Influence of Model Resolution on Antibody Aggregation Simulations

Kasra Manavi and Lydia Tapia
Department of Computer Science
University of New Mexico
Albuquerque, New Mexico, 87131
Email: kazaz@cs.unm.edu, tapia@cs.unm.edu

***Abstract*—Up to 40% of the world's population suffer from allergies. A primary mediator for allergies is the aggregation of antigens and IgE antibodies bound to cell-surface receptors FcϵRI. Antibody/antigen aggregate formation causes stimulation of mast cells and basophils, initiating cell degranulation and resulting in the release of immune mediators which produce an allergic or anaphylactic response. Understanding of the shape and structure of aggregates can provide critical insights into allergic response. However, due to the large size and number of molecules involved in aggregation, traditional techniques such as MD and coarse grained energetic models are computationally infeasible. Alternative methods such as ODEs and rule-based models are able to simulate the process of aggregation, however, these methods exclude critical geometric details.**

In our previous work, we presented methods to geometrically model, simulate and analyze antibody aggregation inspired by rigid body robotic motion simulation. Our polygon-based models of antibody-receptor complexes and antigens capture the 3-D shape of molecular structures while reducing structural complexity. Due to our simulation techniques, the number of polygon interaction comparisons are directly impacted by geometric model complexity. This is similar to collision detection calls in robotic motion planning, a primitive operation that has major implications for run time. Recent advances in polygon reduction techniques allows us to reduce the complexity of the models involved in the simulation. However, reductions could translate into qualitative changes in the molecules being simulated. In this paper we analyze the impact of model resolution on our simulations of antibody aggregation. Our exploration is focused on two antigens, a trivalent man-made antigen and a common shrimp allergen.

I. Introduction

The ability to computationally design and predict large scale molecular structures and their components offer ways to investigate nano-scale phenomena where experimental methods have difficulties. Methods developed to design molecules focus on molecular cages and 2-D/3-D crystalline arrays [12], whereas predicting large scale structures focus on protein or polymer aggregation [2, 17]. Computational methods have been developed to design protein-protein interfaces [7, 10] as well as self-assembling molecules [11].

The human allergy immune response is caused by cell surface molecule aggregation. This immune responses is triggered by a tyrosine kinase cascade initiated by the transmembrane signalling caused by the crosslinking of IgE-FcϵRI (receptor complexes) via antigen (ligand) binding. This crosslinking stimulates mast cells and basophils and induces degranulation, resulting in the release of histamine and other allergic response mediators [22]. The ability to predict the structures of these aggregates is key in determining how the spatial organization of receptor complexes affect transmembrane signaling.

The geometric impacts of molecular aggregation have not been well studied due to the fact that it is presently computationally infeasible to model hundreds of large molecules at an all-atom resolution. Methods such as ODEs or rule-based modeling can simulate the aggregation process, do not provide any geometric details of aggregates [5, 26]. To address this problem, we have previously presented methods in [14] to simulate and analyze receptor complex aggregation using polygon-based models. The geometric complexity is an important feature of the molecular models. Lower complexity translates into fewer polygon to polygon comparisons and faster run times. On the other hand, the reduction in geometric detail impacts model realism. In this paper we explore the impacts model resolution has on our aggregation simulation. We present results on one synthetic ligand that has been studied experimentally and one natural allergen with some experimental analysis. These ligands greatly differ in size, structure and valency.

II. Related Work

Design of Molecular Structures. Methods for designing protein based assemblies come in two forms, stochastic (resulting in irregular structures with probability-derived attributes) and deterministic (producing exactly specified geometric features). Principles for the design of ordered protein assemblies are discussed in [12]. A majority of the computational design methods have focused on interface construction [7, 10], but newer methods go further and fully design self-assembling molecules [11]. These methods are similar in that they generally start out performing rigid body docking followed by iterative design/minimization steps to refine the interface.

Molecular Aggregation Prediction. Many fields stand to benefit from the use of computational methods to deter unintended aggregation. Most of the methods developed use lattice models with force fields and focus on the interactions of proteins with both denatured [2] and native [27] conformations. A coarse grained molecular dynamics based approach to study polymer-drug aggregation was done in [17]. We note all of these models have energetics included in their computation

and would not be feasible to simulate a system of the size we do.

IgE Aggregation Experiments. Studies using nanoparticles have shown ligand size and valency impact degranulation of RBL mast cells [8]. Spatiotemporal analysis of IgE aggregation has been done using nanoscale imaging and motion tracking techniques. The locations of static gold nanoparticle labeled IgE-FcϵRI have been imaged using transmission electron microscopy [24]. Tracking of quantum dot labeled IgE-FcϵRI has determined temporal information such as diffusion rates [1]. While these experimental methods have been able to measure attributes about receptor dynamics, neither retain information about the aggregate binding patterns. Because of this, distinguishing bound from simply proximal receptors is challenging.

IgE Aggregation Models & Simulations. There have been many methods developed to model ligand-receptor interactions. A majority of these methods focused on systems comprised of trivalent ligands and bivalent receptors. One initial model was based on thermodynamic equilibrium, taking into account interactions between free ligand and free receptors and between crosslinking receptors via ligand [5]. An updated kinetics based version of this model was introduced in [26] and was extended to consider steric constraints in [16].

Ligand Structure and Valency. There have been numerous studies of natural allergens initiating degranulation. One that has been had particular interest is the common shrimp allergen, Pen a 1. The immune response to Pen a 1 is triggered by the shrimp tropomyosin molecule, a 40 nm coiled coil structured, (Figure 1), which crosslinks IgE. The allergen has been predicted to have 5 binding regions [19] and a total of 16-18 binding sites [9, 21].

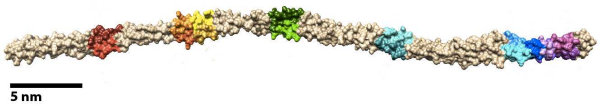

Fig. 1: The molecular structure of Pen a 1, the common shrimp allergen (tan). The binding sites (various colors) are located in 5 regions on the coiled coil structure.

Alongside experimentation on natural antigen, synthetic ligands have been generated to provide insights into antibody aggregation [18, 23, 25]. These synthetic ligands are primed with DNP, a linker used to bind to engineered antibodies. We explore the ligand DF3, a trivalent molecule with long, extended linkers protruding from the molecule (Figure 2).

III. Methods

A. Model Construction

Our method is based on simulating 3-D rigid body models of the receptor complex and ligand molecules. We use our own generated models as well as contributed models. An all-atom structure of the receptor complex was used from [13]. The IgE structure, composed of both heavy and light chains are modeled bound to the α subunit of FcϵRI. The receptor complex is made up of 1,709 amino acids totaling 13,477 atoms.

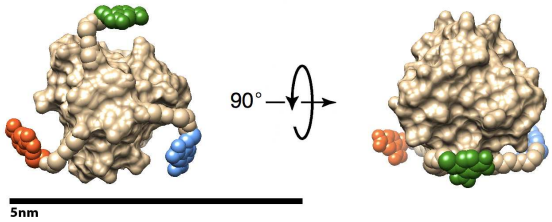

Fig. 2: The molecular structure of DF3 (tan), a synthetic ligand. The fibritin trimer has 3 DNP linkers (various colors) attached to the N-terminus of each subunit of the trimer.

The ligand DF3 was generated by starting with the base fibritin trimer (PDB:1RFO) and adding DNP linkers to the N-terminus of the each fibritin subunit. DF3 is comprised of 81 amino acids with 1,365 atoms total. A model structure for Pen a 1 was available in the Structural Database of Allergenic Proteins (SDAP Model #284). This tropomyosin model was composed of 568 amino acids totalling 4,580 atoms.

Since it would be computationally prohibitive to use these all-atom models at the molecule counts we simulate, we reduce the complexity of the model (Figure 3). To construct our models, we begin with an all-atom structure. Using the multiscale model extension of UCSF Chimera [4], we generate isosurface models of the molecules at a variety of resolutions, 4Å for DF3 and 6Å for the receptor complex and Pen a 1. The resulting model of the occupied volume, referred to henceforth as the base model, is considered to be the model with the highest resolution, i.e., the most detailed model.

This base model can now be reduced in complexity using standard polygon reduction techniques [3] Since the amount of detail in the base model hinders performance, we want to observe the costs versus benefits of decreasing the model resolution. We the polygon reduction algorithm in Maya [15], a modeling software package, that allows the generation of models with a specified reduction in the percentage of polygons.

The ligand binding sites are unique to each ligand, and thus had to be modeled accordingly. The model of DF3 with DNP linkers has very flexible binding sites [13]. The model has the linkers compressed to half their length and has a binding site located at the end of the linker. We model a spherical binding volume with a radius of half the DNP linker length (7.5Å) that is centered at the end of the compressed linker.

The binding sites of Pen a 1 are located on its surface. The location of binding sites were determined by surface amino acids of the ligand epitopes [9]. Vertices located on the center

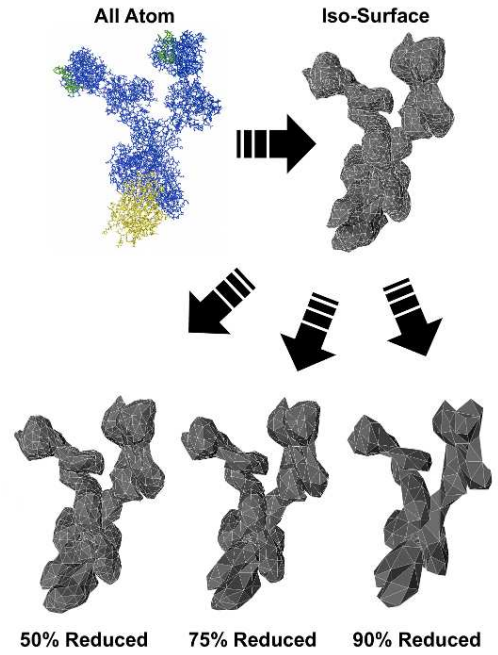

Fig. 3: The model construction process starting with an all-atom models, generating the isosurface model, and then applying polygon reduction to the isosurface model to generate models with lower resolution.

of the epitopes and on the exterior of the surface were used and a binding radius of 3Å was used. The same epitope locations were used regardless of the resolution since the locations of the binding sites on the original all-atom structure doesn't change.

B. Simulation Methods

As outlined in our previous publication [14], the simulation begins with all molecules randomly placed on a grid in a collision free state with no molecules bound. At each time step, every molecule gets an updated position and orientation generated via random sampling. The updated positions and orientations are generated with the consideration of biological constraints, e.g., diffusion constants [1] and rotational correlation times [20]. Over the course of the simulation a receptor binding site will end up within the binding volume of the ligand. The probability of a binding event occurring is dependent on the association rates specified. Dissociation rates are specified as well, each bond at every time step is evaluated for dissociation. We note that DNP linkers have been well studied and have values for association/disassociation rates. These parameters are not known for Pen a 1, thus we can not determine kinetics from our simulation, only possible packing structures.

C. Aggregate Model and Analysis

We define the state of the system as graph $G\{V, E\}$ where V is the set of molecules and E is the set of edges. Given two vertices $\{v_l, v_r\} \in V$, an undirected edge $e\{v_l, v_r\} \in E$ if and only if v_l and v_r are bound. Since ligands only bind to receptors and vice versa, the graph is bipartite. Data about the overall aggregation process can be extracted from the graph.

To gain further insights into the aggregate structures, we take our modeled aggregate structures and generate all-atom structures. With these all-atom models we can take measurements of the aggregate structure and analyze features of the aggregate such as steric hindrance and can quantify model construction quality.

IV. Experimental Setup

We simulate a discrete patch of membrane 200 nm x 200 nm (40,000 nm^2) with non-periodic boundaries. We simulate 24 receptors for all our experiment, resulting in a density of $\sim$600 receptors/μm^2 In two different experiments, we simulate 12 DF3 and 1 Pen a 1 ligand molecules. We used the diffusion constant 0.09μm^2/s of IgE-FcϵRI found in [1] for all molecules. We use a time step of 10 μs and run experiments 500,000 time steps, long enough for the simulations to reach a stable state. Association and disassociation rates of 1.0 and 0.01 s^{-1}, from [25] were used for both ligands. As aggregates get larger, they have been shown to slow down [1]. This is incorporated into the simulation by diffusing aggregates inversely proportional to their size. Rotations of large aggregates are limited by the diffusion rate of the fastest moving receptor of the aggregate. The receptor furthest from the center of the aggregate is limited by its diffusion constant, thus limiting the amount of rotation an aggregate can make. Simulation code was developed using the Parasol Motion Planning Library (PMPL). Simulations were run on a super computer housed at UNM utilizing single cores of Intel Xeon E5645 processors with 4 GB of ram per processor. Thirty (30) runs of each experiment were performed.

V. Results

A. Volume and Timing

We begin with analyzing the impact of polygon reduction on the volume of the model. Table I shows the number of polygons and volume for each model. The polygon reduction algorithm works by specifying a percentage of the polygons to reduce, leading to the close correspondence between the reduction percentage and the number of polygons. We find that volumes decrease with increased reduction. Such decrease is expected, and can be quite dramatic (nearly 50% for 90% reduced Pen a 1). We note that volume reduction does not necessarily mean less realistic results, "soft docking" approaches [6] allow a certain degree of inter-protein penetration to approximate flexibility given rigid structures.

As seen in Figure 4, the reduction in polygons has a clear effect on runtime. We see a linear increase in runtime vs. model polygon count. This is due in part to the nature of the rigid body modeling, collision detection is a major factor in computation time and is highly dependent on model complexity. We attribute binding site interaction calculation to the difference in slope between the two lines.

TABLE I: Model reduction statistics including polygon counts and volumes of the molecular models generated at a variety of resolutions.

Molecule Name	Model Property	Model Percent Reduction 0%	50%	75%	90%
Receptor	Polygons (#)	4876	2438	1216	490
	Volume (nm^2)	234.98	227.90	208.73	162.31
	Volume (%)	100.00	96.99	88.83	69.07
DF3	Polygons (#)	1208	604	302	120
	Volume (nm^2)	15.83	14.90	13.16	9.74
	Volume (%)	100.00	94.13	83.13	61.53
Pen a 1	Polygons (#)	2328	1164	582	234
	Volume (nm^2)	51.60	49.95	44.86	28.80
	Volume (%)	100.00	96.80	86.94	55.81

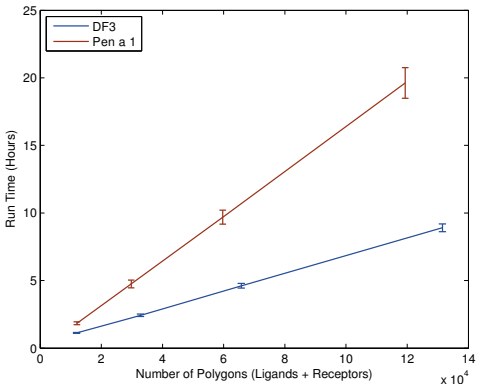

Fig. 4: Runtimes of the different resolutions of the same model. The x-axis is the sum of the number of polygons used to describe all of the models for any given experiment. The y-axis is the runtime in hours.

B. Impact on Quality of Results

Simulations are run until a stable state is reached. Figure 5 shows the number of edges in G for DF3 in blue. We see that for DF3, all of the reductions generate similar numbers of connections. The more reduced models produce slightly more connections but all the values are very close. The mean of each reduction is contained in the overlap of the standard deviation of all reductions.

In Figure 6 we see that Pen a 1 model resolution has more of an impact on the number of connections that are made. The 90% reduced model made nearly 2 more connection than the unreduced model for a single ligand. This is one of the side effects of reducing model volume. With the reduction there is more open volume around the binding site, reducing steric hindrance of receptors trying to bind to sites in the same or adjacent regions.

To further analyze the implications of model reduction we plot a histogram of aggregate size versus percentage of occurrence. We see in Figure 7 that there is minimal impact on aggregate size distributions for the DF3 experiment. The distribution for each model reduction seem to be the same.

This is not the case for Pen a 1, seen in Figure 8. The

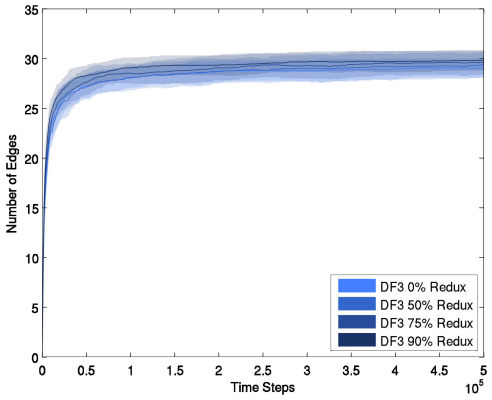

Fig. 5: The x-axis is simulation time step and the y-axis is the number of edges in G. DF3 connections do not seem to be affected much by model resolution.

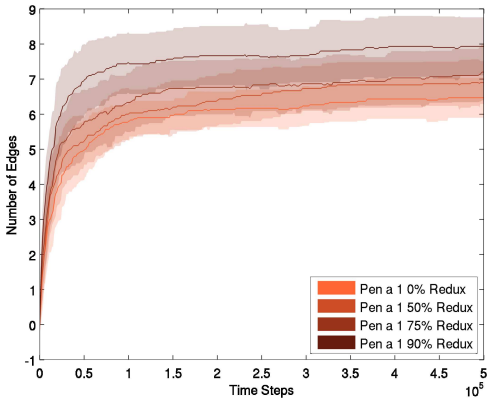

Fig. 6: The x-axis is simulation time step and the y-axis is the number of edges in G. Pen a 1 binding is affected by the use of different resolutions, lower resolutions generate more connections than higher resolutions.

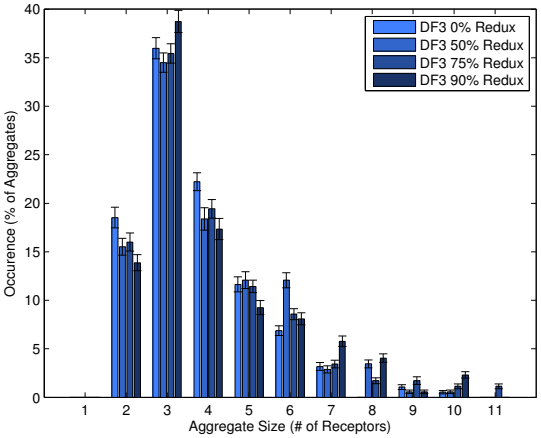

Fig. 7: Aggregate size vs. percentage of aggregates of that size. The different resolutions do not affect the distribution of aggregate sizes for DF3.

distribution has the two least reduced models peaking near aggregates of size 7 whereas the two most reduced models

peak near aggregates of size 8. This is attributed to the volume reduction of the model which in effect relaxes the steric constraints. With a smaller volume, more free space is available for a molecule to pack into a tight space within a given aggregate.

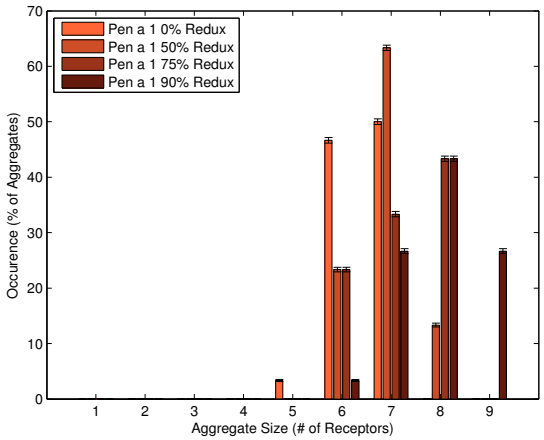

Fig. 8: Aggregate size vs. percentage of aggregates of that size. Aggregate size seems to be dependent on model resolution, lower resolution Pen a 1 models produce larger aggregates.

Finally, we investigate the impact of model reduction on the aggregate structure. In order to evaluate this, after aggregates are constructed with low-resolution polygon models, we construct the corresponding all-atom structure. However, since the polygon models are much simpler than the all atom structures, there may be unintended interactions. For example, when non-bonded atoms are too close, repulsion occurs due to van der Waals interactions. In order to evaluate the possible effects of transitioning between polygon and all-atom models, we counted the number of C_αs and DNP linker carbon rings within 7Å for IgE-FcϵRI and DF3. For Pen a 1, distances were calculated between C_αs for the aggregated molecules. In order to indicate these proximal non-binding residues, we refer to them as *potential collisions*. Also, ligand binding site epitopes are not included in the enumeration.

TABLE II: Percentage (%) of residues that exhibit a *potential collision*. Ligand residues involved in binding are not included.

Ligand Simulated	Model Percent Reduction 0%	50%	75%	90%
DF3	0.0074	0.0122	0.0215	0.1016
Pen a 1	0.0158	0.0350	0.0816	0.2238

We can see from the results in Table II that model resolution does have an impact on the number of *potential collisions* that exist in aggregate structures. *Potential collision* residues increase as resolution decreases. We see that DF3 is not significantly impacted by model resolution up to 50%. However, at 90% there is a relatively large increase in *potential collisions*. This trend exists for Pen a 1 as well, however, it is not as pronounced.

We see that Pen a 1 model reduction generally has a higher percentage of residues in *potential collision* compared to DF3. This is attributed to the flexibility of the DF3 binding site. The DNP linker has a large, relatively open volume that can be bound. The binding sites of Pen a 1 are smaller in volume since they are on the molecular surface and are partially occupied by the molecular volume. Therefore, the antibodies have to be closer to the allergen surface in Pen a 1.

We note that overall, the number of residues in *potential collision* is minimal. Even at 90% reduced for the Pen a 1, aggregates generated have about 0.2% of residues in *potential collision*. These interactions could be first addressed though locally evaluated energetics and perturbations. An example Pen a 1 all-atom aggregate is shown in Figure 9.

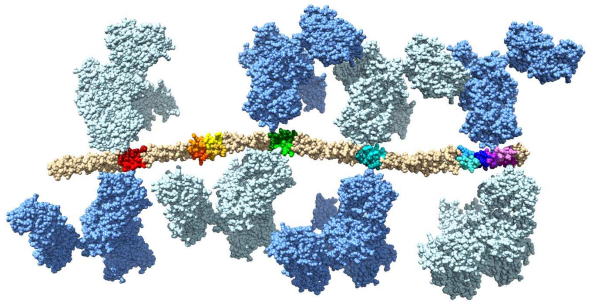

Fig. 9: A resulting aggregate structure generated using our method. The 8 IgE-FcϵRI (light/medium blue) are bound to the Pen a 1 ligand (tan) at various binding site on the antigen (various colors)

VI. Conclusion

We investigated the impact of model resolution on simulations of antibody aggregation. We generated models of both synthetic and natural allergens at a variety of resolutions and analyze the generated aggregates. We found that model reduction affects fundamental attributes such as the polygon count, volume change, and simulation time in predictable ways.

Acknowledgments

This work supported in part by the National Institutes of Health (NIH) Grant P50GM085273 supporting the New Mexico Spatiotemporal Modeling Center and NIH Grant P20RR018754 supporting the Center for Evolutionary and Theoretical Immunology. Special thanks to Chang-Shung Tung from Los Alamos National Laboratories for providing structures and the UNM Center for Advanced Research Computing for providing the computational resources used in this work.

References

[1] Nicholas L. Andrews, Janet R. Pfeiffer, A. Marina Martinez, David M. Haaland, Ryan W. Davis, Toshiaki Kawakami, Janet M. Oliver, Bridget S. Wilson, and Diane S. Lidke. Small, mobile FcϵRI receptor aggregates are signaling competent. *Immunity*, 31:469–479, 2009.

[2] Dusan Bratko, Troy Cellmer, John M. Prausnitz, and Harvey W. Blanch. Molecular simulation of protein aggregation. *Biotech. and Bioeng.*, 96(1):1–8, 2007.

[3] P. Cignoni, C. Montani, and R. Scopigno. A comparison of mesh simplification algorithms. *Computers & Graphics*, 22(1):37 – 54, 1998.

[4] Thomas D. Goddard, Conrad C. Huang, and Thomas E. Ferrin. Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. *Structure*, 13(3):473 – 482, 2005.

[5] B. Goldstein and A.S. Perelson. Equilibrium theory for the clustering of bivalent cell surface receptors by trivalent ligands. Application to histamine release from basophils. *Biophysical Journal*, 45(6):1109 – 1123, 1984.

[6] Inbal Halperin, Buyong Ma, Haim Wolfson, and Ruth Nussinov. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, 47(4):409–443, 2002.

[7] Po-Ssu Huang, John J. Love, and Stephen L. Mayo. A de novo designed protein-protein interface. *Protein Science*, 16(12):2770–2774, 2007.

[8] Yu-Fen Huang, Haipeng Liu, Xiangling Xiong, Yan Chen, and Weihong Tan. Nanoparticle–mediated ige–receptor aggregation and signaling in rbl mast cells. *Journal of the American Chemical Society*, 131(47): 17328–17334, 2009.

[9] Ovidiu Ivanciuc, Catherine H. Schein, and Werner Braun. SDAP: Database and computational tools for allergenic proteins. *Nucleic Acids Research*, 31(1):359–362, 2003.

[10] Ramesh K. Jha, Andrew Leaver-Fay, Shuangye Yin, Yibing Wu, Glenn L. Butterfoss, Thomas Szyperski, Nikolay V. Dokholyan, and Brian Kuhlman. Computational design of a $\{$PAK1$\}$ binding protein. *Journal of Molecular Biology*, 400(2):257 – 270, 2010.

[11] Neil P. King, William Sheffler, Michael R. Sawaya, Breanna S. Vollmar, John P. Sumida, Ingemar Andr, Tamir Gonen, Todd O. Yeates, and David Baker. Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science*, 336(6085):1171–1174, 2012.

[12] Yen-Ting Lai, Neil P. King, and Todd O. Yeates. Principles for designing ordered protein assemblies. *Trends in Cell Biology*, 22(12):653 – 661, 2012. Special Issue Synthetic Cell Biology.

[13] Avanika Mahajan, Dipak Barua, Patrick Cutler, Diane S. Lidke, Flor A. Espinoza, Carolyn Pehlke, Rachel Grattan, Yuko Kawakami, Chang-Shung Tung, Andrew R. M. Bradbury, William S. Hlavacek, and Bridget S. Wilson. Optimal aggregation of fcϵri with a structurally defined trivalent ligand overrides negative regulation driven by phosphatases. *ACS Chemical Biology*, May 2014.

[14] Kasra Manavi, Bridget S. Wilson, and Lydia Tapia. Simulation and analysis of antibody aggregation on cell surfaces using motion planning and graph analysis. *Proc. ACM Conf. on Bioinfo., Comp. Bio. and Biomed. (ACM-BCB)*, pages 458–465, 2012.

[15] Autodesk Maya. URL http://www.autodesk.com/.

[16] Michael I. Monine, Richard G. Posner, Paul B. Savage, James R. Faeder, and William S. Hlavacek. Modeling multivalent ligand-receptor interactions with steric constraints on configurations of cell-surface receptor aggregates. *Biophysical Journal*, 98:48–56, 2010.

[17] Lili X. Peng, Lei Yu, Stephen B. Howell, and David A. Gough. Aggregation properties of a polymeric anticancer therapeutic: A coarse-grained modeling study. *Journal of Chemical Information and Modeling*, 51(12):3030–3035, 2011.

[18] Richard G. Posner, Kala Subramanian, Byron Goldstein, James Thomas, Toni Feder, David Holowka, and Barbara Baird. Simultaneous cross-linking by two nontriggering bivalent ligands causes synergistic signaling of IgE FcRI complexes. *J. of Immunology*, 7:3601–3609, 1995.

[19] Ayuso R, B Lehrer S, and Reese G. Identification of continuous, allergenic regions of the major shrimp allergen pen a 1 (tropomyosin). *Int Arch Allergy Immunol*, 127: 27–37, 2002.

[20] N.A. Rahman, I. Pecht, D.A. Roess, and B.G. Barisas. Rotational dynamics of type i fc epsilon receptors on individually-selected rat mast cells studied by polarized fluorescence depletion. *Biophysical Journal*, 61(2):334 – 346, 1992.

[21] Gerald Reese, Julia Viebranz, Susan M. Leong-Kee, Matthew Plante, Iris Lauer, Stefanie Randow, Mar San-Miguel Moncin, Rosalia Ayuso, Samuel B. Lehrer, and Stefan Vieths. Reduced allergenic potency of vr9-1, a mutant of the major shrimp allergen pen a 1 (tropomyosin). *The Journal of Immunology*, 175(12): 8354–8364, 2005.

[22] Juan Rivera and Alasdair M. Gilfillan. Molecular regulation of mast cell activation. *Journal of Allergy and Clinical Immunology*, 117(6):1214 – 1225, 2006.

[23] Dwaipayan Sil, Jong Bum Lee, Dan Luo, David Holowka, and Barbara Baird. Trivalent ligands with rigid dna spacers reveal structural requirements for ige receptor signaling in rbl mast cells. *ACS Chemical Biology*, 2(10):674–684, 2007.

[24] Bridget S. Wilson, Janet M. Oliver, and Diane S. Lidke. Spatio-temporal signaling in mast cells. *Advances in Experimental Medicine and Biology*, 716:91–106, 2011.

[25] K. Xu, B. Goldestein, D. Holowka, and B. Baird. Kinetics of multivalent antigen DNP-BSA binding to IgE-FceRI in relationship to the stimulated tyrosine phosphorylation of FceRI. *J. Immunol.*, 160:3225–3235, 1998.

[26] Jin Yang, Michael I. Monine, James R. Faeder, and William S. Hlavacek. Kinetic monte carlo method for rule-based modeling of biochemical networks. *Phys. Rev. E*, 78, Sep 2008.

[27] Lin Zhang, Diannan Lu, and Zheng Liu. How native proteins aggregate in solution: A dynamic monte carlo simulation. *Biophys. Chem.*, 133(13):71 – 80, 2008.