Predictive Modeling for Geometric Rule-Based Methods

Brittany R. Hoard

Department of Computer Science
University of New Mexico
Albuquerque, NM, United States
bhoard@unm.edu

***Abstract*—Previously, we proposed a method for incorporating molecular geometry in a biological rule-based model by encoding molecular curvature into the rules and associated binding rate constants. We combined this method with a 3D rigid-body Monte Carlo simulation to model antigen-antibody aggregation. In this work, we use our geometric rule-based method to develop a model for predicting the output of the full-resolution Monte Carlo simulation given the output of lower resolution simulations. The purpose of this predictive model is to reduce the computational cost of the Monte Carlo simulation. We develop this model by first choosing a rule set for each molecular geometry and varying only the binding rate constant for each Monte Carlo resolution, and then fitting the resulting data to a function. We examine the calculation time needed for each predictive model to demonstrate how this model is more efficient than running a full-resolution simulation. We find that this method can reduce the computational time of the Monte Carlo simulation by up to 20%.**

***Index Terms*—rule-based modeling, geometric modeling, molecular geometry, Monte Carlo simulation, computational efficiency**

I. Introduction

The allergic response is produced by the release of immune mediators by mast cells and basophils [1]. This process, in turn, is initiated by the aggregation of antigens and IgE-FcϵRI antibody-receptor complexes [1]. Computational modeling of antibody-antigen aggregate formation as well as the size and structure of these aggregates is an important tool for greater understanding of the allergic response. Allergens can possess multiple conformations with different geometries, which may affect aggregate structure and size. It is known that the allergen fold affects the IgE reactivity of its epitopes, or binding sites [2]. Consequently, the incorporation of molecular geometry into aggregation models can more accurately capture details of the aggregation process, and may lead to insights into how geometry affects aggregate formation. However, it is challenging to simulate aggregation due to the computational cost of simulating large molecules. Methods to geometrically model antibody aggregation inspired by rigid body robotic motion simulations have previously been developed; however, high computational cost mandates that the resolution of the 3D molecular models be reduced, which affects the results of the simulation. Rule-based modeling can be used to model aggregation with low computational cost, but traditional rule-based modeling approaches do not include details of molecular geometry.

In previous work, we proposed a novel implementation of rule-based modeling that encodes details of molecular geometry into the rules and the binding rate constant associated with each rule [3], [4]. We then performed a study of antigen-antibody aggregation using our proposed method combined with a previously developed 3D rigid-body Monte Carlo simulation.

We first simulated the binding of IgE antibodies bound to cell surface receptors FcϵRI to various binding regions of the allergen Pen a 1 using the aforementioned Monte Carlo simulation, and we analyzed the distribution of the sizes of the aggregates that form during the simulation [3], [4]. Then, using our novel rule-based approach, we optimized a rule-based model according to the geometry of the Pen a 1 molecule and the data from the Monte Carlo simulation. In particular, we used the distances between the binding regions of the Pen a 1 molecule to optimize the rules and associated binding rate constants. The optimized rule-based models provide information about the average steric hindrance between binding regions and the probability that IgE-FcϵRI receptor complexes will bind to these regions. In addition, the optimized rule-based models provide a means of quantifying the variation in aggregate size distribution that results from differences in molecular geometry.

We performed this procedure for seven resolutions and three molecular conformations of Pen a 1 [3], [4]. We then analyzed the impact of resolution and conformation on the aggregate size distribution and on the optimal rule-based model.

In this work, we created models to enable the prediction of Monte Carlo results at full model resolution by only running the simulation at lower resolutions. If the number of low-resolution simulations is small enough, the computational time required to obtain results will be less than running a single set of full-simulation calculations. To achieve this, we analyzed trends in the variation of the parameters of our optimized rule-based models with resolution. We first fixed the rule set and varied only the binding rate constant for each resolution, and then fitted the resulting rate constant versus resolution data to a function. We then used that function to predict the appropriate rate constant for a full-resolution model. Finally, we used this

978-1-7281-1867-3/19/$31.00

rate constant in our geometric rule-based model, which we ran to obtain the aggregate size distribution, which is the result in which we are most interested.

The main contributions of this work are as follows:

1) Outline a method for creating a model that can be used to predict the results of a 3D full-resolution rigid-body Monte Carlo simulation by running lower resolution Monte Carlo simulations.
2) Assess the computation time required for each predictive model to show that these predictive models can save computational time.

II. Related Work

Our work builds on computational methods for modeling molecular geometry in biochemical processes, as well as experimental methods to study IgE antibody binding behavior.

A. Rule-Based Modeling

Biological signaling systems are often comprised of macromolecules that can exist in a large number of functionally distinct states. This number scales exponentially with the amount of modification possibilities [5]. One problem that arises when modeling these systems is the specification problem, i.e. how to specify such a large system.

One solution is implicit specification, which involves the coarse-graining of sets of reactions and parameters into "rules"; the only explicitly specified features in a reaction rule are those which affect the reaction. Rules define the conditions for molecular transformations and interactions, and are associated with rate laws [6]. Some rules define multiple reactions, which means that all of these reactions are associated with the same rate law. The rules can usually be specified independently. Rule-based specification methods include Kappa-calculus [7], BioNetGen [8], ANC [9], and ML-Rules [10]. The Simmune project and the SSC allow the specification of molecules within spatial regions of arbitrary geometries [11].

The rule-based methods can be population-based, particle-based, or hybrid. Population-based methods include ODE/PDE numerical integration and the stochastic Gillespie algorithm. In these methods, the application of a rule changes the size of one of the populations, each of which consists of all molecules that share the same state and same species. The system state space can be very large, so methods to reduce it have been introduced [5].

Particle-based rule evaluation involves tracking individual "particles" (molecules and molecular complexes) through the simulation [6]. This is a network-free method; at any time point, only the existing particles, their states, and the possible reactions for the existing particles are necessary. Spatial particle-based methods include an explicit specification of space, and include SRSim [12] and MCell [13].

Our method differs from traditional approaches to rule-based modeling in that the rules used in our method are constrained by molecular geometry.

B. Geometric Molecular Modeling

The spatial simulation software SRSim [12] is a rule-based modeling method that allows for the specification of molecule geometry. SRSim integrates rule-based modeling, molecular dynamics, and a stochastic, diffusing-particle simulator. Molecular geometry is provided by the user via data files. Our proposed method is different in that it is a purely rule-based ODE model that does not require any additional data files to run, as the molecular geometry is encoded into the rules themselves. In addition, our method only requires the BioNetGen software to run.

The stochastic, particle-based Meredys software [14] uses Brownian dynamics to simulate reaction-diffusion systems at the mesoscopic level. It requires the specification of details such as molecule positions, molecular geometry, reaction site positions, and reaction types. Our rule-based method is population-based and only requires the distances between pairs of binding regions on a single antigen molecule to create the model.

Computational methods for modeling two-molecule ligand-receptor docking simulate systems on a smaller scale than those studied using our method. Our method uses more realistic geometric molecular models than do existing methods for self-assembly of molecular structures, such as those employing simple bead models [15].

C. Experimental Methods

Nanoprobe labeling and transmission electron microscopy (TEM) of cell membranes are used to study cell signaling. Methods for the spatial analysis of these nanoprobes, including statistical analysis of clustering, were developed in [16]. Quantum dot (QD)-IgE probes that bind FcϵRI have been used to study the mobility of receptors by single-particle tracking [17]. The kinetics of DNP-BSA binding to IgE has been studied by observing fluorescence quenching [18], and the data was analyzed using a mathematical model in which the IgE binding sites are transiently exposed, allowing binding and cross-linking to occur.

III. Methods

Our cutoff distance model has two variable parameters: the cutoff distance range, which corresponds to the rule set, and the rate constant k_{f2}. For the purpose of creating a model that can be used to predict the aggregate size data of higher resolutions given only lower resolution data, it is helpful to fix one parameter and only allow the other parameter to vary. Then, a trend can be more easily observed.

Due to the difficulty of fixing k_{f2} at a constant value for all resolutions and still being able to obtain rule-based models that fit the Monte Carlo data well, we have fixed the cutoff distance range for all resolutions and allowed k_{f2} to vary. The rate constant k_{f2} was optimized using an adaptive method based on the Metropolis algorithm that minimized the difference between the Monte Carlo output and the output of the rule-based model. We have chosen to fix the cutoff distance range for all resolutions at the optimal range for the 25% reduced

resolution. This range is 5.6-6.2 nm for the native type, 5.8-6.0 nm for the S-shaped type, and 6.8-8.6 nm for the U-shaped type. Since our goal is to be able to predict the results of the 0% reduced resolution, the rule set of the 25% reduced resolution strikes a balance between allowing good fits to the 0% reduced Monte Carlo data and still allowing reasonable fits to the lower resolution Monte Carlo data. The aggregate size data for the 0% reduced resolution, along with the optimal rate constant versus resolution data, are shown in Figure 1 for the native type Pen a 1, Figure 2 for the S-shaped type, and Figure 3 for the U-shaped type. These data were fitted to exponential functions.

Ultimately, we want to develop data sets for fitting our models using only rate constant versus resolution data for lower resolutions, so that the computationally costly higher resolution Monte Carlo simulations do not need to be run. For this reason, we developed models using a method similar to that used for the models described above, with the exception that the exponential functions are fitted to the data sets with the following data points *omitted*: 0% reduced data; 0% and 25% reduced data; 0%, 25%, and 90% reduced data; 0%, 25%, and 65% reduced data; and 0%, 25%, 65%, and 90% reduced data.

A possible source of error is the Monte Carlo simulation data due to the somewhat small number of runs used to fit our predictive exponential functions and the Monte Carlo data not always fitting the expected trend of higher aggregate sizes with higher percent reduction in resolution (see Figure 2). Another source of error is the fitted function itself, which provides a predicted rate constant that is different from the actual rate constant.

IV. Results and Discussion

A. Results and Model Accuracy

For comparison purposes, Figures 1, 2, and 3 display the rule-based modeling histogram data corresponding to both the actual optimized rate constant for the 0% reduction Monte Carlo data and the rate constant predicted by the exponential function to fit the 0% reduction data. We observe that these models accurately predict the optimized aggregate size distribution.

The models that omit only the 0% reduced data, along with the corresponding aggregate size histogram predicted for the 0% reduced resolution, are shown in Figure 4 for the native type, Figure 5 for the S-shaped type, and Figure 6 for the U-shaped type. The aggregate size distributions predicted by these models are similar to the optimized distributions, although the predicted distribution for the native type shows a significant difference in the aggregate size probability values. Our time analysis (Section IV-B) shows that these models are computationally expensive.

The aggregate size histogram predicted for the 0% reduced resolution for the models that omit the 0%, 25%, and 90% reduced data are shown in Figure 7 for the native type, Figure 8 for the S-shaped type, and Figure 9 for the U-shaped type. We observe that these models continue to accurately capture the optimized aggregate size distribution, although there is some difference in the probability values. Our time analysis (Section IV-B) shows that these models are computationally expensive.

The next two sets of models that we will discuss are especially important since our time analysis showed that the total run times for these models are less than that of the full-resolution model.

The aggregate size histogram predicted for the 0% reduced resolution for the models that omit the 0%, 25%, and 65% reduced data are shown in Figure 10 for the native type, Figure 11 for the S-shaped type, and Figure 12 for the U-shaped type. We observe that these models accurately capture the location of the first and second most probable aggregate sizes, and the probability values are closer to that of the Monte Carlo model than the previously described models.

The aggregate size histogram predicted for the 0% reduced resolution for the models that omit the 0%, 25%, 65%, and 90% reduced data are shown in Figure 13 for the native type, Figure 14 for the S-shaped type, and Figure 15 for the U-shaped type. Similarly to the previously described model, We observe that these models accurately capture the location of the first and second most probable aggregate sizes, and the probability values are reasonably close to that of the Monte Carlo model.

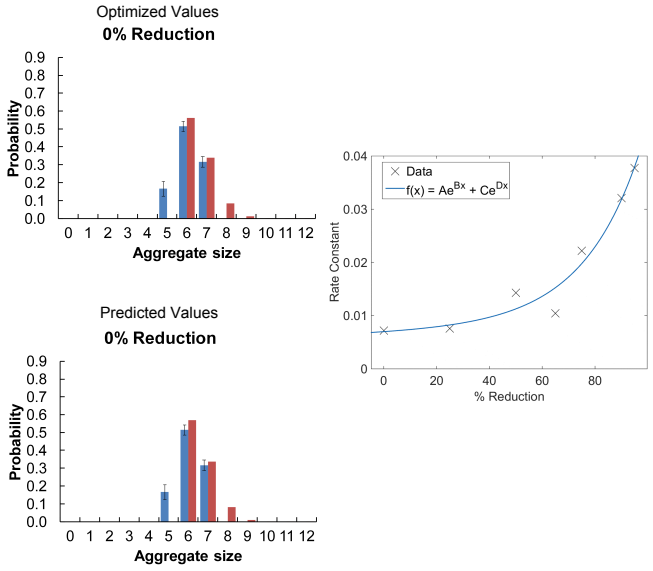

Figure 1. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the native Pen a 1 for the optimized rate constant (top left) and predicted rate constant (bottom left). Rate constant versus resolution data for the native Pen a 1 conformation (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 5.6-6.2 nm (see rule set 2 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the native type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=0.0066, B=0.0027, C=0.0004, and D=0.0462.

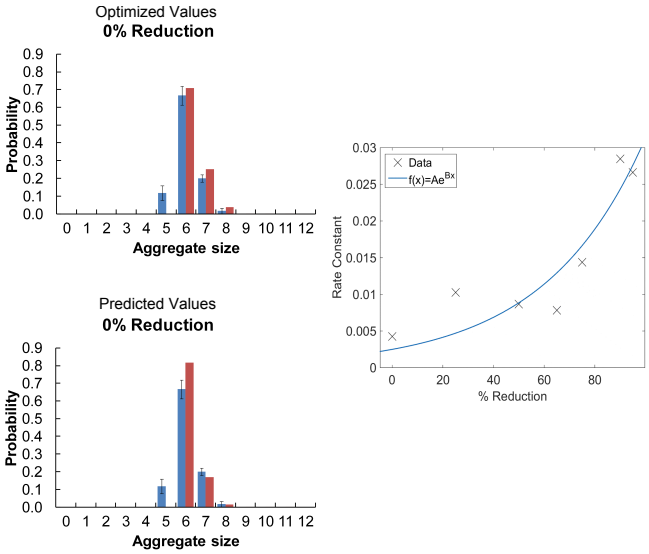

Figure 2. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the S-shaped Pen a 1 for the optimized rate constant (top left) and predicted rate constant (bottom left). Rate constant versus resolution data for the S-shaped Pen a 1 conformation (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 5.8-6.0 nm (see rule set 6 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the S-shaped type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=0.0025 and B=0.0253.

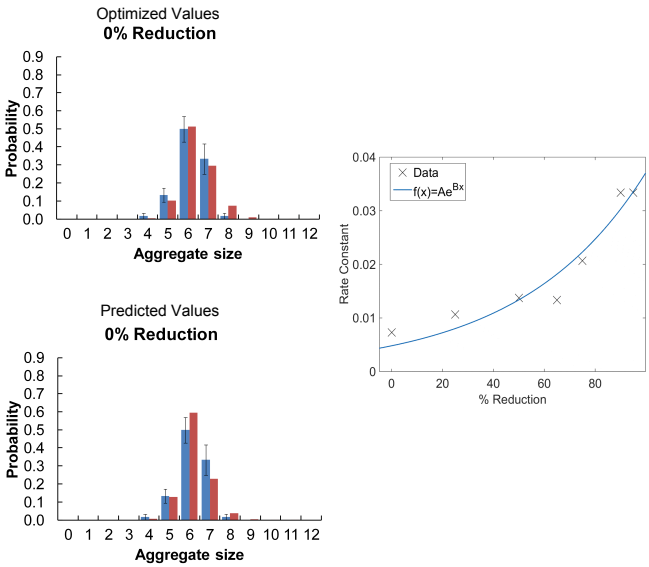

Figure 3. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the U-shaped Pen a 1 for the optimized rate constant (top left) and predicted rate constant (bottom left). Rate constant versus resolution data for the U-shaped Pen a 1 conformation (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 6.8-8.6 nm (see rule set 9 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the U-shaped type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=0.0048 and B=0.0204.

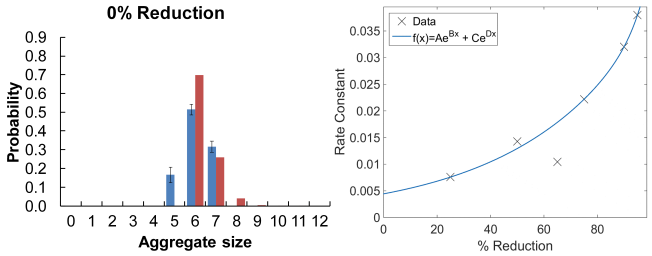

Figure 4. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the native Pen a 1 for the predicted rate constant (left). Rate constant versus resolution data for the native Pen a 1 conformation, omitting the 0% reduction data (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 5.6-6.2 nm (see rule set 2 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the native type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=7.539e-11, B=0.1875, C=0.004443, and D=0.02138.

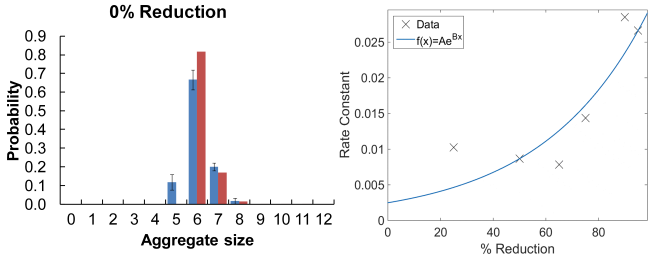

Figure 5. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the S-shaped Pen a 1 for the predicted rate constant (left). Rate constant versus resolution data for the S-shaped Pen a 1 conformation, omitting the 0% reduction data (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 5.8-6.0 nm (see rule set 6 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the S-shaped type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=0.002491 and B=0.02494.

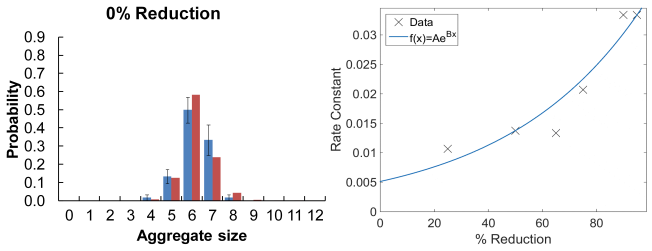

Figure 6. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the U-shaped Pen a 1 for the predicted rate constant (left). Rate constant versus resolution data for the U-shaped Pen a 1 conformation, omitting the 0% reduction data (right). The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 6.8-8.6 nm (see rule set 9 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the U-shaped type Pen a 1. The data points (X-shaped markers) were fitted to an exponential function (boxed equation) which is plotted as a solid line. The equation constants are A=0.005139 and B=0.0197.

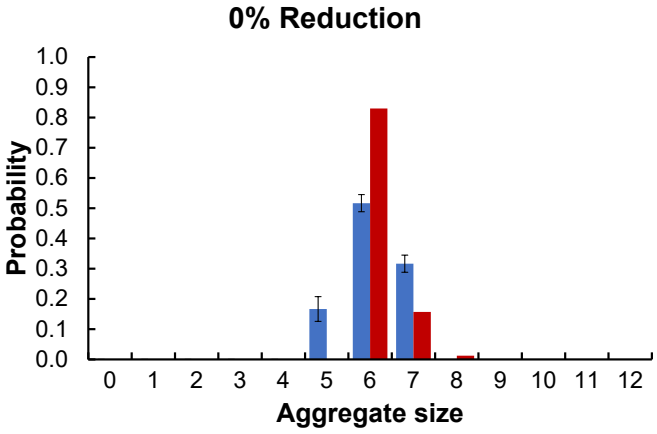

Figure 7. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the native Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 5.6-6.2 nm (see rule set 2 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the native type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.002296 and B=0.029456.

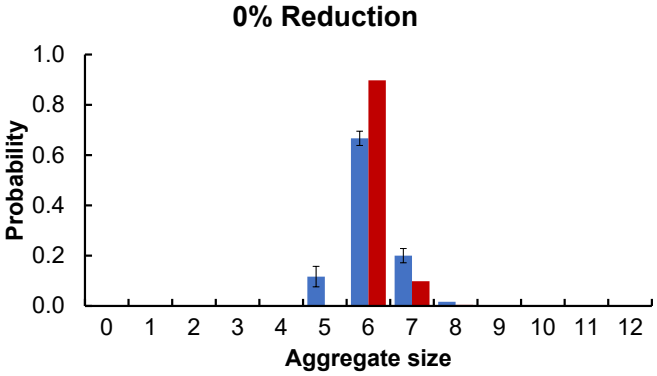

Figure 8. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the S-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 5.8-6.0 nm (see rule set 6 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the S-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.001332 and B=0.031452.

B. Time Analysis

In previous work, our collaborators measured the run time of the Monte Carlo simulation for the models with 0%, 50%, 75%, and 90% polygon reduction [19]. They found that the trend in run time with percentage reduction is linear. We interpolate this trend to determine the run time for the models with 25% and 65% polygon reduction, and we extrapolate to estimate the run time for the model with 95% polygon reduction. We present the polygon reduction percentages and their associated run times for a single run in Table I.

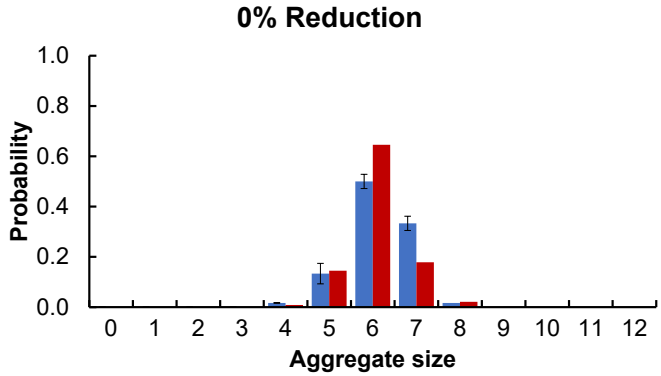

Figure 9. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the U-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 6.8-8.6 nm (see rule set 9 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the U-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.003447 and B=0.023714.

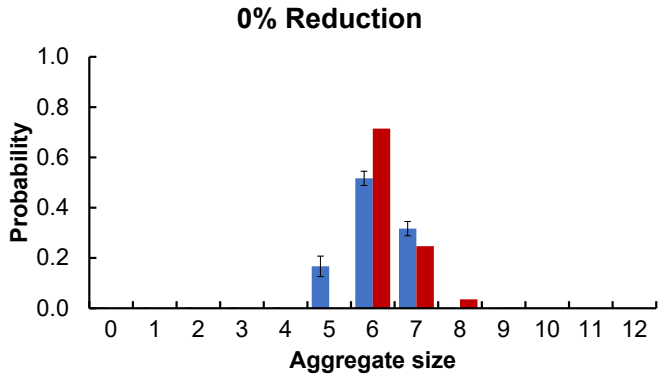

Figure 10. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the native Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 65% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). For these plots, the cutoff distance range has been fixed at 5.6-6.2 nm (see rule set 2 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the native type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.004157 and B=0.022989.

Table I
RUN TIMES FOR A SINGLE RUN OF SEVEN 3D MONTE CARLO MODEL RESOLUTIONS

Polygon Reduction Percentage	Run Time (Hours)
0	20
25	15
50	10
65	7
75	5
90	2
95	1

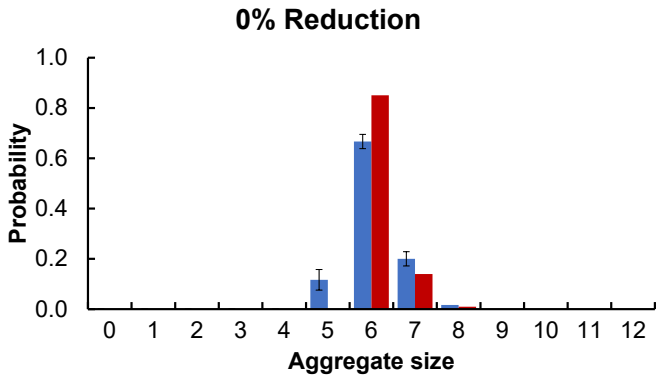

Figure 11. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the S-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 65% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 5.8-6.0 nm (see rule set 6 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the S-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.001992 and B=0.028116.

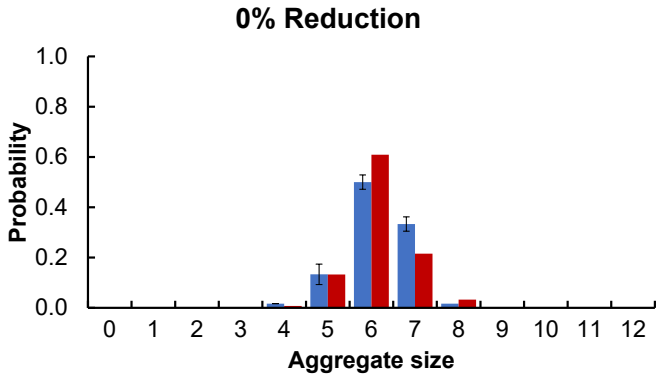

Figure 12. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the U-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, and 65% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 6.8-8.6 nm (see rule set 9 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the U-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.004434 and B=0.021575.

Next, assuming we perform 30 runs for each model resolution, we calculate the total run time needed for each predictive model and list this data, along with the run time needed for a set of 30 full-resolution runs, in Table II. We also include the percentage change of the run time as compared to a set of full-resolution runs, which takes 600 hours.

From Table II, we observe that some of the predictive models are more time-consuming than a set of full-resolution runs. As this defeats the purpose of the predictive model, we should use the predictive models that cost less time to create as long as they provide us with the desired accuracy of results.

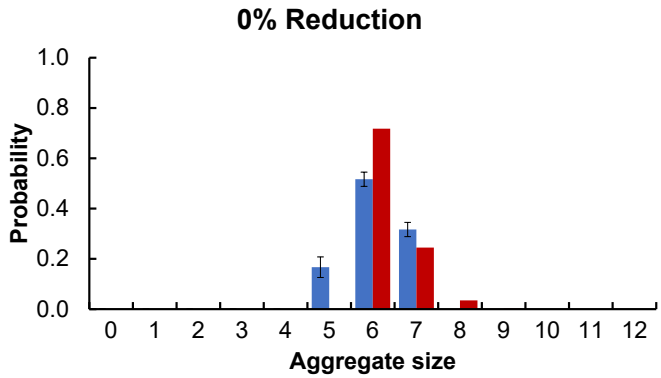

Figure 13. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the native Pen a 1 for the predicted rate constant, omitting the 0%, 25%, 65%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 5.6-6.2 nm (see rule set 2 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the native type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.004106 and B=0.023293.

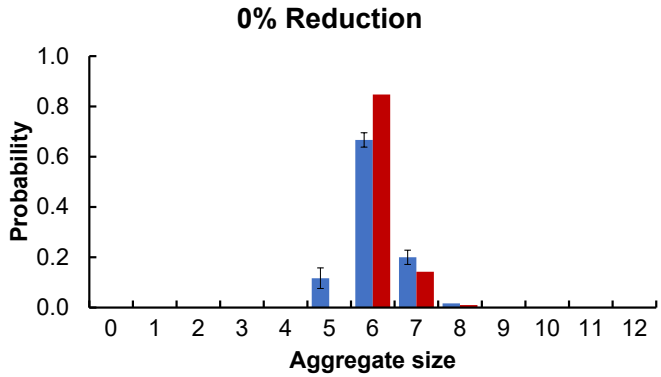

Figure 14. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the S-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, 65%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 5.8-6.0 nm (see rule set 6 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the S-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.002037 and B=0.026951.

Table II
RUN TIMES FOR PREDICTIVE MODELS AND PERCENTAGE CHANGE FROM 600 HOURS (FOR A SET OF 30 RUNS PER RESOLUTION)

Predictive Model Description	Run Time (Hours)	Percentage Change
Full resolution	600	0
Only omit 0% reduction	1200	+100
Omit 0% and 25% reduction	750	+25
Omit 0%, 25%, and 90% reduction	690	+15
Omit 0%, 25%, and 65% reduction	540	-10
Omit 0%, 25%, 65%, and 90% reduction	480	-20

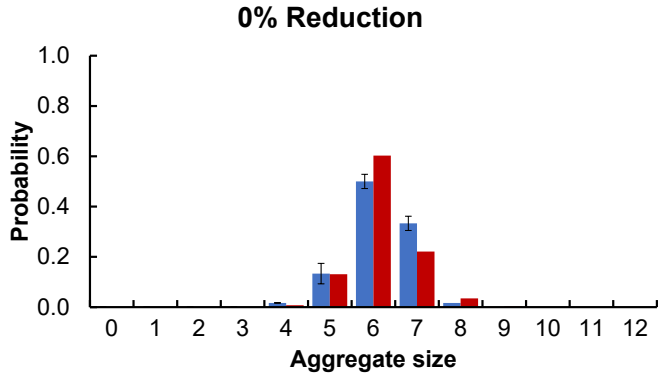

Figure 15. Comparison of Monte Carlo (blue) and rule-based model (red) aggregate size distributions for the 0% reduced resolution of the U-shaped Pen a 1 for the predicted rate constant, omitting the 0%, 25%, 65%, and 90% reduction data. The standard error of the mean was calculated to obtain the error bars for the Monte Carlo model (blue). The cutoff distance range has been fixed at 6.8-8.6 nm (see rule set 9 in the Appendices), which was found to be the optimal range for the 25% reduced resolution of the U-shaped type Pen a 1. The data points were fitted to an exponential function. The equation constants are A=0.004606 and B=0.020665.

V. Conclusions

We proposed a method for developing a predictive model for predicting results for high-resolution Monte Carlo molecular simulations using only lower-resolution Monte Carlo simulations and a biological rule-based model. We found that this method can accurately predict the most probable aggregate sizes of the full-resolution Monte Carlo simulation. We also found that the computational cost of the Monte Carlo simulation can be reduced by up to 20% using this method.

Future work could involve investigating the differences we found between the predictive models for the three conformations of Pen a 1 that we studied. Our results showed that our predictive models for the U-shaped type tended to predict the third most probable aggregate size with significantly greater accuracy than the models for the native and S-shaped types. We want to improve our understanding of the cause of this difference so that we can improve the accuracy of our models.

We can also examine how the number of runs used in the Monte Carlo simulation affects the output. Furthermore, in future research, we will use experimental electron microscopy (EM) data in place of the aggregation data from our Monte Carlo simulations. Cryo-EM data, which provides information about molecular geometry, is now becoming available. The locations of receptor probes in EM images can be used to estimate receptor clustering. We will also consider the effects of allergen cross-linking on our geometric rule-based method.

ACKNOWLEDGMENT

Thank you very much to the peer reviewers for their valuable feedback on this paper. Also, special thanks to Dr. Bruna Jacobson for her role in the early phases of this research project.

REFERENCES

[1] J. Rivera and A. M. Gilfillan, "Molecular regulation of mast cell activation," *Journal of Allergy and Clinical Immunology*, vol. 117, no. 6, pp. 1214–1225, 2006.

[2] R. Valenta, B. Linhart, I. Swoboda, and V. Niederberger, "Recombinant allergens for allergen-specific immunotherapy: 10 years anniversary of immunotherapy with recombinant allergens," *Allergy*, vol. 66, pp. 775–783, 2011.

[3] B. Hoard, B. Jacobson, K. Manavi, and L. Tapia, "Extending rule-based methods to model molecular geometry," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2015, pp. 587–594.

[4] ——, "Extending rule-based methods to model molecular geometry and 3D model resolution," *BMC Systems Biology*, vol. 10, no. 48, 2016.

[5] J. Yang, M. I. Monine, J. R. Faeder, and W. S. Hlavacek, "Kinetic Monte Carlo method for rule-based modeling of biochemical networks," *Phys. Rev. E*, vol. 78, no. 3, p. 031910, 2008.

[6] L. A. Chylek, L. A. Harris, C.-S. Tung, J. R. Faeder, C. F. Lopez, and W. S. Hlavacek, "Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems," *WIRESBM*, vol. 6, pp. 13–36, 2014.

[7] V. Danos and C. Laneve, "Formal molecular biology," *Theoretical Computer Science*, vol. 325, no. 1, pp. 69–110, 2004.

[8] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek, "BioNetGen: Software for rule-based modeling of signal transduction based on the interactions of molecular domains," *Bioinformatics*, vol. 20, no. 17, pp. 3289–3291, 2004.

[9] J. F. Ollivier, V. Shahrezaei, and P. S. Swain, "Scalable rule-based modelling of allosteric proteins and biochemical networks," vol. 6, no. 11, Nov. 2010.

[10] C. Maus, S. Rybacki, and A. M. Uhrmacher, "Rule-based multi-level modeling of cell biological systems," *BMC Systems Biology*, vol. 5, p. 166, 2011.

[11] M. Stefan, T. Bartol, T. Sejnowski, and M. Kennedy, "Multi-state modeling of biomolecules," *PLoS Computational Biology*, vol. 10, no. 9, p. e1003844, 2014.

[12] G. Gruenert, B. Ibrahim, T. Lenser, M. Lohel, T. Hinze, and P. Dittrich, "Rule-based spatial modeling with diffusing, geometrically constrained molecules," *BMC Bioinformatics*, vol. 11, p. 307, 2010.

[13] R. Kerr, T. Bartol, B. Kaminsky, M. Dittrich, J. Chang, S. Baden, T. Sejnowski, and J. Stiles, "Fast Monte Carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces," *SIAM J. Sci. Comput.*, vol. 30, no. 6, pp. 3126–3149, 2009.

[14] D. P. Tolle and N. L. Novère, "Meredys, a multi-compartment reaction-diffusion simulator using multistate realistic molecular complexes," vol. 4, no. 24, 2010.

[15] H. W. Hatch, J. Mittal, and V. K. Shen, "Computational study of trimer self-assembly and fluid phase behavior," *The Journal of Chemical Physics*, vol. 142, no. 16, 2015.

[16] J. Zhang, K. Leiderman, J. R. Pfeiffer, B. S. Wilson, J. M. Oliver, and S. L. Steinberg, "Characterizing the topography of membrane receptors and signaling molecules from spatial patterns obtained using nanometer-scale electron-dense probes and electron microscopy," *Micron*, vol. 37, no. 1, pp. 14–34, 2006.

[17] N. L. Andrews, J. R. Pfeiffer, A. M. Martinez, D. M. Haaland, R. W. Davis, T. Kawakami, J. M. Oliver, B. S. Wilson, and D. S. Lidke, "Small, mobile FcεRI receptor aggregates are signaling competent," *Immunity*, vol. 31, no. 3, pp. 469–479, 2009.

[18] K. Xu, B. Goldstein, D. Holowka, and B. Baird, "Kinetics of multivalent antigen DNP-BSA binding to IgE-FcεRI in relationship to the stimulated tyrosine phosphorylation of FcεRI," vol. 160, no. 7, pp. 3225–3235, 1998.

[19] K. Manavi, B. Jacobson, B. Hoard, and L. Tapia, "Influence of model resolution on geometric simulations of antibody aggregation," *Robotica*, vol. 34, no. 8, pp. 1754–1776, 2016.