Geometrical Insights into the Process of Antibody Aggregation

Kasra Manavi, Alan Kuntz, Lydia Tapia
Department of Computer Science
University of New Mexico
Albuquerque, NM 87131

Abstract
IgE antibodies bound to cell-surface receptors, FcεRI, crosslink through the binding of antigens on the cell surface. This formation of aggregates is what stimulates mast cells and basophils in order to initiate degranulation, resulting in an allergic response. Nearly 1,500 Americans die each year from anaphylactic shock precipitated by aggregation. Experimental studies have shown the spatial organization of the aggregated IgE-FcεRI complexes affects transmembrane signaling that initiates allergic response. There are many factors that can affect the shape and size of aggregates. However, one critical factor may be the conformational structure of the antigen (ligand). This structure can affect the number of receptors that can bind to a single ligand, e.g., the valency of the ligand. For example, a common hay fever antigen has a valency of four while the common shrimp antigen has a valency of eighteen.

3-D simulation of hundreds of antibodies aggregating can be computationally infeasible. However, we present methods based on robotic representations of molecular structures and Monte Carlo simulation that provide 3-D details of aggregate formation. In this paper, we demonstrate the utility of our methods on ligands of different valences: a bivalent DCT-cys (DCT) and a trivalent fibrinogen trimmer (DF3). We show that we can capture experimentally measured properties while enabling a detailed look into the geometry of aggregation formation.

Introduction
It is predicted that nearly 40% of the world’s population suffers from allergies (Pawankar et al. 2012) and each year in the United States, about 1,500 people die from anaphylactic shock (Sompayrac 2012). This is caused by a tyrosine kinase cascade initiated by antigen-mediated crosslinking of IgE antibodies, bound to FcεRI receptors. This stimulates both mast cells and basophils to then release histamine and other mediators of allergic reactions (Rivera and Gilfillan 2006). Previous studies suggest the spatial organization of clustered IgE-FcεRI complexes is what affects the transmembrane signaling which initiates these responses (Fewtrell and Metzger 1980; Wilson et al. 2001; Wilson, Pfiffer, and Oliver 2000; Wilson, Oliver, and Lidke 2011). It has also been shown that the size of aggregates plays a role in the strength of transmembrane signaling (Posner et al. 1995). In theory, the geometry and valency of a ligand will affect the size and structure of resulting aggregates. Synthetic trivalent ligands were revealed to play a major role in the structural requirements for IgE aggregation, and experimentation has been done regarding receptor signaling in Rat Basophilic Leukemic (RBL) mast cells (Sil et al. 2007). However, there exist other ligands which vary in valency and structure that initiate the same response (Metzger 1992; Posner et al. 1995; Xu et al. 1998).

We presented these methods in (Manavi, Wilson, and Tapia 2012). In this paper we extend our previous results with ligands of varied valency. The method works as follows. First, 3-D rigid-body representations are made of the molecular structures involved in aggregation. These models are used in a Monte Carlo approach with relaxed constraints to study the aggregation process. Techniques based on random configuration space sampling in robotic motion planning are used in order to find valid molecular placements during the simulation. Projected 2-D graphs of molecules (nodes) and bindings (edges) are used to analyze aggregates.

We show results on two synthetic ligands that have been studied experimentally. These ligands are similar in size. However, they vary in valency (bivalent and trivalent). We show that our simulations are able to capture experimentally derived results such as time to aggregate convergence and the size of the resulting aggregates. We then look at the effects valency has on aggregate structure by analyzing resulting aggregates from each ligand.

Related Work
In this section, we review the related work in motion planning of molecules along with IgE aggregation studies, both experimental and computational.

Motion Planning
Motion planning (Reif 1979), the problem of finding a valid, collision-free path for a given robot and environment, is a significant problem in robotics. Motion planning has been studied extensively and been applied to a wide array of domains, including molecular motions (Singh, Latombe, and Brutlag 1999; Amato, Dill, and Song 2003).
In motion planning, a robot is decomposed into parameters used to describe a pose, or configuration, of the robot in an environment. These $d$ parameters define the robot’s position and the orientation of the robot’s joints, referred to as degrees of freedom (DOFs). We view these $d$ parameters as $d$-dimensional continuous space and call this space configuration space, or $C_{space}$ (Lozano-Pérez and Wesley 1979). $C_{space}$ can be divided into two regions, $C_{free}$, the set of all valid configurations and $C_{obst}$, all invalid or infeasible configurations. Finding a valid, collision-free path can be seen as finding a path completely within $C_{free}$ from a specified start configuration to a goal configuration. For our work, no particular goal configuration is specified, instead we explore $C_{free}$ until some convergence criteria is met.

A wide array of methods have been developed to study molecular motions. A majority of these methods are based on Monte Carlo simulations with discrete time intervals (Covell 1992; Kolinski and Skolnick 1994). Problems such as protein folding (Amato, Dill, and Song 2003; Thomas et al. 2007; Tapia, Thomas, and Amato 2010), RNA folding (Tang et al. 2008), and ligand binding (Singh, Latombe, and Brutlag 1999; Bayazit, Song, and Amato 2001) have been approached using motion planning techniques.

**IgE Aggregation Experiments**

Spatial/temporal studies of IgE aggregation require nanoscale resolution imaging of cell membrane sheets. Studies have shown that ligand size and valency impact signalling in RBL mast cells (Huang et al. 2009). Gold nano-particle labeled IgE-FcεRI have been imaged on cell membranes using Transmission Electron Microscopy (TEM) techniques (Wilson, Oliver, and Lidke 2011). This method provides the locations of IgE-FcεRI on the membrane, but does not provide temporal information. Temporal properties such as molecule speed have been determined using Quantum Dot tracking methods (Andrews et al. 2009). However, both of these experimental methods retain no information about the binding patterns of aggregates. This makes it difficult, if not impossible, to determine the difference between simply proximal and aggregated IgE.

**IgE Aggregation Models & Simulations**

To model ligand-receptor interactions a wide array of models have been implemented (Goldstein and Perelson 1984; Yang et al. 2008; Monine et al. 2010; Manavi, Wilson, and Tapia 2012). Most of these models were designed to study bivalent cell surface receptors interaction with trivalent ligands in a well mixed system. Two assumptions made are: 1) ligand binds to receptors, no ligand-ligand or receptor-receptor interactions, and 2) a single ligand cannot bind to both sites of a receptor. One of the initial models developed was the Goldstein-Perelson model (Goldstein and Perelson 1984). This model was based on thermodynamic equilibrium and only considers two interactions, those between free ligands and receptors and the crosslinking of two receptors by a ligand. Only considering these two interactions means structures such as cycles are not accounted for. The TLBR model is a kinetics based version of the Goldstein-Perelson model and includes a wider array of interactions (Yang et al. 2008). The TLBR model takes into account cyclic dimers and heximers which are generated using ideal molecule placement of bound molecules. A simulation version of the TLBR model was created that considered steric constraints of the molecules in the system (Monine et al. 2010). Dynamic bond trees are a graph formulation of aggregates (Chang and Yang 2011) which were intended for accelerated bookkeeping, not aggregate analysis.

**Ligand Structure and Valency**

A lot of work has been done studying synthetic ligands. Synthetic ligands such as the bivalent DCT, trivalent DF3, and multivalent BSA$_n$ have been constructed to study receptor aggregation (Posner et al. 1995; Sil et al. 2007; Xu et al. 1998). The structures of these ligands are well known and documented. All these structures use DNP, a hapten used in molecular biology for their high immunogenicity, to bind IgE antibodies. Each binding site is a DNP linker bound to the base ligand. DCT is a synthetic bivalent ligand (Figure 1 left) with 2 binding sites on opposite sides of the molecule. DF3 is a newer synthetic ligand (Figure 1 right) with 3 binding sites. BSA$_n$ is a synthetic multi-valent ($n$ being the valency) ligand used to analyze receptor aggregation. The number of DNP linkers bound to BSA$_n$ can vary (2-25 binding sites). However, there is no control of spatial distribution of the binding sites on BSA$_n$, so there is no guarantee of binding site uniformity.

![Figure 1: Synthetic ligands used in our simulations. Bivalent DCT (left) and trivalent DF3 (right).](image)

**Experimental Methods**

In this section, we address the methods used for model construction, simulation techniques, and analysis techniques.

**Model Construction**

Our simulation is performed using three dimensional models of the receptor complex molecules and ligand molecules. We use constructed all-atom models (Mahajan et al. 2013) as well as generate our own. The construction of these models use a combination of motif binding geometries, molecular dynamics, and homology modeling. In (Mahajan et al. 2013), available molecular PDB structures were used to construct the model of the receptor complex (PDBs: 1OAU, 2VWE, 100V, 1F6A). The IgE structure is modeled bound to FcεRI and consists of 1.532 amino acids (11,850 atoms).
We use two ligands in our experiments with varying structure. The bivalent ligand we use is DCT and the trivalent ligand is a fibrin trimer (PDB: 1RFO). To create DCT, DNP linkers were coupled to the alpha amino groups of a pair of bound L-tyrosine. To create DF3, the N-terminus of each fibrin of the trimer was extended with a DNP linker.

It would be computationally prohibitive to use all-atom models for the size of simulations we are pursuing, so we reduce model complexity. We compute the iso-surface of the molecules, constructing a high resolution structure of the molecule’s occupied volume. Iso-surface construction was done using Chimera, a molecular modeler. We end up with high-resolution object files which contain large amounts of detail. This detail hinders performance, such as collision detection, when simulating hundreds of molecules.

To make the models more computationally feasible, we further reduce the complexity by applying a polygonal reduction algorithm using the Maya modeling tool. This decreases the amount of detail in the model while still maintaining the occupied volume. These simplified structures are what we use as molecular models in our simulation. The process for model construction is shown in Figure 2.

Because we are modeling molecules as rigid bodies, a method had to be created to account for DNP linker flexibility. We model the DNP linkers with a structure half of the length of a linker. During the simulation we model a sphere with diameter the length of the DNP linker and it’s origin at the end of the partial-length linker. This sphere is treated as the binding area, any outside molecule with a binding site present binding sites within the radius of the modeled linker will be evaluated for potential binding. This is a simple probabilistic calculation based on experimental association rates. Dissociation rates are handled by evaluating every bond in the simulation at each time step for dissociation based on experimental dissociation rates. As bonds are formed and broken, the graph $G$ is updated with the addition and removal of edges respectively.

Over time, the ligands and receptors bind to form aggregates. An aggregate is simulated as a single body and moves slower depending on its size. Simulations are run for a predefined time, chosen to be well beyond the time stable graph formation is determined. Figure 3 shows a small scale example of our experiments.

### Aggregate Model and Analysis

We analyze our results using a variety of techniques based on state of the graph $G$. We define the system as a graph $G\{V,E\}$ where $V$ is the set of all molecules in the simulation, and $E$ is the set of edges in the graph and $e\{v_a, v_b\} \in E$ iff $v_a$ is bound to $v_b$. We note that since ligand only bind to receptors and vice versa, this is a bipartite graph.

This graph represents the structure of all aggregates in the system and can be analyzed using standard graph metric tools. For example, the number of edges in $G$ should stabilize when the simulation reaches a stable state. Another example is the number of connected components in $G$ measures the number of aggregates and singletons in the simulation. Each aggregate is represented by a connected component, $g$, in $G$. Each $g$ can be individually analyzed in order to quantify important characteristics.
Results

For each ligand, we simulate a variety of ratios of ligand to receptor. Receptor count is consistent at 90 and we varied the ligand count, setup to match experimental analysis (constant receptor concentration, varied ligand concentration). We ran counts of 30, 45, 90, and 180 for both ligands. We simulated the molecules on a 400nm by 400nm membrane patch. A fixed time interval of 10.0ms was used for all simulations. The association and dissociation rates used were 1.0 and 0.01, respectively, and were determined experimentally (Xu et al. 1998). The speed for all molecules is \( s = 0.09 \mu \text{m}^2/\text{s} \), based on the speed of IgE-Fc\( \epsilon \)RI. Recent experimental evidence suggests unbound molecules move faster than bound molecules (Andrews et al. 2009). To account for this, as molecules aggregate, the speed of the aggregate \( i \) is reduced to \( s/v_i \), \( v_i \) being the number of molecules in \( i \). Multiple (10) runs of each experiment were performed. Simulations were created using PMPL, a motion planning library developed at Texas A&M University. Experiments were run on single cores of a super computer at UNM with Intel Xeon E5645 processors and 4GB RAM per processor.

Equilibrium of Aggregate Formation

The metrics developed in (Manavi, Wilson, and Tapia 2012) can be applied at anytime during the experiment. However, aggregates are the most interesting and complex when they have had time to stabilize. To analyze the stability of the system, we look at the number of edges in the graph \( G \). As the number of edges in \( G \) stabilizes, aggregates structures don’t change significantly. The average number of edges in \( G \) over the course of the experiments are shown in Figure 4.

As can be seen in Figure 4, the number of edges initially grows quickly in all of the experiments. At around the 2 minute marker, the rate at which edges are being added to \( G \) starts to slow down and we see a leveling off of the edge count. We infer this as the aggregates becoming stable and fully formed. This result is found to be consistent with observations seen in IgE-Fc\( \epsilon \)RI, where changes in mobility are associated with aggregation. IgE-receptor aggregation slow down has been observed experimentally within 20 secs of ligand addition and are typically complete within 60-90 secs (Andrews et al. 2009). We also see that in Figure 4 DF3 produces more edges than DCT for the same ratio of ligand to receptor. This is expected due to the valency of the ligand, DF3 has 50% more binding sites accessible per molecule.

Aggregate Size

The TEM approach has shown that large Fc\( \epsilon \)RI “signaling patches” form within 1-2 minutes of the addition of ligand. One limitation of this technique is that it is not possible to estimate aggregate sizes within these patches. In this section, we show how using a graph based approach can be used to understand the differences in aggregate sizes.

Results are shown in Figure 5. Aggregate size was measured for every connected component in \( G \). Since experimental studies are only able to distinguish receptor position, we measure aggregate size as the cardinality of the vertices labeled receptors. After the simulations were run, aggregate size counts were collected and averaged for each ligand, aggregate sizes of two or larger were reported.

We see two trends in Figure 5, the first being the relation between ratios and the second between the ligands. If we look at different ratios for both ligands, the middle ratios (45 and 90 ligand counts) produce more aggregates of any given size relative to the extreme ratios (30 and 180 ligand counts). Not only that, but these median ratios also produce larger aggregates compared to the extreme ratios.
Figure 5: Aggregate sizes and number of occurrences. Aggregate sizes (number of receptors) were enumerated at the end of run and averaged. The x-axis is the size of aggregate and y-axis is the average number of aggregates of that size.

In Figure 5 there are clear aggregate size differences depending on the ratio of ligand to receptor, and the results are consistent across both ligands. Looking at low ligand ratio, ligands have a hard time binding already bound receptors since there are so few. Because of this aggregates tend to stay small. Looking at high ligand ratio, we see a saturation of ligand making for smaller aggregates. This is attributed to the receptor binding sites quickly filling up with unbound ligands. We see the more moderate ratios produce more aggregates and larger aggregates as well. Starting from a receptor saturated (low ligand ratio) and increasing the ligand count, aggregate count of a particular size increases. This trend continues until ligand saturation occurs and we see a decline in the number of aggregates of a particular size. This trend is seen in the bell shape curve in the number of occurrences of any given aggregate size and molecule counts.

We also see in Figure 5 that there are differences in aggregate size dependent on the ligand. We see that the trivalent DF3 produces larger aggregates than the bivalent DCT. We account this to the valency difference, trivalent DF3 has more accessible binding sites and can produce more complex structures (cycles, chains and trees) than bivalent DCT (cycles and chains).

**Resulting Aggregates**

Ligands with different valences can produce different aggregate formations. The resulting aggregates seen in Figure 6 were constructed during our simulation. We see that trivalent ligands are capable of generating aggregates that cannot be made using bivalent ligands i.e. trees.

Details of aggregate structures (Figure 6) are interesting because aggregate binding patterns are difficult to see using experimental imaging techniques. For example, notice the compactness of Figure 6 (A). Even though the DCT ligand is only able to produce simple structures, the receptor positions are compact, similar to the DF3 aggregate (Figure 6 (B)).

**Conclusion**

In this paper, we analyze aggregate formation from the cross-linking of an antibody to bivalent and trivalent ligands. By developing simplified models based on experimentally derived data, we are able to study aggregate formation under biologically-relevant conditions. This is particularly helpful since experimental techniques report proximity of clustered receptors but fail to provide a measurement of aggregate sizes within a complex topography. Our 3-D simulations have provided insight into the aggregation process by reporting both realistic timescales and by providing detailed information on receptor aggregates and ligand binding.

**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, the Adaptive Motion Planning Research Group member Andrea Howells for help with the modeling, Brid-
get S. Wilson for her advisement on the project, and the UNM Center for Advanced Research Computing.

References
Bayazit, O. B.; Song, G.; and Amato, N. M. 2001. Ligand binding with OBPRM and haptic user input: Enhancing automatic motion planning with virtual touch. 954–959.
Fewtrell, C., and Metzger, H. 1980. Larger oligomers of IgE are more effective than dimers in stimulating rat basophilic leukemia cells. 125:701–710.