarily complex forms to represent frictional forces. For coarser representations which demonstrate inertial dynamics, the inverse of the diffusion tensor may not exist, yielding an ill-posed optimization problem. However, using the Bayesian formulation, a generalized force-matching approach can be derived for obtaining both conservative and dissipative forces. This method essentially weights individual configurations within the force-matching functional by taking into account the relative contributions from thermal noise.

Inspired by this approach, Dequidt and Canchaya [75] devised a numerical scheme for parameterizing a DPD model with pairwise interactions. Using a Bayesian formulation, the authors derived an optimization function, which can be analytically solved for models with limited complexity. The resulting method is again a generalized force-matching (later termed as “trajectory matching”), where a least squares fit of the average force (weighted by a frictional contribution) over a short trajectory is performed. The method can be extended to simultaneously target the average pressure of the system, placing a restriction on the form of the conservative potential. The approach was tested on a one-site model of *n*-pentane, allowing direct comparison with the work of Trément et al. [50], who applied the Green–Kubo relation (Equation (8)) to determine the dissipative forces. The static properties of the resulting model accurately represent that of the reference simulation, not only for the pure liquid but also for liquid–vapor phase equilibria, when employing reference data from multiple temperatures [76]. However, significant discrepancies appear in the dynamical properties. In particular, the diffusion constant and TCFs demonstrate similar errors to the model of Trément et al. [50], indicating a limitation in the functional form of the frictional forces rather than a limitation of the optimization procedure. Similar results were demonstrated by Kempfer et al. [77] applying the same methodology to a polymer melt, highlighting the difficulty of simultaneously matching both short and

long time scale properties in more complex systems. Considering these discrepancies, extensions to more flexible basis functions for the frictional forces will help to assess the capabilities of the approach.

Harmandaris and coworkers [78,79] recently formulated an elegant theoretic framework which expands on the variational approaches for equilibrium coarse-graining [19,73]. In addition to generalizing the force-matching (i.e., MS-CG) method to nonlinear mappings and recasting structure-based methods in a probabilistic framework [78], the authors proposed a variational inference approach for the non-equilibrium regime [79]. This approach avoids the time discretization of the previous methods by employing the “relative entropy rate”—a full path-space entropy functional for continuous-time representations [80]. This work provides a detailed investigation into the proposed framework, elucidating the connections between force-matching-based and relative-entropy-based non-equilibrium methods as well as equivalency conditions for a completely data-driven approach based on generic time-series data. To the best of our knowledge, the method has not yet been tested on molecular simulation models.

### 2.7. Application to Proteins

Guenza and coworkers developed the Langevin equation for protein dynamics (LE4PD) method [81] by applying the projector operator technique to describe the system dynamics in terms of the  $\alpha$ -carbons along the protein backbone. The resulting model represents protein fluctuation dynamics using an overdamped Langevin equation along with structural information about the important metastable configurations of the protein as input. An extended Rouse–Zimm model is employed at the residue level, which specifically accounts for the difference in hydrodynamic interactions depending on the relative exposure to the solvent. In contrast to the standard Rouse–Zimm model, LE4PD allows for global anisotropy and sequence specificity while incorporating internal dissipation and free-energy barriers in a physically-motivated approach. LE4PD takes as input either short molecular dynamics simulations or experimental conformal structures and is capable of accurately representing both global and local motions of folded protein structures without any free parameters through a mode analysis approach. Through multiple investigations considering nine different proteins [81–83], the authors demonstrated that LE4PD is capable of reproducing NMR relaxation time scales as well as nuclear Overhauser effects. The LE4PD model has also recently been employed to identify universal hierarchical scaling features underlying sequence-specific protein dynamics, resembling the directed polymer in random media model [84].

## 3. Time Rescaling Relationships

If there exists a true time scale separation between the CG and removed degrees of freedom, then the time dependence of the frictional forces will vanish. In this case, there are two factors which determine the complexity of the frictional forces. First, the chosen conservative force modulates the form of the dissipative forces, although this relationship is not well understood. Additionally, the chosen projection operator within the MZ formalism determines the partitioning between the frictional and random forces [23]. Thus, with the appropriate representation, conservative interactions, and projection operator it may be possible to remove the coordinate, momentum, and time dependence from the friction kernel, resulting in a simple Langevin description for the CG dynamical evolution. In this case, a uniform time rescaling would recover all relevant dynamical properties at the CG level of resolution. Note however that, even in this case, the scalar friction term generally depends on the thermodynamic state point, e.g., temperature, pressure, or density.

### 3.1. Polymers

Many features of polymer systems can be understood using simple theoretical models which idealize monomer–monomer and monomer–solvent interactions. Amongst the simplest is the Rouse model [1], which represents segments of the polymer (i.e., one or several monomers) as CG sites connected along the chain via harmonic springs. These segments are assumed to move independently

from one another and, in the simplest case, the sites do not interact. This model is appropriate for chain lengths shorter than the entanglement length, beyond which more sophisticated reptation models are required. Extensions such as hydrodynamic-like interactions between segments to better model chain dynamics in dilute solution conditions can also be applied [1]. These models act as an important reference point from which “real” polymer dynamics can be compared. For example, in the absence of entanglement, divergence from the Rouse model implies significant correlations between segments that can arise from chain end effects, chain stiffness or semiflexibility, coupling between intra- and inter-molecular forces, or the presence of strong specific interactions [85,86]. Each of these effects may lead to non-trivial frictional forces from the MZ perspective. Although homogeneous polymeric systems exhibit static and dynamic universality beyond characteristic length and time scales [87], which can be reproduced with generic bead–spring models [88], chemically-specific molecular simulation models offer advantages for making direct comparisons with experimental measurements and opportunities for extensions to more complex, e.g., heterogeneous, systems.

Harmandaris, van der Vegt, Kremer, and coworkers performed detailed investigations into the dynamical properties generated from CG models for polystyrene-based systems with a two-site per monomer representation [89–97]. The conservative forces were parameterized to reproduce local structural features of a higher-resolution model. The authors assumed a scalar time-rescaling factor, which may depend on the molecule identity as well as the thermodynamic state point, determined by matching self-diffusion coefficients to those obtained from higher-resolution simulations. Although the rescaling factor depends on the molecular weight (i.e., chain length), it converges for relatively short chains, allowing the quantitative prediction of dynamics for system sizes intractable for the underlying reference model. The rescaled CG dynamics show favorable comparison with dynamic structure factors from neutron spin echo spectroscopy, rotation relaxation time scales from NMR, and dielectric spectroscopy. This simple rescaling approach is accurate on length scales of 0.5–1 nm and time scales of approximately 100 ps, allowing the description of long time diffusion as well as local chain dynamics. By considering distinct CG representations, the authors demonstrated that the precise rescaling factor is a sensitive function of the CG mapping, although the overall accuracy of the rescaling was similar for the different representations. By employing an empirically-determined Arrhenius-like time-rescaling factor as a function of concentration and temperature, they also extended the method to describe penetrant diffusion in polystyrene melts. However, the rescaling of individual components may differ, complicating the methodology in cases without a clear time scale separation between the dynamics of distinct components. The time rescalings are largely transferable to higher pressures, although discrepancies emerge in local side-chain dynamics. Remarkably, the rescaling approach results in very good agreement with respect to dielectric spectroscopy experiments that probe segmental dynamics of polystyrene/oligostyrene blends which display heterogeneous dynamics [98]. More recently, the approach has been extended to the nonequilibrium regime (e.g., polymers under shear flow), although difficulties arise far from equilibrium due to the entropy–enthalpy compensation in CG models. The reader is directed to [13,97] for a more detailed overview of time rescaling in CG models for polymers in the equilibrium and nonequilibrium regimes, respectively.

Salerno and Grest [99] performed a systematic investigation of the time-rescaling dependence on resolution and chain length for relatively high-resolution representations. They found that the time rescaling is independent of chain length (as demonstrated earlier by Harmandaris et al. [91]) but can depend sensitively on the CG representation. Ohkuma and Kremer [100] demonstrated that both the stress relaxation time scale and shear rate is recovered through time rescaling for two different models of *cis*-polyisoprene, with and without the incorporation of a pressure correction to the conservative force. More recently, Volgin et al. [101] investigated time rescalings for nanoparticle diffusion in a polyimide-based melt employing both one- and three-site per monomer representations. Interestingly, they found that, while the uniform time rescaling breaks down for intermediate time scales, taking into account the local viscosity of the polymer chain segments (whose size is comparable with the nanoparticle size) rescues the time-rescaling approach.

The time-rescaling approach suffers from several limitations. First, there is no guarantee that a scalar time-rescaling factor exists. For example, Agrawal et al. [102] showed that time-dependent rescaling factors are necessary for accurate prediction of polymer viscoelasticity for polyurea. Additionally, the time rescaling will generally be a function of the system identity, thermodynamic state point, and the chosen CG representation. For example, the time-rescaling behavior in glass-forming polymers can be related to the entropy–enthalpy compensation in CG models [103]. Lyubimov and Guenza [71,104] addressed this issue through an analytic rescaling method which connects monomer-level dynamics with global diffusion of a polymer chain. Using this approach, they considered both a monomer-level representation, i.e., a bead–spring model, and a soft-sphere representation which treats the entire polymer chain with a single CG site. The radius of gyration of the polymer as well as a hard-sphere parameter, which describes the excluded volume of monomer units, are taken as input. Using a clever combination of MZ theory, the Ornstein-Zernike equation from liquid state theory, and a model for cooperative polymer chain dynamics, the authors were able to reconstruct the local dynamics of a polymer chain that is consistent with the overall diffusive motion of the polymer. They demonstrated that for polyethylene and polybutadiene this approach is capable of reproducing the dynamic structure factors generated from united-atom molecular dynamics simulations. Moreover, they showed that the input parameters can be approximated using a freely-rotating chain model with only a small loss in accuracy as the system approaches the highly-entangled regime.

### 3.2. Liquids

There is a long history of work relating structural or thermodynamic properties with the emergent dynamical properties of liquids [105–107]. For example, free volume theories relate the available space for molecular packing with dynamic coefficients for transport properties [108,109]. Perhaps most notably, Rosenfeld demonstrated that a range of transport properties (diffusivity, viscosity, and thermal conductivity) are directly related to the excess entropy of the liquid [12]. These relationships have been validated for a wide class of liquids [110–114] but typically only appear when the appropriate transformations to reduced thermodynamic variables are applied.

Armstrong and Ballone [115] proposed that the Rosenfeld relationship could be utilized to interpret CG diffusion time scales in liquids. They demonstrated the linear relationship between the excess entropy and the diffusion constant for a one-site model of water. While the results demonstrate the potential to predict CG time scales, they also highlight the potential limitations of simple rescaling approaches for CG dynamics. If distinct chemical groups have different excess entropies, then a single rescaling is not possible, and consistent dynamics cannot be easily extracted. One predominant reasoning for the existence of thermodynamic–kinetic relationships in liquids is that all liquids can be considered to be perturbations of the venerable soft-sphere model [116]. Shell [117] leveraged this perspective to optimize CG interaction potentials based on a minimization of the relative entropy with respect to a higher-resolution model. Using this approach, he obtained the optimal soft-sphere scaling exponents, which in turn determine the time-rescaling factors for a set of binary Lennard–Jones systems. This rescaling resulted in a collapse of self-diffusivities onto a master curve for a range of temperatures, although the relationship breaks in the supercooled regime. We note that Accary and Teboul [118] showed that time rescalings for CG models of low-temperature liquids may be limited to correcting long-time diffusion at the cost of accurately representing the mechanism of local diffusion. Velhorst et al. [119] investigated the extension of the Rosenfeld relation to larger, flexible molecules, using Lennard–Jones particles connected by harmonic springs. They demonstrated that the harmonic bonds destroy correlations between the potential energy and the virial. These correlations give rise to the typical structural–kinetic relationships through the generation of isomorphs on the phase diagram—curves along which structure, dynamics, and excess entropy are invariant. In this case, isomorphs still exist, but the trend in excess entropy deviates from the other structural and dynamical properties. The authors also made a connection to polymers, where the well known power law scaling with density indicates that isomorph theory may be applicable.

Recently, Douglas and coworkers [103,120,121] proposed a different method which determines temperature-dependent CG potentials by matching the Debye–Waller factor, which can be related to the relaxation time scales in the localization model of glass forming liquids. In other words, the approach provides a calibration of the activation energy of glass formation via the cohesive interaction strength parameter in the CG potentials. The constructed three-site CG models for *ortho*-terphenyl reproduced the glass transition temperature, relaxation dynamics, and self-diffusion coefficients of the underlying AA model, as well as the reduced compressibilities [120]. The method was also applied to one- and two-site-per-monomer models of three glass-forming polymers with distinct segmental structures, where the form of the interaction potentials were extended to be frequency dependent while taking dynamic fragility into account in order to accurately represent the viscoelastic properties of the polymers under a small-amplitude oscillatory shear force [103,121]. The results demonstrate very good agreement with both AA simulations and experiments, although the authors identify that differences in the CG representation (i.e., level of resolution) can complicate the comparative analysis of CG models for distinct systems. Additionally, the temperature-dependent energy rescaling factors obtained for the three polymers, as well as for an analytically-solvable model, appear to exhibit a universal sigmoidal structure [121]. Overall, this work suggests that energy renormalization may provide a useful route for constructing CG models of glass-forming liquids and polymers with predictive capabilities.

#### 4. Free-Energy-Landscape (FEL) Perspective

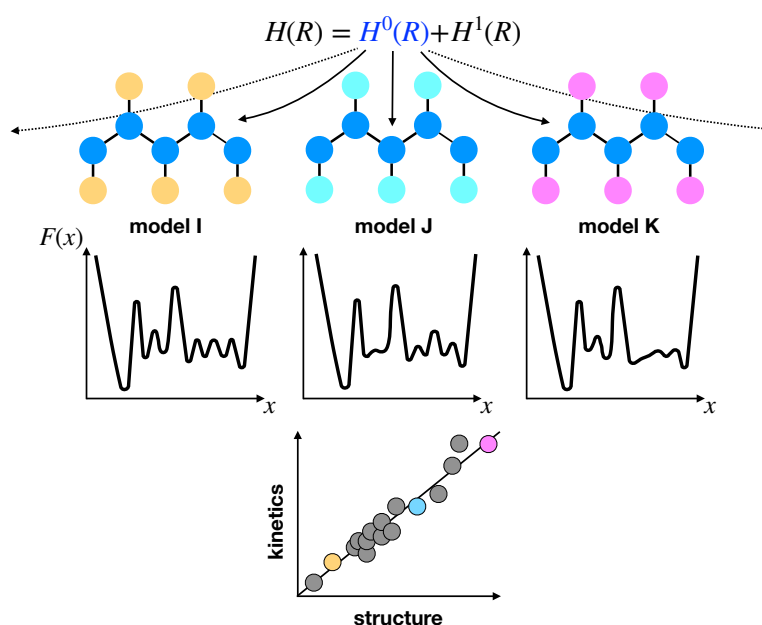
Despite the dynamical limitations of CG models, there is extensive research which utilizes these models to gain insight into the mechanisms of structure formation in soft matter. In particular, the energy-landscape formalism [122,123] provides the conceptual footing to interpret CG dynamics in a meaningful way based on metastable basins and the dominant set of barriers between these states. Perhaps most notable is the application of energy-landscape theory for protein folding [124], which asserts that the energy landscape for globular proteins is funnel shaped, with the native state as the global minimum. By constructing simple models based on this zeroth-order approximation, much of our foundational knowledge about the dominant driving forces of the protein folding process was obtained [3,125]. Although absolute time scales cannot be directly recovered from these models, connections between structural and dynamical properties arise through the restricted topology of the energy landscape [126], resulting in increased confidence in the extraction of insight from simulated pathways. Unlike the approaches described above, this perspective places a focus on the barrier-crossing dynamics along the dominant features of the FEL, while allowing for the diffusive motion within local free-energy minimum to differ from the true dynamics.

##### 4.1. Structural–Kinetic–Thermodynamic Relationships

The identification of structure-activity relationships has been widely adopted as a successful route for materials design [127]. The implicit assumption of this approach is that differences in the chemical features of distinct systems give rise to dominant distinctions with respect to the FEL, which then dictate larger-scale properties of the system. This link to the FEL qualifies the approach as a general means for understanding relationships between chemistry, structure, thermodynamics, and dynamics. More specifically, characteristic chemical or physical properties of a system or process enforce constraints on particular features of the FEL, such that simple relationships between structural, thermodynamic, and dynamical properties may emerge [128–131] (see Figure 2). In the context of CG modeling, these structural–kinetic or thermodynamic–kinetic relationships are somewhat similar in principle to the time-scaling relationships described above, as they provide a handle for quantitatively interpreting the dynamical properties generated by the CG model. Thus, such relationships might be useful for determining effective scaling relationships for building kinetic consistency into CG models, without the use of dissipative forces.

Inspired by high-throughput studies, Rudzinski and Berau [132] recently performed a detailed investigation into structural–kinetic–thermodynamic relationships for helix–coil transitions.

They employed a simple CG model which represents hydrogen-bonding and hydrophobic side chain interactions with simple attractive potentials, while maintaining an atomically-detailed description of excluded volume to avoid sampling of sterically-forbidden conformations. The authors demonstrated that steric effects alone enforce a rather strict restriction of the attainable FELs, characterized through a kinetic network description. This restriction gives rise to very simple relationships between structure and kinetics, e.g., between the average fraction of helical segments and the ratio of nucleation to elongation rates, irrespective of the specific details of the model. The results further demonstrate a link between cooperativity (i.e., thermodynamic properties) and the topology of the kinetic network (e.g., randomness of pathways from the helix to the coil state) at a single temperature. In a follow-up study [133], the same authors investigated the role of conformational entropy in these structural–kinetic relationships and found that a shift in the relationships occurs depending on the representation of steric interactions.



**Figure 2.** Illustration of a high-throughput-inspired investigation to identify structural–kinetic relationships. First, a Hamiltonian is defined with fixed ( $H^0$ ) and variable ( $H^1$ ) terms. Various models are then constructed by varying the parameters of  $H^1$ . The resulting FELs are then investigated (e.g., using kinetic networks as a proxy for the features of the landscape) to identify features that are preserved due to the physical components that are present in every model (i.e.,  $H^0$ ). Finally, structural–kinetic relationships are identified by finding correlations between various properties.

#### 4.2. Markov State Models (MSMs)

MSMs are discrete-time and -space master equation [134] models that can be parameterized directly from molecular simulation trajectories [135]. The problem of constructing an MSM is akin to finding a finite-dimensional, linear approximation to the dynamical propagator of the probability density of the system, which is intimately related to the Koopman operator approach from dynamical systems [136]. MSMs describe the dynamics of the system as memoryless jumps between *microstates* which represent volume elements of configuration space (i.e., sets of entire configurations of the system), in contrast to the particle-based CG models discussed above which explicitly represent the details of individual molecules. The microstate definition is crucial for the accuracy of the model, and is routinely determined in a semi-automated fashion using a combination of dimensionality reduction [137–143] and clustering [144–146] methodologies. Given the configuration space discretization, the transition probabilities between microstates (i.e., the MSM parameters) can be inferred from the number of jumps between pairs of microstates observed in the simulation trajectory within a chosen lag time [147–149]. Noé and coworkers introduced the variational approach to conformational dynamics [150,151], which

generalizes this procedure to arbitrary basis function representations—beyond the indicator functions used for traditional MSMs with discrete-state representations. More specifically, the variational scheme aims to directly approximate the eigenfunctions of the exact dynamical propagator [136]. Many other developments including cross-validation techniques [152] and the incorporation of machine learning algorithms into the kinetic modeling workflow [153] have contributed to the rapid emergence of these kinetic models as powerful tools for extracting insight from molecular simulations, particularly popular within the atomistic simulation community. (The reader is directed to a recent review by Husic and Pande [154] for a detailed account of MSM methodologies and applications.) Importantly, MSMs provide a link between microscopic details of a system and the hierarchy of long time scale processes. Although MSMs are CG models in their own right (representing an explicit coarse-graining in both space and time), the following considers the use of these models as a tool for assisting in the parameterization or interpretation of particle-based CG models.

Rudzinski and Bereau [155] investigated the refinement of two distinct CG peptide models to better reproduce the kinetic properties of given reference models, by altering the conservative force field to better represent the dominant barrier-crossing dynamics. Their approach relies on an MSM-based Bayesian dynamical reweighting technique [156] which determines the minimal adjustments to the kinetic model necessary for consistency with given kinetic reference data. By comparing this “biased” MSM with the MSM which most accurately represents the simulation data alone, they demonstrated that deficiencies of the CG model can be identified and subsequently targeted in a reparameterization. For example, for a C- $\alpha$ -based CG model of tetraalanine, the authors demonstrated that by allowing relatively small errors in the ensemble-averaged structural properties—the target property for the original parameterization—the refined model nearly quantitatively reproduces the ratio of mean first passage times between metastable states while also more accurately representing intermediate structures (i.e., the mechanism of transition between metastable states). These results demonstrate that the inaccurate representation of intermediate states—in this case, the sampling of sterically-forbidden states—destroys the connection between reproducing structural properties and reproducing the hierarchy of kinetic processes. Furthermore, this inaccuracy can be linked to the limited representation of cross correlations between degrees of freedom governing terms in the CG potential, due to the additivity of these terms in standard molecular mechanics Hamiltonians.

While higher-order interactions can be employed to reproduce these correlations in principle, parameterization and implementation of these interactions are often prohibitively expensive. Bereau and Rudzinski [157] recently proposed a scheme for reproducing complex cross correlations, without expanding the standard molecular mechanics basis set. Analogous to surface-hopping schemes for electronic transitions, the approach employs different CG force fields for distinct intramolecular conformations of a molecule and transitions between these force fields on the fly during the simulation. Remarkably, by applying the scheme to the C- $\alpha$  representation of tetraalanine described above, the authors demonstrated that accurate modeling of cross correlations leads to very high accuracy in the sampling of intermediate structures between metastable states (i.e., a better description of free-energy barriers), which in turn leads to the reproduction of barrier-crossing dynamics without adjusting the intra-basin friction (i.e., without the incorporation of dissipative forces).

Chen et al. [158] employed MSMs in conjunction with a framework for refining simulation models by integrating information from multiple experimental measurements in order to refine a CG model to more faithfully represent the folding mechanism of the Fip35 protein. To validate the proposed procedure, the investigation employed AA reference simulations as a proxy for experimental data. Through the consideration of one- and two-site per amino acid representations, the results highlight the importance of local frustration in reproducing the dominant features of the folding landscape. Very recently, Nüske et al. [159] proposed a “spectral matching” method, which attempts to parameterize a CG potential to accurately represent the long time scale processes by targeting the dominant eigenfunctions of the system’s dynamical propagator. This method utilizes the variational approach to conformational dynamics [150,151] along with methods for parameter estimation for



stochastic dynamics [160]. The method was demonstrated on a set of toy models as well as a CG model of alanine dipeptide which represents the system along the  $\phi/\psi$  dihedral angles. The results demonstrate that the method is capable of refining a force-matching-based model to more accurately reproduce the characteristic time scales of the underlying system. The proposed methodology appears to provide both a rigorous framework for the FEL perspective of CG dynamics presented in this section and also a practical alternative to the variational approaches described in Section 2.6.

## 5. Outstanding Challenges through Representative Examples

There are countless investigations using CG models which assess either quantitative dynamical properties or the qualitative pathways of structure formation generated by a particular model. Here, we discuss a few representative topics or studies which, in our opinion, highlight important aspects of the outstanding challenges in the context of interpreting and correcting CG dynamical properties.

As we have discussed in many examples above, coarse-graining of molecular liquids leads to disparate dynamical properties, including both short time scale TCFs and long time scale diffusion properties. Correcting both scales simultaneously can be challenging for both MZ and time-rescaling approaches, depending on the chemical features of the system. These difficulties are especially severe for ionic liquids, where strong non-covalent interactions lead to both structural and dynamic heterogeneity that is notoriously difficult to represent [161]. The presence of two species alone negates the use of traditional time-rescaling approaches, due to inhomogeneous rescaling for the different components [162]. Moreover, the simultaneous reproduction of local packing through the delicate balance of conservative forces while also correcting for lost friction presents a formidable challenge for equilibrium coarse-graining approaches and MZ-based methodologies. While certain dynamical properties may be universally conserved [163], the consistent representation of transport properties, essential for connection with real-world applications, remains difficult even for highly optimized models that accurately reproduce both structural and thermodynamic properties [164].

Mukherjee et al. [165] have illustrated a related but distinct example of challenging time rescaling for a CG model of azobenzene-based liquid crystals. The CG model was developed by applying traditional structure-based techniques to a reference simulation in the supercooled state. As a result, the model not only accurately reproduces structural properties of the amorphous phase, but also very closely matches the transition temperature from the amorphous to smectic phase with respect to the underlying AA model [166]. Both the AA and CG models display two mechanisms of translocation between layers in the smectic phase: (i) the straight mechanism, where a molecule jumps straight between two layers without reorientation; and (ii) the parking-lot mechanism, where a molecule rotates 90 degrees when exiting a layer and remains in the interlayer spacing for some period of time before entering the second layer. Despite the accurate ensemble-averaged properties, the CG model not only exhibits distinct time rescalings for various diffusion processes (each of the two translocation pathways, interlayer diffusion in the smectic phase, and diffusion in the amorphous phase) but also displays qualitatively incorrect propensities of the two translocation mechanisms. These two examples demonstrate fundamental challenges that will appear for modeling a wide range of industrially-relevant systems including heterogeneous polymeric systems [167] (e.g., polymer nanocomposites) as well as for investigating the pathways of molecular assembly processes, which may evolve along very rugged free-energy landscapes with predominant kinetic traps [168].

One might naively assume that the problem of interpreting CG dynamics would be necessarily more complicated for the conformational dynamics of macromolecules in solution compared with the diffusion processes of liquids. However, there are cases where chemically-specific CG models for biomolecules are capable of generating rates and pathways of structure formation that are consistent with experimental measurements. For example, the OxDNA model [169]—a two-site per nucleotide model for DNA parameterized to reproduce the melting temperatures of short duplexes—generates relative rates of duplex hybridization [170,171], strand displacement [172], and hairpin formation [173] that appear directly comparable to experiments. Apparently, the melting-curve-based parameterization

probes the dominant energetic driving forces (i.e., base pairing interactions) for a whole range of larger-scale structure formation. On the other hand, despite the success of simple CG protein models in clarifying the essential physics of the protein folding process, the difficulties of interpreting CG dynamical properties plague the characterization of sequence-specific folding, even with rather specialized CG models. For example, Habibi et al. [174] investigated the folding processes generated by three different CG models: a C- $\alpha$  G $\ddot{o}$  model [3], an AA G $\ddot{o}$  model [175], and the AWSEM model [176]. They demonstrated that the three different models provide disparate descriptions of the forced unfolding process of a 110 residue peptide, despite the capability of all models to fold the peptide to the proper native structure. Furthermore, the fidelity of the folding process, with respect to an AA reference simulation, did not correlate with the complexity of the model.

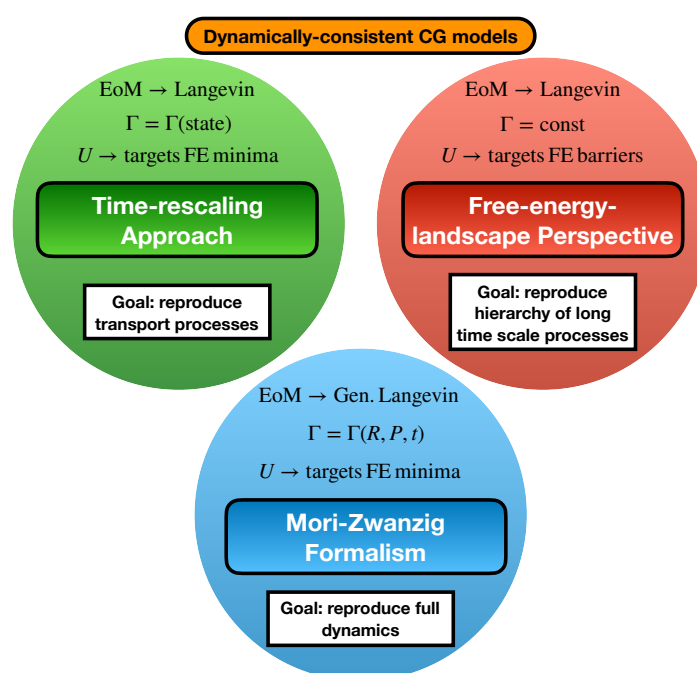
## 6. Discussion and Outlook

The last 5–10 years have seen a noticeable shift towards developing dynamically-consistent CG models, including: (i) the development of advanced Mori–Zwanzig-based methodologies for applications without clear time scale separation; (ii) the repurposing of well-known thermodynamic–kinetic relationships for time rescaling of diffusion processes in liquids; (iii) the identification of structural–kinetic–thermodynamic relationships for interpreting the conformational dynamics of macromolecules; and (iv) the utilization of kinetic models to characterize and correct the hierarchy of long time scale kinetic processes (see Figure 3 for a schematic overview of these approaches).

While the Mori–Zwanzig formalism has been traditionally applied to describe the motion of a system along one or a small set of coarse observables, it has now been established as a practical tool for extending the development of CG models beyond the parameterization of conservative forces. At the same time, there is still much to be learned about the appropriate complexity of configuration- and momentum-dependent frictional forces needed to accurately reproduce the dynamical properties for particular systems and CG representations. Additionally, systematic methods are required for choosing the corresponding set of hyperparameters (e.g., cut-off distances), which can dramatically affect the resulting short- and long-time dynamical behavior [72]. In the case that there is no clear time scale separation between the CG and removed degrees of freedom, time-dependent random forces are necessary to reproduce detailed features of the time correlation functions, although simpler models may be adequate for reproducing longer time scale properties [22]. Efficient parameterization and implementation of colored-noise thermostats, such as the use of fictitious beads coupled to the CG degrees of freedom [31], are of critical importance for alleviating the computational cost of these models. The development of variational approaches that target dynamical properties of the model represents a significant step towards not only more systematic investigations into the optimization of these models but also a better fundamental understanding of the required forces for reproducing particular properties. However, despite the development of several elegant theoretical frameworks [74,79,159], there seems to be a lack of applications to molecular systems. It is currently unclear whether this is due to practical difficulties of the approaches or simply a lack of adoption from external researchers. One limiting factor in successful applications of these approaches may be that there is a lack of methodology for systematically choosing the CG mapping that maximizes the time scale separation between the CG and removed degrees of freedom.

In addition to the optimization of the dissipative forces, a major goal moving forward should be to gain a better understanding of the interplay between the chosen CG representation and the conservative and dissipative forces required for reproducing particular dynamical properties. Although some studies [35] have shown that the conservative forces have little impact on resulting dynamical properties, others [22] indicate that the conservative and dissipative forces are more strongly coupled as the complexity of the system increases, e.g., through the introduction of distinct components. The CG mapping determines the idealized free-energy surface (i.e., the many-body potential of mean force) for representing the structural and thermodynamic properties of the underlying system. However, the set of chosen interactions between CG sites typically provide a rather poor approximation

to this complex landscape (Figure 1). In practice, one often attempts to represent the dominant features of this landscape by targeting certain lower-dimensional properties (e.g., pair correlation functions). The incorporation of dissipative forces is similar in the sense that, given the many-body potential of mean force as the CG conservative potential, there exists a perfect representation of the CG dynamics (i.e., the full generalized Langevin equation), although in practice we can only attempt to describe the dominant features of the “true” friction kernel. However, the conditional dependence on the CG conservative force complicates the construction of the appropriate set of dissipative forces. In other words, similar to the dependence of the optimal conservative force on the CG representation (i.e., mapping), the optimal dissipative forces depend on both the conservative force and the CG representation. There have been shockingly few investigations into the optimization of the CG mapping or conservative potential to reduce the complexity of the dissipative forces [177]. Therefore, future work should focus on the continued development of methods which couple the choice of the CG representation with the optimization of the conservative and dissipative forces.



**Figure 3.** Schematic of three distinct approaches to CG dynamics. **(Top left)** The time-rescaling approach attempts to reproduce long time scale target properties (e.g., the diffusion constant) of the underlying system. A simple Langevin equation motion (EoM) is applied, without the incorporation of dissipative forces. However, an effective scalar friction constant  $\Gamma$ , which can depend on the identity and thermodynamic state of the system, is employed a posteriori in order to calibrate the dynamics against the target property. The conservative potential is usually obtained via some standard coarse-graining approach, which aims to reproduce ensemble-averaged properties of the system (i.e., effectively targeting free-energy minima). **(Top right)** The free-energy-landscape perspective attempts to reproduce the dominant hierarchy of long time scale processes of the underlying system. A simple Langevin EoM is applied, without the incorporation of dissipative forces, which corresponds to the incorporation of a scalar friction coefficient  $\Gamma$ . The conservative potential is determined in order to not only reproduce features of the free-energy minima, but to also accurately represent the dominant barriers along the free-energy landscape. **(Bottom)** The Mori–Zwanzig approach attempts to reproduce the full dynamics of the underlying system by employing a generalized Langevin EoM to model the movement of CG sites. The corresponding friction kernel  $\Gamma$  generally depends on the coordinates and momenta of all CG sites, as well as on time. The conservative potential is usually obtained via some standard coarse-graining approach, which aims to reproduce ensemble-averaged properties of the system (i.e., effectively targeting free-energy minima).

Perhaps it is useful to look towards the free-energy-landscape perspective for inspiration, where researchers have devised heuristic methods for modeling the dominant kinetic features, dictated by the major barriers along the landscape. In comparison to the rigorous bottom-up methodologies, these approaches take advantage of large length and time scale behavior and appear to provide a greater range of transferability, not unlike the difference between bottom-up and top-down methods that target equilibrium properties [6]. The use of kinetic models (e.g., Markov state models) for characterizing the hierarchy of long time scale processes and associated pathways of structure formation in a tractable fashion seems to be an advantageous route forward. Additionally, the identification of structural–kinetic–thermodynamic relationships can significantly simplify the interpretation of the dynamical properties generated by CG models. While some relationships determine universal time-rescaling factors for a whole class of systems, others may be limited to a particular type of process or within a limited range of molecular identities. Nevertheless, these connections can provide confidence in the interpretation of properties away from the system or state point of parameterization, enhancing the predictive capabilities of CG models. Furthermore, these relationships (or lack thereof) inform the “representability” of the model, i.e., the capability of the model to simultaneously reproduce various structural, thermodynamic, and dynamical properties. In particular, the connection between reproducing various properties may be compromised by the limitations of the standard molecular mechanics functions employed to represent interactions between CG sites. The adoption of extended interaction sets or alternative models [157] that can more accurately represent cross correlations between CG degrees of freedom will generally simplify the problem of CG dynamics, regardless of the approach taken.

At the end of the day, bottom-up CG models rely on a higher-resolution reference model, which precludes predictive capabilities *unless* the variation of model parameters over system identity and thermodynamic state point is known a priori. The combination of bottom-up formalisms with analytic theories [71] as well as data-driven techniques [69] can assist in predicting or learning, respectively, these variations, propelling the utility of CG models. Overall, interfacing between the approaches discussed in this review (Figure 3) should lead to improved methods for interpreting and correcting the dynamical properties generated by CG simulation models, providing us not only with quantitatively interpretable models but also with confidence in the fidelity of predicted pathways.

**Funding:** This research received no external funding.

**Acknowledgments:** J.F.R. thanks Tristan Bereau, Will Noid, and David Rosenberger for critical review of the manuscript. J.F.R. is especially grateful to Tristan Bereau for many thoughtful discussions and guidance.

**Conflicts of Interest:** The author declares no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

CG	coarse-grained
MB-PMF	many-body potential of mean force
MZ	Mori–Zwanzig
AA	all-atom
dof	degree of freedom
GLE	generalized Langevin equation
DPD	dissipative particle dynamics
TCF	time correlation function
MS-CG	multiscale coarse-graining
LE4PD	Langevin equation for protein dynamics
FEL	free-energy landscape
MSM	Markov state model
EoM	equation of motion

## References

1. Doi, M.; Edwards, S.F. *The Theory of Polymer Dynamics*; Clarendon Press: Oxford, UK, 1986.
2. Taketomi, H.; Ueda, Y.; Gō, N. Studies on protein folding, unfolding and fluctuations by computer-simulation. 1. Effect of specific amino-acid sequence represented by specific inter-unit interactions. *Int. J. Pept. Protein Res.* **1975**, *7*, 445–459. [[CrossRef](#)] [[PubMed](#)]
3. Clementi, C. Coarse-grained models of protein folding: Toy models or predictive tools? *Curr. Opin. Struct. Biol.* **2008**, *18*, 10–15. [[CrossRef](#)] [[PubMed](#)]
4. Onsager, L. Theories of concentrated electrolytes. *Chem. Rev.* **1933**, *13*, 73–89. [[CrossRef](#)]
5. Kirkwood, J.G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313. [[CrossRef](#)]
6. Noid, W.G. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.* **2013**, *139*, 090901. [[CrossRef](#)] [[PubMed](#)]
7. Peter, C.; Kremer, K. Multiscale simulation of soft matter systems—From the atomistic to the coarse-grained level and back. *Soft Matter* **2009**, *5*, 4357–4366. [[CrossRef](#)]
8. Riniker, S.; Allison, J.R.; van Gunsteren, W.F. On developing coarse-grained models for biomolecular simulation: A review. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12423–12430. [[CrossRef](#)]
9. Mori, H. A continued-fraction representation of time-correlation functions. *Prog. Theor. Phys.* **1965**, *34*, 399–416. [[CrossRef](#)]
10. Zwanzig, R. Memory effects in irreversible thermodynamics. *Phys. Rev.* **1961**, *124*, 983–992. [[CrossRef](#)]
11. Hijón, C.; Español, P.; Vanden-Eijnden, E.; Delgado-Buscalioni, R. Mori–Zwanzig formalism as a practical computational tool. *Faraday Discuss.* **2010**, *144*, 301–322. [[CrossRef](#)]
12. Rosenfeld, Y. A quasi-universal scaling law for atomic transport in simple fluids. *J. Phys. Condens. Matter* **1999**, *11*, 5415–5427. [[CrossRef](#)]
13. Padding, J.T.; Briels, W.J. Systematic coarse-graining of the dynamics of entangled polymer melts: The road from chemistry to rheology. *J. Phys. Condens. Matter* **2011**, *23*, 233101. [[CrossRef](#)] [[PubMed](#)]
14. Zwanzig, R. On identity of three generalized master equations. *Physica* **1964**, *30*, 1109–1123. [[CrossRef](#)]
15. Robertson, B. Equations of motion in nonequilibrium statistical mechanics. *Phys. Rev.* **1966**, *144*, 151–161. [[CrossRef](#)]
16. Kawasaki, K.; Gunton, J. Theory of nonlinear transport processes: Nonlinear shear viscosity and normal stress effects. *Phys. Rev. A* **1973**, *8*, 2048–2064. [[CrossRef](#)]
17. Koide, T.; Kodama, T. Transport coefficients of non-Newtonian fluid and causal dissipative hydrodynamics. *Phys. Rev. E* **2008**, *78*, 051107. [[CrossRef](#)] [[PubMed](#)]
18. Li, P.; Dong, Y.; Zhao, N.; Hou, Z. Distance fluctuation of a single molecule in Lennard-Jones liquid based on generalized Langevin equation and mode coupling theory. *J. Chem. Phys.* **2014**, *140*, 154109. [[CrossRef](#)]
19. Noid, W.G.; Chu, J.W.; Ayton, G.S.; Krishna, V.; Izvekov, S.; Voth, G.A.; Das, A.; Andersen, H.C. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* **2008**, *128*, 244114. [[CrossRef](#)]
20. Kinjo, T.; Hyodo, S. Equation of motion for coarse-grained simulation based on microscopic description. *Phys. Rev. E* **2007**, *75*, 051109. [[CrossRef](#)]
21. Izvekov, S.; Rice, B.M. Multi-scale coarse-graining of non-conservative interactions in molecular liquids. *J. Chem. Phys.* **2014**, *140*, 104104. [[CrossRef](#)]
22. Deichmann, G.; van der Vegt, N.F.A. Bottom-up approach to represent dynamic properties in coarse-grained molecular simulations. *J. Chem. Phys.* **2018**, *149*, 244114. [[CrossRef](#)] [[PubMed](#)]
23. Izvekov, S. Microscopic derivation of particle-based coarse-grained dynamics. *J. Chem. Phys.* **2013**, *138*, 134106. [[CrossRef](#)] [[PubMed](#)]
24. Duenweg, B.; Schiller, U.D.; Ladd, A.J.C. Statistical mechanics of the fluctuating lattice Boltzmann equation. *Phys. Rev. E* **2007**, *76*, 036704. [[CrossRef](#)] [[PubMed](#)]
25. Español, P.; Warren, P.B. Perspective: Dissipative particle dynamics. *J. Chem. Phys.* **2017**, *146*, 150901. [[CrossRef](#)] [[PubMed](#)]
26. Junghans, C.; Praprotnik, M.; Kremer, K. Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat. *Soft Matter* **2008**, *4*, 156–161. [[CrossRef](#)]
27. Di Pasquale, N.; Hudson, T.; Icardi, M. Systematic derivation of hybrid coarse-grained models. *Phys. Rev. E* **2019**, *99*, 013303. [[CrossRef](#)] [[PubMed](#)]

28. Kauzlarić, D.; Español, P.; Greiner, A.; Succi, S. Three routes to the friction matrix and their application to the coarse-graining of atomic lattices. *Macromol. Theory Simul.* **2011**, *20*, 526–540. [[CrossRef](#)]
29. Izvekov, S.; Voth, G.A. Modeling real dynamics in the coarse-grained representation of condensed phase systems. *J. Chem. Phys.* **2006**, *125*, 151101. [[CrossRef](#)]
30. Markutsya, S.; Lamm, M.H. A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions. *J. Chem. Phys.* **2014**, *141*, 174107. [[CrossRef](#)]
31. Davtyan, A.; Dama, J.F.; Voth, G.A.; Andersen, H.C. Dynamic force matching: A method for constructing dynamical coarse-grained models with realistic time dependence. *J. Chem. Phys.* **2015**, *142*, 154104. [[CrossRef](#)]
32. Izvekov, S. Mori-Zwanzig theory for dissipative forces in coarse-grained dynamics in the Markov limit. *Phys. Rev. E* **2017**, *95*, 013303. [[CrossRef](#)] [[PubMed](#)]
33. Izvekov, S. Microscopic derivation of particle-based coarse-grained dynamics: Exact expression for memory function. *J. Chem. Phys.* **2017**, *146*, 124109. [[CrossRef](#)] [[PubMed](#)]
34. Español, P.; de la Torre, J.A.; Duque-Zumajo, D. Solution to the plateau problem in the Green-Kubo formula. *Phys. Rev. E* **2019**, *99*, 022126. [[CrossRef](#)] [[PubMed](#)]
35. Lei, H.; Caswell, B.; Karniadakis, G.E. Direct construction of mesoscopic models from microscopic simulations. *Phys. Rev. E* **2010**, *81*, 026704. [[CrossRef](#)] [[PubMed](#)]
36. Yoshimoto, Y.; Kinefuchi, I.; Mima, T.; Fukushima, A.; Tokumasu, T.; Takagi, S. Bottom-up construction of interaction models of non-Markovian dissipative particle dynamics. *Phys. Rev. E* **2013**, *88*, 043305. [[CrossRef](#)] [[PubMed](#)]
37. Li, Z.; Bian, X.; Caswell, B.; Karniadakis, G.E. Construction of dissipative particle dynamics models for complex fluids via the Mori-Zwanzig formulation. *Soft Matter* **2014**, *10*, 8659–8672. [[CrossRef](#)] [[PubMed](#)]
38. Li, Z.; Bian, X.; Li, X.; Karniadakis, G.E. Incorporation of memory effects in coarse-grained modeling via the Mori-Zwanzig formalism. *J. Chem. Phys.* **2015**, *143*, 243128. [[CrossRef](#)]
39. Li, Z.; Lee, H.S.; Darve, E.; Karniadakis, G.E. Computing the non-Markovian coarse-grained interactions derived from the Mori-Zwanzig formalism in molecular systems: Application to polymer melts. *J. Chem. Phys.* **2017**, *146*, 014104. [[CrossRef](#)] [[PubMed](#)]
40. Español, P.; Serrano, M.; Zúñiga, I. Coarse-graining of a fluid and its relation with dissipative particle dynamics and smoothed particle dynamics. *Int. J. Modern Phys. C* **1997**, *8*, 899–908. [[CrossRef](#)]
41. Flekkøy, E.; Coveney, P. From molecular dynamics to dissipative particle dynamics. *Phys. Rev. Lett.* **1999**, *83*, 1775–1778. [[CrossRef](#)]
42. Flekkøy, E.; Coveney, P.; De Fabritiis, G. Foundations of dissipative particle dynamics. *Phys. Rev. E* **2000**, *62*, 2140–2157. [[CrossRef](#)] [[PubMed](#)]
43. Ayton, G.; Tepper, H.; Mirijanian, D.; Voth, G. A new perspective on the coarse-grained dynamics of fluids. *J. Chem. Phys.* **2004**, *120*, 4074–4088. [[CrossRef](#)] [[PubMed](#)]
44. Eriksson, A.; Jacobi, M.N.; Nyström, J.; Tunström, K. Using force covariance to derive effective stochastic interactions in dissipative particle dynamics. *Phys. Rev. E* **2008**, *77*, 016707. [[CrossRef](#)] [[PubMed](#)]
45. Eriksson, A.; Jacobi, M.N.; Nyström, J.; Tunström, K. On the microscopic foundation of dissipative particle dynamics. *Europhys. Lett.* **2009**, *86*, 44001. [[CrossRef](#)]
46. Hadley, K.R.; McCabe, C. On the Investigation of Coarse-Grained Models for Water: Balancing Computational Efficiency and the Retention of Structural Properties. *J. Phys. Chem. B* **2010**, *114*, 4590–4599. [[CrossRef](#)] [[PubMed](#)]
47. Han, Y.; Dama, J.F.; Voth, G.A. Mesoscopic coarse-grained representations of fluids rigorously derived from atomistic models. *J. Chem. Phys.* **2018**, *149*, 044104. [[CrossRef](#)] [[PubMed](#)]
48. Bock, H.; Gubbins, K.E.; Klapp, S.H.L. Coarse graining of nonbonded degrees of freedom. *Phys. Rev. Lett.* **2007**, *98*, 267801. [[CrossRef](#)] [[PubMed](#)]
49. Izvekov, S.; Rice, B.M. On the importance of shear dissipative forces in coarse-grained dynamics of molecular liquids. *Phys. Chem. Chem. Phys.* **2015**, *17*, 10795–10804. [[CrossRef](#)]
50. Trément, S.; Schnell, B.; Petitjean, L.; Couty, M.; Rousseau, B. Conservative and dissipative force field for simulation of coarse-grained alkane molecules: A bottom-up approach. *J. Chem. Phys.* **2014**, *140*, 134113. [[CrossRef](#)]
51. Lemarchand, C.A.; Couty, M.; Rousseau, B. Coarse-grained simulations of cis- and trans-polybutadiene: A bottom-up approach. *J. Chem. Phys.* **2017**, *146*, 074904. [[CrossRef](#)]

52. Deichmann, G.; Marcon, V.; van der Vegt, N.F.A. Bottom-up derivation of conservative and dissipative interactions for coarse-grained molecular liquids with the conditional reversible work method. *J. Chem. Phys.* **2014**, *141*, 224109. [[CrossRef](#)] [[PubMed](#)]
53. Brini, E.; Marcon, V.; van der Vegt, N.F.A. Conditional reversible work method for molecular coarse graining applications. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10468–10474. [[CrossRef](#)] [[PubMed](#)]
54. Español, P.; Serrano, M.; Pagonabarraga, I.; Zúñiga, I. Energy-conserving coarse-graining of complex molecules. *Soft Matter* **2016**, *12*, 4821–4837. [[CrossRef](#)] [[PubMed](#)]
55. Faure, G.; Delgado-Buscalioni, R.; Español, P. The entropy of a complex molecule. *J. Chem. Phys.* **2017**, *146*, 224106. [[CrossRef](#)] [[PubMed](#)]
56. Español, P. Dissipative particle dynamics with energy conservation. *Europhys. Lett.* **1997**, *40*, 631–636. [[CrossRef](#)]
57. Soheilifard, R.; Makarov, D.E.; Rodin, G.J. Rigorous coarse-graining for the dynamics of linear systems with applications to relaxation dynamics in proteins. *J. Chem. Phys.* **2011**, *135*, 54107. [[CrossRef](#)] [[PubMed](#)]
58. Zgorski, A.; Lyman, E. Toward hydrodynamics with solvent free lipid models: STRD Martini. *Biophys. J.* **2016**, *111*, 2689–2697. [[CrossRef](#)] [[PubMed](#)]
59. Malevanets, A.; Kapral, R. Mesoscopic model for solvent dynamics. *J. Chem. Phys.* **1999**, *110*, 8605–8613. [[CrossRef](#)]
60. Gompper, G.; Ihle, T.; Kroll, D.M.; Winkler, R.G. Multi-Particle Collision Dynamics: A Particle-Based Mesoscale Simulation Approach to the Hydrodynamics of Complex Fluids. In *Advanced Computer Simulation Approaches for Soft Matter Sciences III*; Holm, C., Kremer, K., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–87.
61. Sterpone, F.; Derreumaux, P.; Melchionna, S. Protein simulations in fluids: Coupling the OPEP coarse-grained force field with hydrodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 1843–1853. [[CrossRef](#)]
62. Gnutt, D.; Timr, S.; Ahlers, J.; Ko, B.; Manderfeld, E.; Heyden, M.; Sterpone, F.; Ebbinghaus, S. Stability effect of quinary interactions reversed by single point mutations. *J. Am. Chem. Soc.* **2019**, *141*, 4660–4669. [[CrossRef](#)] [[PubMed](#)]
63. Hu, J.; Korotkin, I.A.; Karabasov, S.A. A multi-resolution particle/fluctuating hydrodynamics model for hybrid simulations of liquids based on the two-phase flow analogy. *J. Chem. Phys.* **2018**, *149*, 084108. [[CrossRef](#)]
64. Wang, Y.; Li, Z.; Xu, J.; Yang, C.; Karniadakis, G.E. Concurrent coupling of atomistic simulation and mesoscopic hydrodynamics for flows over soft multi-functional surfaces. *Soft Matter* **2018**, *15*, 1747–1757. [[CrossRef](#)] [[PubMed](#)]
65. Cao, Z.; Dama, J.F.; Lu, L.; Voth, G.A. Solvent free ionic solution models from multiscale coarse-graining. *J. Chem. Theory Comput.* **2013**, *9*, 172–178. [[CrossRef](#)] [[PubMed](#)]
66. Jung, G.; Hanke, M.; Schmid, F. Iterative reconstruction of memory kernels. *J. Chem. Theory Comput.* **2017**, *13*, 2481–2488. [[CrossRef](#)] [[PubMed](#)]
67. Jung, G.; Hanke, M.; Schmid, F. Generalized Langevin dynamics: Construction and numerical integration of non-Markovian particle-based models. *Soft Matter* **2018**, *14*, 9368–9382. [[CrossRef](#)] [[PubMed](#)]
68. Davtyan, A.; Voth, G.A.; Andersen, H.C. Dynamic force matching: Construction of dynamic coarse-grained models with realistic short time dynamics and accurate long time dynamics. *J. Chem. Phys.* **2016**, *145*, 224107. [[CrossRef](#)] [[PubMed](#)]
69. Lei, H.; Bakera, N.A.; Li, X. Data-driven parameterization of the generalized Langevin equation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14183–14188. [[CrossRef](#)]
70. Yoshimoto, Y.; Li, Z.; Kinefuchi, I.; Karniadakis, G.E. Construction of non-Markovian coarse-grained models employing the Mori-Zwanzig formalism and iterative Boltzmann inversion. *J. Chem. Phys.* **2017**, *147*, 244110. [[CrossRef](#)] [[PubMed](#)]
71. Lyubimov, I.; Guenza, M.G. First-principle approach to rescale the dynamics of simulated coarse-grained macromolecular liquids. *Phys. Rev. E* **2011**, *84*, 031801. [[CrossRef](#)]
72. Fu, C.C.; Kulkarni, P.M.; Shell, M.S.; Leal, L.G. A test of systematic coarse-graining of molecular dynamics simulations: Transport properties. *J. Chem. Phys.* **2013**, *139*, 094107. [[CrossRef](#)]
73. Shell, M.S. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J. Chem. Phys.* **2008**, *129*, 144108. [[CrossRef](#)] [[PubMed](#)]

74. Español, P.; Zúñiga, I. Obtaining fully dynamic coarse-grained models from MD. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10538–10545. [[CrossRef](#)] [[PubMed](#)]
75. Dequidt, A.; Canchaya, J.G.S. Bayesian parametrization of coarse-grain dissipative dynamics models. *J. Chem. Phys.* **2015**, *143*, 084122. [[CrossRef](#)] [[PubMed](#)]
76. Canchaya, J.G.S.; Dequidt, A.; Goujon, F.; Malfreyt, P. Development of DPD coarse-grained models: From bulk to interfacial properties. *J. Chem. Phys.* **2016**, *145*, 054107. [[CrossRef](#)] [[PubMed](#)]
77. Kempfer, K.; Devemy, J.; Dequidt, A.; Couty, M.; Malfreyt, P. Development of coarse-grained models for polymers by trajectory matching. *ACS Omega* **2019**, *4*, 5955–5967. [[CrossRef](#)]
78. Kalligiannaki, E.; Harmandaris, V.; Katsoulakis, M.A.; Plecháč, P. The geometry of generalized force matching and related information metrics in coarse-graining of molecular systems. *J. Chem. Phys.* **2015**, *143*, 084105. [[CrossRef](#)] [[PubMed](#)]
79. Harmandaris, V.; Kalligiannaki, E.; Katsoulakis, M.; Plecháč, P. Path-space variational inference for non-equilibrium coarse-grained systems. *J. Comput. Phys.* **2016**, *314*, 355–383. [[CrossRef](#)]
80. Katsoulakis, M.A.; Plechac, P. Information-theoretic tools for parametrized coarse-graining of non-equilibrium extended systems. *J. Chem. Phys.* **2013**, *139*, 074115. [[CrossRef](#)]
81. Copperman, J.; Guenza, M.G. Coarse-grained Langevin equation for protein dynamics: Global anisotropy and a mode approach to local complexity. *J. Phys. Chem. B* **2015**, *119*, 9195–9211. [[CrossRef](#)]
82. Caballero-Manrique, E.; Bray, J.K.; Deutschman, W.A.; Dahlquist, F.W.; Guenza, M.G. A theory of protein dynamics to predict NMR relaxation. *Biophys. J.* **2007**, *93*, 4128–4140. [[CrossRef](#)]
83. Copperman, J.; Guenza, M. Predicting protein dynamics from structural ensembles. *J. Chem. Phys.* **2015**, *143*, 243131. [[CrossRef](#)] [[PubMed](#)]
84. Copperman, J.; Dinpajoo, M.; Beyerle, E.R.; Guenza, M.G. Universality and specificity in protein fluctuation dynamics. *Phys. Rev. Lett.* **2017**, *119*, 158101. [[CrossRef](#)]
85. Guenza, M. Cooperative dynamics in unentangled polymer fluids. *Phys. Rev. Lett.* **2002**, *88*, 025901. [[CrossRef](#)] [[PubMed](#)]
86. Guenza, M. Theoretical models for bridging timescales in polymer dynamics. *J. Phys. Condens. Matter* **2008**, *20*, 033101. [[CrossRef](#)]
87. Takahashi, K.Z.; Nishimura, R.; Yamato, N.; Yasuoka, K.; Masubuchi, Y. Onset of static and dynamic universality among molecular models of polymers. *Sci. Rep.* **2017**, *7*, 12379. [[CrossRef](#)] [[PubMed](#)]
88. Kremer, K.; Grest, G. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **1990**, *92*, 5057–5086. [[CrossRef](#)]
89. Harmandaris, V.A.; Adhikari, N.P.; van der Vegt, N.F.A.; Kremer, K. Hierarchical modeling of polystyrene: From atomistic to coarse-grained simulations. *Macromolecules* **2006**, *39*, 6708–6719. [[CrossRef](#)]
90. Harmandaris, V.A.; Adhikari, N.P.; van der Vegt, N.F.A.; Kremer, K.; Mann, B.A.; Voelkel, R.; Weiss, H.; Liew, C.C. Ethylbenzene diffusion in polystyrene: United atom atomistic/coarse grained simulations and experiments. *Macromolecules* **2007**, *40*, 7026–7035. [[CrossRef](#)]
91. Harmandaris, V.A.; Reith, D.; van der Vegt, N.F.A.; Kremer, K. Comparison between coarse-graining models for polymer systems: Two mapping schemes for polystyrene. *Macromol. Chem. Phys.* **2007**, *208*, 2109–2120. [[CrossRef](#)]
92. Harmandaris, V.A.; Kremer, K. Predicting polymer dynamics at multiple length and time scales. *Soft Matter* **2009**, *5*, 3920–3926. [[CrossRef](#)]
93. Fritz, D.; Harmandaris, V.A.; Kremer, K.; van der Vegt, N.F.A. Coarse-grained polymer melts based on isolated atomistic chains: Simulation of polystyrene of different tacticities. *Macromolecules* **2009**, *42*, 7579–7588. [[CrossRef](#)]
94. Harmandaris, V.A.; Kremer, K. Dynamics of polystyrene melts through hierarchical multiscale simulations. *Macromolecules* **2009**, *42*, 791–802. [[CrossRef](#)]
95. Fritz, D.; Koschke, K.; Harmandaris, V.A.; van der Vegt, N.F.A.; Kremer, K. Multiscale modeling of soft matter: Scaling of dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10412–10420. [[CrossRef](#)] [[PubMed](#)]
96. Harmandaris, V.A.; Floudas, G.; Kremer, K. Temperature and pressure dependence of polystyrene dynamics through molecular dynamics simulations and experiments. *Macromolecules* **2011**, *44*, 393–402. [[CrossRef](#)]
97. Harmandaris, V.A. Quantitative study of equilibrium and non-equilibrium polymer dynamics through systematic hierarchical coarse-graining simulations. *Korea-Aust. Rheol. J.* **2014**, *26*, 15–28. [[CrossRef](#)]



98. Harmandaris, V.A.; Kremer, K.; Floudas, G. Dynamic heterogeneity in fully miscible blends of polystyrene with oligostyrene. *Phys. Rev. Lett.* **2013**, *110*, 165701. [[CrossRef](#)] [[PubMed](#)]
99. Salerno, K.; Agrawal, A.; Perahia, D.; Grest, G. Resolving dynamic properties of polymers through coarse-grained computational studies. *Phys. Rev. Lett.* **2016**, *116*, 058302. [[CrossRef](#)]
100. Ohkuma, T.; Kremer, K. Comparison of two coarse-grained models of cis-polyisoprene with and without pressure correction. *Polymer* **2017**, *130*, 88–101. [[CrossRef](#)]
101. Volgin, I.V.; Larin, S.V.; Lyulin, A.V.; Lyulin, S.V. Coarse-grained molecular-dynamics simulations of nanoparticle diffusion in polymer nanocomposites. *Polymer* **2018**, *145*, 80–87. [[CrossRef](#)]
102. Agrawal, V.; Holzworth, K.; Nantasetphong, W.; Amirkhizi, A.V.; Oswald, J.; Nemat-Nasser, S. Prediction of viscoelastic properties with coarse-grained molecular dynamics and experimental validation for a benchmark polyurea system. *J. Polym. Sci. B* **2016**, *54*, 797–810. [[CrossRef](#)]
103. Song, J.; Hsu, D.D.; Shull, K.R.; Phelan, F.R.; Douglas, J.F.; Xia, W.; Keten, S. Energy renormalization method for the coarse-graining of polymer viscoelasticity. *Macromolecules* **2018**, *51*, 3818–3827. [[CrossRef](#)] [[PubMed](#)]
104. Lyubimov, I.Y.; Guenza, M.G. Theoretical reconstruction of realistic dynamics of highly coarse-grained cis-1,4-polybutadiene melts. *J. Chem. Phys.* **2013**, *138*, 12A546. [[CrossRef](#)] [[PubMed](#)]
105. Cohen, M.; Turnbull, D. Molecular transport in liquids and glasses. *J. Chem. Phys.* **1959**, *31*, 1164–1169. [[CrossRef](#)]
106. Adam, G.; Gibbs, J. On temperature dependence of cooperative relaxation properties in glass-forming liquids. *J. Chem. Phys.* **1965**, *43*, 139–146. [[CrossRef](#)]
107. Rosenfeld, Y. Relation between transport-coefficients and internal entropy of simple systems. *Phys. Rev. A* **1977**, *15*, 2545–2549. [[CrossRef](#)]
108. Starr, F.; Sastry, S.; Douglas, J.; Glotzer, S. What do we learn from the local geometry of glass-forming liquids? *Phys. Rev. Lett.* **2002**, *89*, 125501. [[CrossRef](#)] [[PubMed](#)]
109. Mittal, J.; Errington, J.R.; Truskett, T.M. Relationships between self-diffusivity, packing fraction, and excess entropy in simple bulk and confined fluids. *J. Phys. Chem. B* **2007**, *111*, 10054–10063. [[CrossRef](#)]
110. Saika-Voivod, I.; Poole, P.; Sciortino, F. Fragile-to-strong transition and polyamorphism in the energy landscape of liquid silica. *Nature* **2001**, *412*, 514–517. [[CrossRef](#)]
111. Sastry, S. The relationship between fragility, configurational entropy and the potential energy landscape of glass-forming liquids. *Nature* **2001**, *409*, 164–167. [[CrossRef](#)]
112. Goel, G.; Krekelberg, W.P.; Pond, M.J.; Mittal, J.; Shen, V.K.; Errington, J.R.; Truskett, T.M. Available states and available space: Static properties that predict self-diffusivity of confined fluids. *J. Stat. Mech.* **2009**, *2009*, P04006. [[CrossRef](#)]
113. Krekelberg, W.P.; Pond, M.J.; Goel, G.; Shen, V.K.; Errington, J.R.; Truskett, T.M. Generalized Rosenfeld scalings for tracer diffusivities in not-so-simple fluids: Mixtures and soft particles. *Phys. Rev. E* **2009**, *80*, 061205. [[CrossRef](#)] [[PubMed](#)]
114. Dyre, J.C. Perspective: Excess-entropy scaling. *J. Chem. Phys.* **2018**, *149*, 210901. [[CrossRef](#)] [[PubMed](#)]
115. Armstrong, J.A.; Chakravarty, C.; Ballone, P. Statistical mechanics of coarse graining: Estimating dynamical speedups from excess entropies. *J. Chem. Phys.* **2012**, *136*, 124503. [[CrossRef](#)] [[PubMed](#)]
116. Likos, C. Effective interactions in soft condensed matter physics. *Phys. Rep.* **2001**, *348*, 267–439. [[CrossRef](#)]
117. Shell, M.S. Systematic coarse-graining of potential energy landscapes and dynamics in liquids. *J. Chem. Phys.* **2012**, *137*, 084503. [[CrossRef](#)]
118. Accary, J.; Teboul, V. Time versus temperature rescaling for coarse grain molecular dynamics simulations. *J. Chem. Phys.* **2012**, *136*, 094502. [[CrossRef](#)]
119. Veldhorst, A.A.; Dyre, J.C.; Schröder, T.B. Scaling of the dynamics of flexible Lennard-Jones chains. *J. Chem. Phys.* **2014**, *141*, 054904. [[CrossRef](#)]
120. Xia, W.; Song, J.; Hansoge, N.K.; Phelan, F.R.; Keten, S.; Douglas, J.F. Energy renormalization for coarse-graining the dynamics of a model glass-forming liquid. *J. Phys. Chem. B* **2018**, *122*, 2040–2045. [[CrossRef](#)]
121. Xia, W.; Hansoge, N.K.; Xu, W.S.; Phelan, F.R., Jr.; Keten, S.; Douglas, J.F. Energy renormalization for coarse-graining polymers having different segmental structures. *Sci. Adv.* **2019**, *5*, eaav4683. [[CrossRef](#)]
122. Stillinger, F.; Weber, T. Hidden structure in liquids. *Phys. Rev. A* **1982**, *25*, 978–989. [[CrossRef](#)]
123. Stillinger, F. A topographic view of supercooled liquids and glass-formation. *Science* **1995**, *267*, 1935–1939. [[CrossRef](#)] [[PubMed](#)]

124. Onuchic, J.; LutheySchulten, Z.; Wolynes, P. Theory of protein folding: The energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545–600. [[CrossRef](#)] [[PubMed](#)]
125. Dill, K.A.; Chan, H.S. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **1997**, *4*, 10–19. [[CrossRef](#)] [[PubMed](#)]
126. Leopold, P.E.; Montal, M.; Onuchic, J.N. Protein folding funnels: A kinetic approach to the sequence structure relationship. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8721–8725. [[CrossRef](#)]
127. Berau, T.; Andrienko, D.; Kremer, K. Research Update: Computational materials discovery in soft matter. *APL Mater.* **2016**, *4*, 053101. [[CrossRef](#)]
128. Schneider, E.V.; Böttcher, J.; Huber, R.; Maskos, K.; Neumann, L. Structure-kinetic relationship study of CDK8/CycC specific compounds. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8081–8086. [[CrossRef](#)] [[PubMed](#)]
129. Winquist, J.; Geschwindner, S.; Xue, Y.; Gustavsson, L.; Musil, D.; Deinum, J.; Danielson, U.H. Identification of structural-kinetic and structural-thermodynamic relationships for thrombin inhibitors. *Biochemistry* **2013**, *52*, 613–626. [[CrossRef](#)]
130. Dickson, C.J.; Hornak, V.; Pearlstein, R.A.; Duca, J.S. Structure-kinetic relationships of passive membrane permeation from multiscale modeling. *J. Am. Chem. Soc.* **2017**, *139*, 442–452. [[CrossRef](#)]
131. Menichetti, R.; Kanekal, K.H.; Kremer, K.; Berau, T. In silico screening of drug-membrane thermodynamics reveals linear relations between bulk partitioning and the potential of mean force. *J. Chem. Phys.* **2017**, *147*, 125101. [[CrossRef](#)]
132. Rudzinski, J.F.; Berau, T. The role of conformational entropy in the determination of structural-kinetic relationships for helix-coil transitions. *Computation* **2018**, *6*, 21. [[CrossRef](#)]
133. Rudzinski, J.F.; Berau, T. Structural-kinetic-thermodynamic relationships identified from physics-based molecular simulation models. *J. Chem. Phys.* **2018**, *148*, 204111. [[CrossRef](#)] [[PubMed](#)]
134. Klippenstein, S.J.; Pande, V.S.; Truhlar, D.G. Chemical kinetics and mechanisms of complex systems: A perspective on recent theoretical advances. *J. Am. Chem. Soc.* **2014**, *136*, 528–546. [[CrossRef](#)] [[PubMed](#)]
135. Bowman, G.R.; Pande, V.S.; Noé, F. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*; Springer Science and Business Media: Dordrecht, The Netherlands, 2014.
136. Klus, S.; Nuske, F.; Koltai, P.; Wu, H.; Kevrekidis, I.; Schuette, C.; Noe, F. Data-Driven Model Reduction and Transfer Operator Approximation. *J. Nonlinear Sci.* **2018**, *28*, 985–1010. [[CrossRef](#)]
137. Mu, Y.; Nguyen, P.; Stock, G. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. *Proteins Struct. Funct. Bioinform.* **2005**, *58*, 45–52. [[CrossRef](#)] [[PubMed](#)]
138. Naritomi, Y.; Fuchigami, S. Slow dynamics in protein fluctuations revealed by time-structure based independent component analysis: The case of domain motions. *J. Chem. Phys.* **2011**, *134*, 065101. [[CrossRef](#)] [[PubMed](#)]
139. Schwantes, C.R.; Pande, V.S. Improvements in Markov state model construction reveal many non-native interactions in the folding of NTL9. *J. Chem. Theory Comput.* **2013**, *9*, 2000–2009. [[CrossRef](#)] [[PubMed](#)]
140. Perez-Hernandez, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F. Identification of slow molecular order parameters for Markov model construction. *J. Chem. Phys.* **2013**, *139*, 015102. [[CrossRef](#)] [[PubMed](#)]
141. Schwantes, C.R.; Pande, V.S. Modeling molecular kinetics with tICA and the kernel trick. *J. Chem. Theory Comput.* **2015**, *11*, 600–608. [[CrossRef](#)] [[PubMed](#)]
142. Chen, W.; Ferguson, A.L. Molecular enhanced sampling with autoencoders: On-the-fly collective variable discovery and accelerated free energy landscape exploration. *J. Comput. Chem.* **2018**, *39*, 2079–2102. [[CrossRef](#)] [[PubMed](#)]
143. Wehmeyer, C.; Noe, F. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *J. Chem. Phys.* **2018**, *148*, 241703. [[CrossRef](#)]
144. Sittel, F.; Stock, G. Robust density-based clustering to identify metastable conformational states of proteins. *J. Chem. Theory Comput.* **2016**, *12*, 2426–2435. [[CrossRef](#)] [[PubMed](#)]
145. Lemke, O.; Keller, B.G. Density-based cluster algorithms for the identification of core sets. *J. Chem. Phys.* **2016**, *145*, 164104. [[CrossRef](#)] [[PubMed](#)]
146. Husic, B.E.; McKiernan, K.A.; Wayment-Steele, H.K.; Sultan, M.M.; Pande, V.S. A minimum variance clustering approach produces robust and interpretable coarse-grained models. *J. Chem. Theory Comput.* **2018**, *14*, 1071–1082. [[CrossRef](#)] [[PubMed](#)]
147. Noé, F. Probability distributions of molecular observables computed from Markov models. *J. Chem. Phys.* **2008**, *28*, 244103. [[CrossRef](#)] [[PubMed](#)]

148. Prinz, J.H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J.D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105. [[CrossRef](#)] [[PubMed](#)]
149. Trendelkamp-Schroer, B.; Noé, F. Efficient Bayesian estimation of Markov model transition matrices with given stationary distribution. *J. Chem. Phys.* **2013**, *138*, 164113. [[CrossRef](#)] [[PubMed](#)]
150. Noe, F.; Nueske, F. A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Model. Simul.* **2013**, *11*, 635–655. [[CrossRef](#)]
151. Nüske, F.; Keller, B.; Pérez-Hernández, G.; Mey, A.S.; Noé, F. Variational approach to molecular kinetics. *J. Chem. Theory Comput.* **2014**, *10*, 1739–1752. [[CrossRef](#)] [[PubMed](#)]
152. McGibbon, R.T.; Pande, V.S. Variational cross-validation of slow dynamical modes in molecular kinetics. *J. Chem. Phys.* **2015**, *142*, 124105. [[CrossRef](#)]
153. Mardt, A.; Pasquali, L.; Wu, H.; Noe, F. VAMPnets for deep learning of molecular kinetics. *Nat. Commun.* **2018**, *9*, 5. [[CrossRef](#)]
154. Husic, B.E.; Pande, V.S. Markov state models: From an art to a science. *J. Am. Chem. Soc.* **2018**, *140*, 2386–2396. [[CrossRef](#)] [[PubMed](#)]
155. Rudzinski, J.F.; Berau, T. Concurrent parametrization against static and kinetic information leads to more robust coarse-grained force fields. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 1373–1389. [[CrossRef](#)]
156. Rudzinski, J.F.; Kremer, K.; Berau, T. Communication: Consistent interpretation of molecular simulation kinetics using Markov state models biased with external information. *J. Chem. Phys.* **2016**, *144*, 051102. [[CrossRef](#)] [[PubMed](#)]
157. Berau, T.; Rudzinski, J.F. Accurate structure-based coarse-graining leads to consistent barrier-crossing dynamics. *Phys. Rev. Lett.* **2018**, *121*, 256002. [[CrossRef](#)] [[PubMed](#)]
158. Chen, J.; Chen, J.; Pinamonti, G.; Clementi, C. Learning effective molecular models from experimental observables. *J. Chem. Theory Comput.* **2018**, *14*, 3849–3858. [[CrossRef](#)] [[PubMed](#)]
159. Nüske, F.; Boninsegna, L.; Clementi, C. Coarse-graining molecular systems by spectral matching. *J. Chem. Phys.* **2019**, *151*, 044116. [[CrossRef](#)]
160. Crommelin, D.; Vanden-Eijnden, E. Diffusion estimation from multiscale data by operator eigenpairs. *Multiscale Model. Simul.* **2011**, *9*, 1588–1623. [[CrossRef](#)]
161. Jeong, D.; Kim, D.; Choi, M.Y.; Kim, H.J.; Jung, Y. Dynamic heterogeneity in room-temperature ionic liquids. In *Ionic Liquids: Theory, Properties, New Approaches*; Kokorin, A., Ed.; Intech Europe: Rijeka, Croatia, 2011; pp. 167–182.
162. Karimi-Varzaneh, H.A.; Mueller-Plathe, F.; Balasubramanian, S.; Carbone, P. Studying long-time dynamics of imidazolium-based ionic liquids with a systematically coarse-grained model. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4714–4724. [[CrossRef](#)]
163. Pal, T.; Vogel, M. Role of dynamic heterogeneities in ionic liquids: Insights from all-atom and coarse-grained molecular dynamics simulation studies. *ChemPhysChem* **2017**, *18*, 2233–2242. [[CrossRef](#)]
164. Moradzadeh, A.; Motevaselian, M.H.; Mashayak, S.Y.; Aluru, N.R. Coarse-grained force field for imidazolium-based ionic liquids. *J. Chem. Theory Comput.* **2018**, *14*, 3252–3261. [[CrossRef](#)]
165. Mukherjee, B.; Peter, C.; Kremer, K. Single molecule translocation in smectics illustrates the challenge for time-mapping in simulations on multiple scales. *J. Chem. Phys.* **2017**, *147*, 114501. [[CrossRef](#)] [[PubMed](#)]
166. Mukherjee, B.; Peter, C.; Kremer, K. Dual translocation pathways in smectic liquid crystals facilitated by molecular flexibility. *Phys. Rev. E* **2013**, *88*, 010502(R). [[CrossRef](#)] [[PubMed](#)]
167. Zeng, Q.H.; Yu, A.B.; Lu, G.Q. Multiscale modeling and simulation of polymer nanocomposites. *Prog. Polym. Sci.* **2008**, *33*, 191–269. [[CrossRef](#)]
168. Sun, L.; Zheng, C.; Webster, T.J. Self-assembled peptide nanomaterials for biomedical applications: Promises and pitfalls. *Int. J. Nanomed.* **2017**, *12*, 73–86. [[CrossRef](#)] [[PubMed](#)]
169. Šulc, P.; Romano, F.; Ouldridge, T.E.; Rovigatti, L.; Doye, J.P.K.; Louis, A.A. Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* **2012**, *137*, 135101. [[CrossRef](#)]
170. Ouldridge, T.E.; Šulc, P.; Romano, F.; Doye, J.P.K.; Louis, A.A. DNA hybridization kinetics: Zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* **2013**, *41*, 8886–8895. [[CrossRef](#)] [[PubMed](#)]

171. Schreck, J.S.; Ouldridge, T.E.; Romano, F.; Šulc, P.; Shaw, L.P.; Louis, A.A.; Doye, J.P.K. DNA hairpins destabilize duplexes primarily by promoting melting rather than by inhibiting hybridization. *Nucleic Acids Res.* **2015**, *43*, 6181–6190. [[CrossRef](#)]
172. Srinivas, N.; Ouldridge, T.E.; Šulc, P.; Schaeffer, J.M.; Yurke, B.; Louis, A.A.; Doye, J.P.K.; Winfree, E. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **2013**, *41*, 10641–10658. [[CrossRef](#)]
173. Mosayebi, M.; Romano, F.; Ouldridge, T.E.; Louis, A.A.; Doye, J.P.K. The role of loop stacking in the dynamics of DNA hairpin formation. *J. Phys. Chem. B* **2014**, *118*, 14326–14335. [[CrossRef](#)]
174. Habibi, M.; Rottler, J.; Plotkin, S.S. As simple as possible, but not simpler: Exploring the fidelity of coarse-grained protein models for simulated force spectroscopy. *PLoS Comput. Biol.* **2016**, *12*, e1005211. [[CrossRef](#)]
175. Whitford, P.C.; Noel, J.K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K.Y.; Onuchic, J.N. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins Struct. Funct. Bioinform.* **2009**, *75*, 430–441. [[CrossRef](#)] [[PubMed](#)]
176. Davtyan, A.; Schafer, N.P.; Zheng, W.; Clementi, C.; Wolynes, P.G.; Papoian, G.A. AWSEM-MD: Protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *J. Phys. Chem. B* **2012**, *116*, 8494–8503. [[CrossRef](#)] [[PubMed](#)]
177. Guttenberg, N.; Dama, J.F.; Saunders, M.G.; Voth, G.A.; Weare, J.; Dinner, A.R. Minimizing memory as an objective for coarse-graining. *J. Chem. Phys.* **2013**, *138*, 094111. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).