

Is Preservation of Symmetry Necessary for Coarse-Graining?

Maghesree Chakraborty, Jinyu Xu, Andrew D White*

Department of Chemical Engineering

University of Rochester

Abstract

There is a need for theory on how to group atoms in a molecule to define a coarse-grained (CG) mapping. This letter investigates the importance of preserving symmetry of the underlying molecular graph of a given molecule when choosing a CG mapping. 26 CG models of seven alkanes with three different techniques CG techniques were examined. We find preserving symmetry has no consistent effect on CG model accuracy.

Introduction:

Coarse-grained (CG) simulations have been widely used to study systems to address length-scale challenges in molecular dynamics[19, 24]. Selecting a CG mapping and obtaining the corresponding potential energy function are the key steps of defining a CG model. Both of these choices determine how closely a CG simulation reproduces results from the corresponding all-atom (AA) simulation. There are many approaches for fitting the potential[29, 36], but the choice of a CG mapping is still made using chemical intuition. There have been recent efforts to develop more systematic approaches to choose CG mappings [38, 34, 6, 13, 9], including our previous work[6]. Webb et al. [38] used spectral grouping iteratively to generate CG representations with successively lower resolutions. Wang and Gómez-Bombarelli [37] recently explored variational auto-encoder CG mappings, which is a promising new data-driven direction. The method, however, has only been tested on a small number of systems and provides no theory or explanation of mapping operator choice. There are pipeline softwares available, like BOCS[15], VOTCA[35] and Auto-Martini[3], to facilitate CG system preparation and subsequent simulation. However, these tools either require the user to select the mapping operator or create mapping based on established rules, like Martini CG mappings. Zavadlav et al. [40] reported a Bayesian framework to compare different CG mappings of water varying in resolution and number of interaction sites. Kanekal and Bereau [22] have also used a Bayesian framework to investigate the limit of effect of varying the number of CG bead types. Despite the recent attention systematic selection of CG mappings, there is a lack of comprehensive studies on different factors that might influence the efficiency of CG mappings. In this study we compare different symmetric and asymmetric CG mapping operators of alkanes to understand the importance of preserving symmetry while choosing a CG mapping.

Symmetries in molecules have a significant impact on their properties. Previously, molecular symmetry has been exploited to simplify calculation of physical properties [20] (like optical

*andrew.white@rochester.edu

activity[28, 5], dipole moment[30], melting point[39, 31],solubility[31], infra-red spectrum[32] and Raman spectrum[1]) and chemical properties[14, 25]. Besides the point symmetry groups[16] of molecules, another type of symmetry is called the topological symmetry[7, 23, 8] and refers to the symmetry of the underlying molecular graph where atoms are represented as nodes and bonds as edges[23]. The topological symmetry groups can be identified by using graph automorphism on the molecular graph[7]. The topological symmetry has recently been used in a recent work by Rosenfeld [33] for molecular synthesis based on topological symmetry. In our previous work[6], we found that only considering topologically symmetric CG mapping operators reduces the number of unique mappings by an order of magnitude for molecules with heavy atoms between 3 and 9. In this work, we test if considering only symmetric mapping operators is valid on alkanes.

We have considered propane and three isomers of hexane and octane for this study. For each molecule, symmetric and asymmetric CG mapping operators were used to perform bottom-up CG simulations. Additionally, we compared how the performances of CG mappings of hexane varied with the choice of different bottom-up approaches to fit the CG potential: force-matching (FM), iterative boltzmann inversion (IBI) and relative entropy (RE). Our goal is not to compare accuracies of these methods, but rather to ensure our conclusions about symmetry are independent of CG potential fitting method. Discussions on comparing these methods, including when they are equivalent, can be found in Kmiecik et al. [24] and Noid [29]. Ruhle et al. [35] also compared FM and IBI for small organic molecules like water, methanol, propane and hexane.

CG alkane simulations have been studied before and we have summarized the variety of CG mapping operators used for alkanes in previous studies in Table 1. The alkanes highlighted in red are included in our study. While we have tried to include relevant previous work, the list is not exhaustive. Previously, there has been limited study on the effect of symmetry on CG model fidelity. We had considered two asymmetric mappings for methanol in a previous work[6]. Recently, Jin et al. [21], mentioned that symmetry mismatch between the FG and CG representations had resulted in failure of MS-CG models in inter-facial systems. They developed the center of symmetry CG in order to preserve the symmetry present in the FG model when it is mapped into a CG model by adding a virtual site. Among the previously studied mapping operators for 16 alkanes listed in Table 1, almost all mappings preserve symmetry except the following: 2-3 mapping for n-pentane, 2-3-3 mapping for n-octane, 2-2-3-2 mapping for nonane, 2-2-3-3 mapping for n-decane, 3-3-3-2 and 2-2-2-3-2 mapping for n-undecane, 2-2-3-3-3 and 2-2-2-2-3-2 mappings for n-tridecane, 2-2-2-3-3-3 mapping for n-pentadecane, 2-2-3-3-3 mapping for n-hexadecane, and 2-2-2-2-3-3-3 and 3-2-3-3-3-3 mappings of n-heptadecane. We compare some of the previously published results with what we observe later.

Table 1: List of CG mappings used for alkanes in previous studies.

Molecule	CG Mappings	Metrics of comparison
Neopentane (2,2-dimethylpropane)	Single site mapping at COM [12]	RDF[12], VACF[12], self-diffusion coefficient[12]

n-Pentane	CG bead at each carbon-atom[18], 2-3[2]	Surface Tension[18, 2], self-diffusion coefficient[2], compressibility[2]
n-Hexane	CG bead at each carbon-atom[18], 2 bead mapping[12, 2], 2-2-2[2, 34], 2-1-1-2[34]	Surface Tension[18], RDF[12], VACF[12], self-diffusion coefficient[12, 2], compressibility[2]
Cyclohexane	Single site mapping at COM [12]	RDF[12], VACF[12], self-diffusion coefficient[12]
n-Heptane	2-3-2[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]
n-Octane	CG bead at each carbon-atom[18], 2-2-2-2[2], 2-3-3[2]	Surface Tension[18], self-diffusion coefficient[2], compressibility[2]
Nonane	3-3-3[2], 2-2-3-2[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]
n-Decane	CG bead at each carbon-atom[18], 2-2-2-2-2[2], 2-2-3-3[2]	Surface Tension[18], self-diffusion coefficient[2], compressibility[2]
n-Undecane	3-3-3-2[2], 2-2-2-3-2[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]
n-Dodecane	CG bead at each carbon-atom[18], CG1[10], CG2[10], CG3[10], CG4[10], 3-3-3-3[2], 2-2-2-2-2-2[2]	Surface Tension[18, 10], Temperature-density relationship[10], self-diffusion coefficient[2], compressibility[2]
n-Tridecane	2-2-3-3-3[2], 2-2-2-2-3-2[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]
Tetradecane	CG bead at each carbon-atom[18], 3-3-2-2-2-2[2], 2-2-2-2-2-2-2[2]	Surface Tension[18, 2], self-diffusion coefficient[2], compressibility[2]
n-Pentadecane	3-3-3-3-3[2], 2-2-2-3-3-3[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]
n-Hexadecane	CG bead at each carbon-atom[18], 2-2-3-3-3-3[2], 2-2-2-2-2-2-2-2[2]	Surface Tension[18, 2], self-diffusion coefficient[2], compressibility[2]
n-Heptadecane	2-2-2-2-3-3-3[2], 3-2-3-3-3-3[2]	Surface Tension[2], self-diffusion coefficient[2], compressibility[2]

n-Tetracosane	CG1,CG2,CG3,CG4[10]	Surface Tension[10], Temperature-density relationship[10]
---------------	---------------------	---

Our work is further motivated by previous CG studies which have yielded results contrary to chemical intuition. Some work has shown that the accuracy of CG mapping with the reference fine-grain (FG) simulation does not monotonically increase with increase in the resolution of the mapping[4]. Foley et al. [17] has shown how the information content in CG mapping seems to have an optimum with respect to CG mapping operator resolution. There are other reports[10, 11, 26], including our previous work[6], that corroborate that higher resolution CG mappings do not always outperform lower resolution mappings. This underlines the need of systematically studying factors which are often deemed trivial while using chemical intuition.

Method:

Symmetric and asymmetric mapping operators were considered for seven molecules: n-Propane, n-Hexane, Isohexane (2-Methylpentane), 2,3-Dimethylbutane, n-Octane, 3-Ethylhexane, and 4-Methylheptane. Three hexane isomers (n-Hexane, Isohexane, 2,3-Dimethylbutane) and three octane isomers (n-Octane, 3-Ethylhexane and 4-Methylheptane) were chosen since we wanted to study linear and branched isomers of 6-carbon and 8-carbon alkanes respectively. The illustrations of the mapping operators considered are shown in Figure 1.

We have also investigated how the choice of method for obtaining CG potentials affects the performance of different mapping operators. To complete this section of our study, we have considered 6 mappings of hexane, labelled in red in Figure 1.





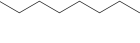
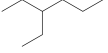


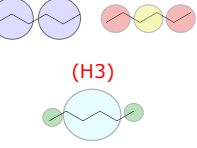
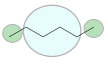
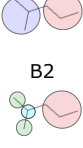
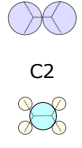
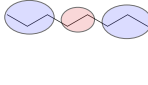
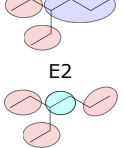
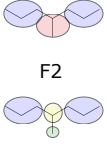

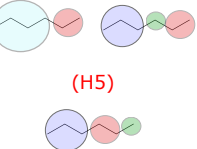
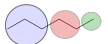

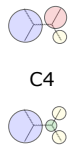
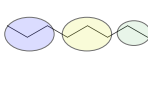
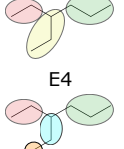
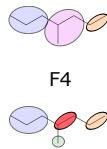
	P  Propane	A  n-Hexane	B  Isohexane	C  2, 3-Dimethylbutane	D  n-Octane	E  3-Ethylhexane	F  4-Methylheptane
Symmetric	P1 	A1 (H1) A2 (H2)  (H3) 	B1 	C1 	D1 	E1 	F1 
Asymmetric	P2 	A3 (H4) A4 (H6)  (H5) 	B3 	C3 	D2 	E3 	F3 

Figure 1: Illustration of symmetric and asymmetric mapping operators of the seven molecules. We have highlighted in red the alternative labels for hexane CG mapping operators that are used to investigate the effect of different methods (FM, IBI and RE) of obtaining CG potentials on performance of the mappings.

We compare the performances of the mapping operators using CG potentials obtained by FM, IBI and RE. The FG simulation for each molecule was performed using GROMACS-2016 for 1 ns with the OPLS-AA force field and 1 fs time step. The densities (in g/cm^3) used for FG simulations are as follows: propane - 0.635, n-hexane - 0.650, 2-methylpentane - 0.655, 2,3-dimethylbutane - 0.660, n-octane - 0.699, 3-ethylhexane - 0.7079 and 4-methylheptane - 0.705.

For each FG simulation, the NVT ensemble was maintained at 300 K for all molecules except propane, for which the FG simulation was conducted at a temperature of 200 K[35]. For all the 7 molecules, FM-CG simulations were conducted according to the methods described in our previous work [6]. The iterative methods, IBI and RE, were also implemented using VOTCA[27, 35]. All the CG simulations were run for 1 ns with a time-step of 2fs. To evaluate how different mapping operators performed, we compared the center of mass (COM) radial distribution functions (RDFs) and the velocity autocorrelation functions (VACFs) of the CG mapping to those obtained from the corresponding FG simulations. For quantitative analysis evaluating the symmetric and asymmetric mappings, we computed the squared error between a CG mapping result and the FG result, normalized over all the CG mappings of a given molecule. For comparing FM, IBI and RE for a particular mapping, we computed the squared error and normalized over the three CG simulation results. The roots of the normalized mean square errors are the final reported values.

Results and Discussion

Figure 2 and Figure 3 compare the mean square COM-RDF errors and the mean square VACF errors respectively of the asymmetric and symmetric mapping operators of the molecules. As seen in Figure 2, preservation of topological symmetry present in the FG model while selecting a CG mapping does not guarantee closer agreement with reference FG COM-RDF. For instance, E1, the symmetric 3-bead CG mapping of 3-ethylhexane, has higher COM-RDF square error than E3, the asymmetric 3-bead CG mapping of 3-ethylhexane, even though both of them have the same degrees of freedom. Similar results are seen based on COM-RDF square error for other symmetric and asymmetric mapping operator pairs with the same degrees of freedom for n-octane (D1-D2), 3-ethylhexane (E2-E4) and 4-methylheptane (F1-F3, F2-F4). When mean square error for VACF is the metric of comparison, we see in the symmetric and asymmetric CG mapping pair, F2-F4 for 4-methylheptane, that the asymmetric mapping yields lower mean square VACF error than the symmetric one. F2 and F4 mappings have comparable degrees of freedom. Note that the symmetric and asymmetric mapping pairs above have an equal number of beads, and thus equal degrees of freedom, but the asymmetric mappings have more bead types. This gives more trainable parameters for asymmetric mappings and could explain the sometimes better performance of asymmetric mappings.

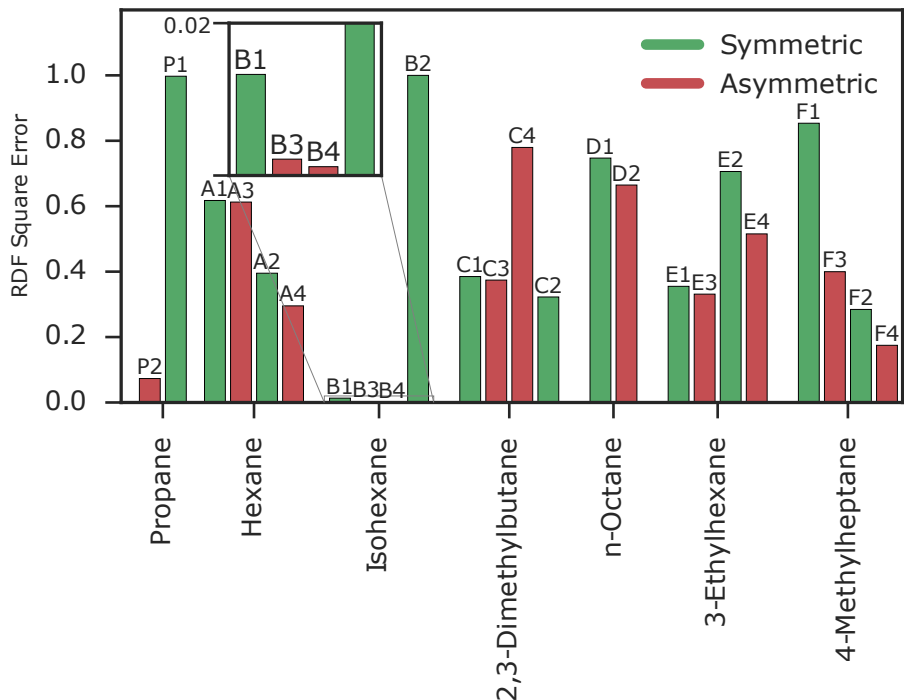


Figure 2: COM-RDF mean square errors of symmetric and asymmetric mappings of the seven molecules: propane, n-hexane, isohexane, 2,3-dimethylbutane, n-octane, 3-ethylhexane and 4-methylheptane. For each molecule, the CG mappings have been arranged in the order of increasing resolution. Additionally, for each molecule, the COM-RDF mean square errors have been normalized over all its mappings.

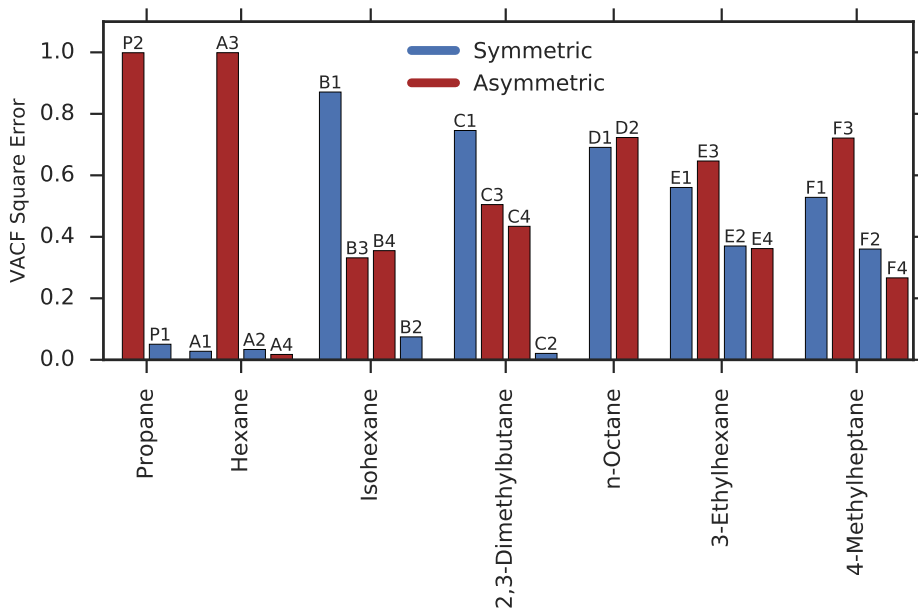


Figure 3: VACF mean square errors of symmetric and asymmetric mappings of the seven molecules: propane, n-hexane, isohexane, 2,3-dimethylbutane, n-octanes, 3-ethylhexane and 4-methylheptane. For each molecule, the CG mappings have been arranged in the order of increasing resolution. Additionally for each molecule, the VACF mean square errors have been normalized over all its mappings.

We also observe that increasing the degrees of freedom by selecting a higher resolution CG mapping does not guarantee a closer agreement with FG results. As seen in Figure 2, the 4-bead asymmetric mapping operator for 2,3-dimethylbutane, C4, gives higher COM-RDF square error than the 2-bead mappings, C1 and C3. This result corroborates with previously reported works[10, 17], which showed that increasing the resolution of a CG mapping operator does not guarantee better agreement with FG results. Additionally, C4, which has higher COM-RDF mean square error than C1 and C3, yields lower VACF mean square error compared to C1 and C3. We also see the reversal of this result where asymmetric mappings which give lower COM-RDF mean square error, give higher VACF mean square error compared to the corresponding symmetric mapping operators. This is seen for propane, n-octane and 3-ethylhexane (E1-E3).

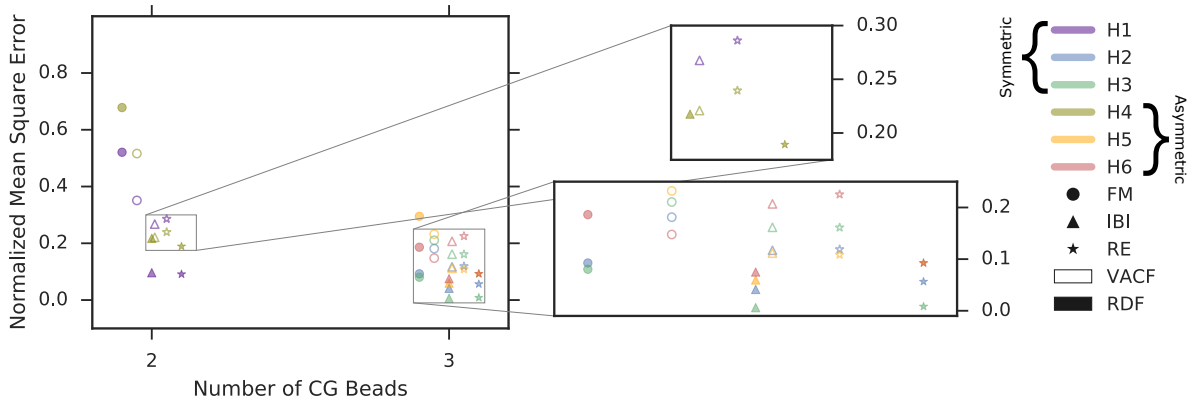


Figure 4: Comparison of the normalized mean square errors for RDF and VACF for the symmetric and asymmetric mapping operators of hexane. Three different methods (FM, IBI and RE) were also contrasted. The three methods are denoted by different markers and the two evaluation metrics, COM-RDF mean square error and VACF mean square error, are differentiated by filled and unfilled markers, respectively.

Figure 4 shows the result obtained by comparing FM, IBI and RE for six hexane mappings highlighted in red in the Figure 1 illustration. The COM-RDF mean square error for H3 is lowest among the 6 hexane mapping operators (H1 through H6) for all the three methods (FM, IBI and RE). Though it is a symmetric 3-bead mapping, it has more skewed mass distribution among its beads than H2, a comparable 3-bead symmetric mapping operator. In contrast, asymmetric mappings H6 (for FM) and H5 (for IBI and RE) give the least VACF mean square errors. Additionally for IBI and RE, the 2-bead asymmetric mapping, H4, yields lower VACF mean square error than the comparable 2-bead symmetric mapping operator, H1.

To summarize, for both parts of the study, involving different molecules and involving the three methods on hexane, we observe that the performance of symmetric and asymmetric mapping operators vary based on the metric of evaluation. Similar results were reported in previous studies as cited in Table 1. An et al. [2], in their work on developing transferable CG models for hydrocarbons, showed that a 3-bead hexane CG mapping better agreed with experimental values of self-diffusion coefficient and expansibility compared to a 2-bead mapping. However, the 2-bead mapping yielded lower error when compressibility and surface tension were considered[2]. In the same work, the asymmetric mapping for n-nonane (2-2-3-2) agreed with experimental values better than the symmetric mapping (3-3-3) when self-diffusion coefficient and compressibility were considered. Both the mappings, 2-2-2-3-2 and 3-3-3-2, for undecane studied by An et al. [2] are asymmetric. The 5 bead 2-2-2-3-2 mapping yielded compressibility and surface-tension values closer to experimental results than the 4 bead 3-3-3-2 mapping. On the contrary, the 4 bead mapping gave expansibility and self-diffusion coefficients closer to the experimentally observed values than the 5 bead mapping. While there are widely used evaluation metrics like the RDF and others as listed in 1, it is still a matter of preference since there is no metric that has been unanimously decided to be most effective for CG mapping evaluation. It should be noted that evaluation metrics to compare CG mappings is still. We have chosen COM-RDF since it is not dependent on the number of beads in a CG mapping. This allows us to compare the COM-RDFs of CG mappings of different resolutions. Our second evaluation

metric, VACF, has the same advantage.

Conclusion

In this work we show that CG mapping operators which break symmetry sometimes perform better than symmetric CG mapping operators with comparable degrees of freedom. To our knowledge, this is the first attempt to systematically study the effect of CG mapping symmetry on the performance. Further, we provide additional evidence to support previously reported hypothesis that the information content of a CG mapping operators do not monotonically increase with resolution[]. These two factors can particularly be useful to systematically select multi-scale CG representation of macro-molecules like polymers and proteins, where it might be desirable to have specific areas of interest at higher resolutions compared to others. The results reported in this work also warrant further exploration of the possible metrics of comparison between FG and CG simulations.

Acknowledgements

This work is supported by the National Science Foundation under grant No. 1764415. We thank the Center for Integrated Research Computing at the University of Rochester for providing the computational resources required to complete this study. We also thank Dr. Tristan Bereau for providing helpful feedback.

References

- [1] Chapter 4 Symmetry of molecules, group theory and its applications in vibrational spectroscopy. In *Mod. Fourier Transform Infrared Spectrosc.*, volume 35 of *Comprehensive Analytical Chemistry*, pages 41–96. Elsevier, 2001. doi: [https://doi.org/10.1016/S0166-526X\(01\)80007-6](https://doi.org/10.1016/S0166-526X(01)80007-6). URL <http://www.sciencedirect.com/science/article/pii/S0166526X01800076>.
- [2] Y. An, K. K. Bejagam, and S. A. Deshmukh. Development of New Transferable Coarse-Grained Models of Hydrocarbons. *J. Phys. Chem. B*, 122(28):7143–7153, 2018. ISSN 15205207. doi: 10.1021/acs.jpcc.8b03822.
- [3] T. Bereau and K. Kremer. Automated Parametrization of the Coarse-Grained Martini Force Field for Small Organic Molecules. *J. Chem. Theory Comput.*, 11(6):2783–2791, 2015. ISSN 15499626. doi: 10.1021/acs.jctc.5b00056.
- [4] P. Buslaev and I. Gushchin. Effects of Coarse Graining and Saturation of Hydrocarbon Chains on Structure and Dynamics of Simulated Lipid Molecules. *Sci. Rep.*, 7(1):1–15, 2017. ISSN 20452322. doi: 10.1038/s41598-017-11761-5. URL <http://dx.doi.org/10.1038/s41598-017-11761-5>.
- [5] J. L. Carlos. Molecular symmetry and optical inactivity. *J. Chem. Educ.*, 45(4):248–251, 1968. ISSN 00219584. doi: 10.1021/ed045p248.

- [6] M. Chakraborty, C. Xu, and A. D. White. Encoding and Selecting Coarse-Grain Mapping Operators with Hierarchical Graphs. 134106, 2018. ISSN 0021-9606. doi: 10.1063/1.5040114. URL <http://arxiv.org/abs/1804.04997>.
- [7] D. Chambers and E. Flapan. Topological Symmetry Groups of Small Complete Graphs. *Symmetry (Basel)*, 6(2):189–209, 2014. ISSN 2073-8994. doi: 10.3390/sym6020189. URL <http://www.mdpi.com/2073-8994/6/2/189/>.
- [8] W. Chen, J. Huang, and M. K. Gilson. Identification of Symmetries in Molecules and Complexes. *J. Chem. Inf. Comput. Sci.*, 44(4):1301–1313, 2004. ISSN 00952338. doi: 10.1021/ci049966a.
- [9] Y. L. Chen and M. Habeck. Data-driven coarse graining of large biomolecular structures. *PLoS One*, 12(8):1–17, 2017. ISSN 19326203. doi: 10.1371/journal.pone.0183057.
- [10] M. Dallavalle and N. F. A. van der Vegt. Evaluation of Mapping Schemes for Systematic Coarse Graining of Higher Alkanes. *Phys. Chem. Chem. Phys.*, 19:23034–23042, 2017. ISSN 1463-9076. doi: 10.1039/C7CP03926C. URL <http://dx.doi.org/10.1039/C7CP03926C><http://xlink.rsc.org/?DOI=C7CP03926C>.
- [11] T. Dannenhoffer-Lafage, A. D. White, and G. A. Voth. A Direct Method for Incorporating Experimental Data into Multiscale Coarse-Grained Models. *J. Chem. Theory Comput.*, 12(5):2144–2153, 2016. ISSN 15499626. doi: 10.1021/acs.jctc.6b00043.
- [12] G. Deichmann, V. Marcon, and N. F. A. Van Der Vegt. Bottom-up derivation of conservative and dissipative interactions for coarse-grained molecular liquids with the conditional reversible work method. *J. Chem. Phys.*, 141(22), 2014. ISSN 00219606. doi: 10.1063/1.4903454.
- [13] P. Diggins, C. Liu, M. Deserno, and R. Potestio. Optimal Coarse-Grained Site Selection in Elastic Network Models of Biomolecules. *J. Chem. Theory Comput.*, 15(1):648–664, 2019. ISSN 15499626. doi: 10.1021/acs.jctc.8b00654.
- [14] J. D. Dunitz. Symmetry arguments in chemistry. *Proc. Natl. Acad. Sci. U. S. A.*, 93(25):14260–14266, 1996. ISSN 00278424. doi: 10.1073/pnas.93.25.14260.
- [15] N. J. Dunn, K. M. Lebold, M. R. Delyser, J. F. Rudzinski, and W. G. Noid. BOCS: Bottom-up Open-source Coarse-graining Software. *J. Phys. Chem. B*, 122(13):3363–3377, 2018. ISSN 15205207. doi: 10.1021/acs.jpcc.7b09993.
- [16] M. Electronics. Chapter 1 Molecular Electronics. pages 7–30. doi: 10.1088/978-1-6817-4637-1ch1.
- [17] T. T. Foley, M. S. Shell, and W. G. Noid. The Impact of Resolution upon Entropy and Information in Coarse-Grained Models. *J. Chem. Phys.*, 143(24):243104, 2015. ISSN 00219606. doi: 10.1063/1.4929836.
- [18] G. Gyawali, S. Sternfield, R. Kumar, and S. W. Rick. Coarse-Grained Models of Aqueous and Pure Liquid Alkanes. *J. Chem. Theory Comput.*, 13(8):3846–3853, 2017. ISSN 15499626. doi: 10.1021/acs.jctc.7b00389.

- [19] H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, and S. J. Marrink. The Power of Coarse Graining in Biomolecular Simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 4(3):225–248, 2014. ISSN 17590884. doi: 10.1002/wcms.1169.
- [20] J. Ivanov. Molecular symmetry perception. *J. Chem. Inf. Comput. Sci.*, 44(2):596–600, 2004. ISSN 0095-2338. doi: 10.1021/ci0341868. URL <http://www.ncbi.nlm.nih.gov/pubmed/15032540>.
- [21] J. Jin, Y. Han, and G. A. Voth. Coarse-graining involving virtual sites: Centers of symmetry coarse-graining. *J. Chem. Phys.*, 150(15), 2019. ISSN 00219606. doi: 10.1063/1.5067274.
- [22] K. H. Kanekal and T. Bereau. Resolution limit of data-driven coarse-grained models spanning chemical space. 1(July), 2019. URL <http://arxiv.org/abs/1907.04082>.
- [23] D. J. Klein and B. Mandal. Local symmetries for molecular graphs. *Match*, 74(2):247–258, 2015. ISSN 03406253.
- [24] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, and A. Kolinski. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.*, 116(14):7898–7936, 2016. ISSN 15206890. doi: 10.1021/acs.chemrev.6b00163.
- [25] S. C. Lan, P. Raghunath, Y. H. Lu, Y. C. Wang, S. W. Lin, C. M. Liu, J. M. Jiang, M. C. Lin, and K. H. Wei. Symmetry and coplanarity of organic molecules affect their packing and photovoltaic properties in solution-processed solar cells. *ACS Appl. Mater. Interfaces*, 6(12):9298–9306, 2014. ISSN 19448252. doi: 10.1021/am501659u.
- [26] S. Markutsya, A. Devarajan, J. Y. Baluyut, T. L. Windus, M. S. Gordon, and M. H. Lamm. Evaluation of Coarse-Grained Mapping Schemes for Polysaccharide Chains in Cellulose. *J. Chem. Phys.*, 138(21):214108, 2013. ISSN 00219606. doi: 10.1063/1.4808025.
- [27] S. Y. Mashayak, M. N. Jochum, K. Koschke, N. R. Aluru, V. Rühle, and C. Junghans. Relative entropy and optimization-driven coarse-graining methods in VOTCA. *PLoS One*, 10(7), 2015. ISSN 19326203. doi: 10.1371/journal.pone.0131754.
- [28] S. F. Mason. Optical activity and molecular dissymmetry. *Contemp. Phys.*, 9(3):239–256, 1968. ISSN 13665812. doi: 10.1080/00107516808204399.
- [29] W. G. Noid. Perspective: Coarse-grained models for biomolecular systems. *J. Chem. Phys.*, 139(9), 2013. ISSN 00219606. doi: 10.1063/1.4818908.
- [30] R. Obaid and M. Leibscher. A molecular symmetry analysis of the electronic states and transition dipole moments for molecules with two torsional degrees of freedom. *J. Chem. Phys.*, 142(6), 2015. ISSN 00219606. doi: 10.1063/1.4907405.
- [31] R. Pinal. Effect of molecular symmetry on melting temperature and solubility. *Org. Biomol. Chem.*, 2(18):2692–2699, 2004. ISSN 14770520. doi: 10.1039/b407105k. URL <http://xlink.rsc.org/?DOI=b407105k>.

- [32] T. Room, L. Peedu, M. Ge, D. Hvonen, U. Nagel, S. Ye, M. Xu, Z. Baic, S. Mamone, M. H. Levitt, M. Carravetta, J. Y. Chen, X. Lei, N. J. Turro, Y. Murata, and K. Komatsu. Infrared spectroscopy of small-molecule endfullerenes. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 371(1998): 20110631, sep 2013. ISSN 1364503X. doi: 10.1098/rsta.2011.0631. URL <https://royalsocietypublishing.org/doi/10.1098/rsta.2011.0631>.
- [33] V. R. Rosenfeld. Looking into the future of molecules with novel topological symmetries. *J. Math. Chem.*, jun 2019. ISSN 0259-9791. doi: 10.1007/s10910-019-01042-z. URL <http://link.springer.com/10.1007/s10910-019-01042-z>.
- [34] J. F. Rudzinski and W. G. Noid. Investigation of coarse-grained mappings via an iterative generalized Yvon-Born-Green method. *J. Phys. Chem. B*, 118(28):8295–8312, 2014. ISSN 15205207. doi: 10.1021/jp501694z.
- [35] V. Ruhle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. Versatile Object-Oriented Toolkit for Coarse-Graining Applications. *J. Chem. Theory Comput.*, 5(12): 3211–3223, 2009. ISSN 15499618. doi: 10.1021/ct900369w.
- [36] M. G. Saunders and G. A. Voth. Coarse-graining methods for computational biology. *Annu. Rev. Biophys.*, 42:73–93, 2013. ISSN 1936-1238. doi: 10.1146/annurev-biophys-083012-130348. URL <http://www.annualreviews.org/doi/pdf/10.1146/annurev-biophys-083012-130348>.
- [37] W. Wang and R. Gmez-Bombarelli. Coarse-Graining Auto-Encoders for Molecular Dynamics. 2018. URL <http://arxiv.org/abs/1812.02706>.
- [38] M. A. Webb, J. Y. Delannoy, and J. J. De Pablo. Graph-Based Approach to Systematic Molecular Coarse-Graining. *J. Chem. Theory Comput.*, 15:1199–1208, 2019. ISSN 15499626. doi: 10.1021/acs.jctc.8b00920.
- [39] J. Wei. Molecular symmetry, rotational entropy, and elevated melting points. *Ind. Eng. Chem. Res.*, 38(12):5019–5027, 1999. ISSN 08885885. doi: 10.1021/ie990588m.
- [40] J. Zavadlav, G. Arampatzis, and P. Koumoutsakos. Bayesian selection of coarse-grained models of liquid water. pages 1–11.