Gaussian Mixture Models with Constrained Flexibility for Fitting Tomographic Tilt Series

Kasra Manavi
University of New Mexico
Albuquerque, New Mexico 87131
kazaz@cs.unm.edu

Sahba Tashakkori
University of New Mexico
Albuquerque, New Mexico 87131
stashakkori@unm.edu

Lydia Tapia
University of New Mexico
Albuquerque, New Mexico 87131
tapia@cs.unm.edu

ABSTRACT

Electron Microscopy (EM) has become an invaluable part of modern structural biology. As the limits of EM techniques are explored, issues arise with model reconstruction and fitting, especially at resolutions lower than 10 Å where structural detail is lost. In this paper, we introduce methods to model and fit molecular structures into low resolution EM datasets (20-40 Å resolution). First, we use Gaussian Mixture Models (GMMs) to describe molecular systems with high flexibility and enable efficient conformation sampling. Then, GMM parameters are optimized to best describe the model molecular structure. Finally, these GMMs are placed into a fitting procedure to generate a conformation of the GMM that fits an input set of tilt series tomograms, a set of 2D images of a 3D molecule taken at a variety of angles. We evaluate our method by fitting a model of the IgE-FcϵRI complex to a variety of simulated single tilt axis datasets. This work was done in preparation for analysis of larger aggregate structures of IgE.

KEYWORDS

Molecular Structural Biology; Flexible Fitting; EM Tomography

1 INTRODUCTION

Fitting a known structure to experimentally imaged molecules is critical to understanding molecular conformations. One such tool to image molecules, Electron Microscopy (EM), has become an essential part of understanding cellular function [13]. EM is performed by preparing a sample and placing it into an electron microscope. The microscope produces an image (tomogram) of the sample. A tomogram can be thought of as a 2D image of a 3D object where the intensity of the image pixel is proportional to the density of the 3D object. Electron Tomography (ET) is a technique that utilizes a series of EM tomograms to produce a 3D reconstruction of the molecular structure. Tomograms of a sample are captured from a variety of angles and are then combined to produce a 3D density map. These tomograms can be collected in an ordered fashion (e.g., single/dual axis tilt series) or an unordered fashion (e.g., single molecule imaged in solution). EM and ET are becoming cornerstones of modern structural biology research, but several challenges remain, including model fitting and reconstruction evaluation [9].

ACMBCB '17, August 2017, Boston, MA, USA

ACM ISBN 978-1-4503-4722-8/17/08…$15.00
https://doi.org/10.1145/3107411.3107504

Depending on the type of sample, density map reconstruction can be challenging. This step, required before an atomic model can be fit, can suffer from issues including distortions in tomogram alignment and the missing wedge problem [15]. Distortions in alignment occur when the tomograms of a dataset are clustered and aligned to produce an averaged tomogram from a given perspective. The quality of the clustering/alignment results are highly dependent on the size of the input dataset and tomogram quality. The missing wedge problem occurs during reconstruction of a tilt series, a set of projections captured by rotating the sample about an axis. If the angle range is narrow, portions of the sample are not imaged, resulting in a loss of information that elongate reconstructions.

In addition to the challenge of reconstruction, fitting structural models, typically all-atom structures, to reconstructed density maps can be difficult. Resolution of a density map is important to the type of fitting technique used. At high resolution (< 10 Å), all-atom fitting techniques work well, but lower resolutions (20-40 Å) still pose a challenge to existing methods due to a lack of detail [26]. This is particularly true for larger asymmetric molecular systems that are typically imaged at lower resolutions.

In order to address both the issues of reconstruction quality and fitting at lower resolutions, we propose a method that uses a flexible reduced-resolution molecular model for fitting directly to a series of tomograms. These tomograms depict images of a static molecular structure at a series of angles which can be efficiently fit with our GMM by comparing projections of our reduced-resolution molecular model at corresponding angles of the tomograms. In order to represent flexibility of the model that may exist in the imaged molecule, we use a multi-body GMM with constrained flexible joints. GMM fit evaluation is performed against all tomograms in the dataset, thus enabling evaluation of a conformation of the GMM against experimental data. While this method does not directly provide all-atom structures, it provides a quick assessment of the imaged molecule and the reasonableness of the conformation.

Our motivation for this approach is determination of structures of large molecular assemblies, specifically antibody/antigen aggregates critical to the understanding of the human allergic immune response. As a first step toward this goal, we focus on one component of these assemblies, the large antibody complex (IgE-FcϵRI). These molecules are potentially highly dynamic, so we are inspired by static images such as datasets captured via Cryo EM where samples are flash frozen before imaging. To evaluate our construction procedure, we generate a GMM of IgE-FcϵRI and assess how well the GMM can model a variety of datasets by optimizing parameters of the GMM. We then evaluate our fitting technique by generating a set of IgE-FcϵRI conformations under a variety of conditions and attempting to fit them to the known native conformation. We show that our GMM is flexible enough to model a variety of datasets and is capable of fitting tilt series data.

2 RELATED WORK

Projection Matching. Several computational geometry methods have been used to model volumes from projection information and vice versa [2, 17, 19, 27]. For example, the optimal packing/covering problems have been solved using Phi-functions which were developed to evaluate the interaction of geometric objects [2]. Another example is shadow art, the idea of occluding light from a source to produce an image. One such method uses light sources and a desired shadow art as input and produces a sculpture that is capable of generating the scene [17]. This work has been expanded to model shadow theater where shadow art is generated by the pose of a single or multiple performance artists [27]. Another similar method turns 2D silhouettes from different perspectives into 3D models [19].

EM and Structure Determination. Integrating comparative modeling and EM data to produce atomic models is reviewed in [25]. Fitting structures to reconstructed 3D EM data can be broken down into two main methodologies, rigid and flexible [7]. Most six degree of freedom rigid fitting is done using methods like geometric techniques [4], GMMs [12], or Zernike descriptors [6]. Flexible fitting focuses on using molecular simulation methods [14, 23], robotic motion planning techniques [1], as well as statistical techniques [5, 22] to determine candidate conformations. GMMs have been applied to other aspects of EM analysis, including reconstruction of single particles imaging [11] and structural dynamic evaluation [10].

Antibody Structure. Antibodies are immunological molecules which identify and target foreign threats to the body. Immunoglobulin (Ig) proteins have been determined to be highly flexible and can form asymmetrical structures [3, 21]. The structures obtained from X-ray diffraction analysis show that IgG is composed of three major structural subunits: two identical binding arms (Fab arms) and a membrane bound constant domain [20]. Since antibodies are known to have very flexible and dynamic structures, populations of different conformations have been found to co-exist in images [24, 28]. Therefore, commonly used methods such as X-ray crystallography, which rely on molecular averaging, often do not reflect protein dynamics and flexibility [29]. In contrast, EM can be used to reconstruct unique and independent samples [20].

3 METHODS

To generate a semi-flexible reduced resolution model of a molecular system, we start with an all-atom molecular structure. This structure is decomposed into rigid subunits, and Principal Component Analysis (PCA) is applied to these subunits to generate a GMM. Flexibility between rigid regions is added, and GMM subunits are refined to complete the model construction. This process is outlined in Figure 1. The resulting model is used to fit tomographic tilt series. In this paper, evaluation focuses on the IgE-FcεRI complex [16]. The model was constructed using available PDB structures (1OAU, 2VWE, 1O0V, 1F6A) and is composed of 1,709 amino acids (13,477 atoms total) [16].

3.1 Model Construction

To begin model construction, we identify rigid subunits of the model by performing rigidity analysis of the all-atom model using Kinari-Web [8]. If a sequence of residues is considered flexible, we identify the rigid subunits at the termini of the flexible sequence.

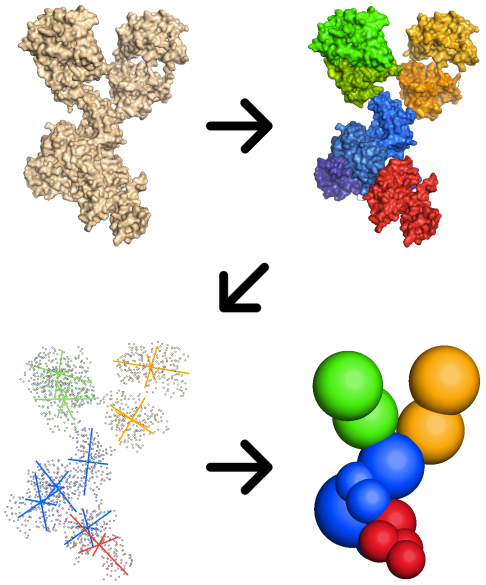


Figure 1: Process of taking an all-atom structure and generating a GMM. First, the atomic model (*top left*) is decomposed into rigid subunits (*top right*). The rigid subunits are then processed using PCA (*bottom left*). The results of PCA are then used to construct a GMM (*bottom right*).

If both ends of the sequence are associated with the same subunit, the sequence is considered a rigid part of that subunit. For our IgE model, ten rigid subunits were classified into five regions outlined in Table 1.

Structural Subunit	Chain & Residues	Region
FcεRIα 1	A 5-84	1
FcεRIα 2	A 88-169	1
Constant 1	H 256-355, I 256-355	1
Constant 2a	H 364-462	1
Constant 2b	I 364-462	1
Constant 3	H 469-571, I 469-571	1
Fab 1 - C terminal	H 151-247, L 140-234	2
Fab 1 - N terminal	H 21-144, L 26-127	3
Fab 2 - C terminal	I 151-247, M 140-234	4
Fab 2 - N terminal	I 21-144, M 26-127	5

Table 1: IgE-FcεRI subunits calculated from rigidity analysis. A flexible region is considered part of a rigid region if both ends of the region are associated with the same structural subunit.

After rigid subunits have been identified, we convert the decomposed subunits into a GMM representation. This is done by performing PCA on the atom positions of each subunit structure identified by rigidity analysis decomposition (Figure 1, *bottom left*).

If the difference between the lowest and highest eigenvalues of the PCA is high, the subunit is divided. From our data we used a ratio of 1.8x or larger as a cutoff for division. For the IgE complex, subunits 1 and 2 of FcϵRIα and constant subunits 2a and 2b were split along their principle component.

Using the results of PCA, molecular subunits are converted to a GMM description with Gaussian functions centered at the subunit's mean μ with standard deviation σ set to the square root of the largest subunit eigenvalue. For each subunit, we include a σ-scaling factor to compensate for potential differences in model resolution, discussed in Section 3.2. This construction method is applied to all subunits of the IgE complex, resulting in a model composed of fourteen Gaussian functions (Figure 1, *bottom right*).

To capture flexibility that may exist between rigid subunits, we incorporate distance constraints into the GMM. Specifically, distances between GMM subunits that are linked together via non-rigid portions of the protein backbone are maintained. If a flexible region is found to connect rigid subunits during rigidity analysis, the α-carbons of the amino acids on both termini of the flexible regions are stored with their associated GMM subunit. These points are evaluated for distance when a new GMM conformation is generated. The distance must be less than $n{\cdot}d$, where n is the number of amino acids in the flexible region and d is the average distance between α-carbons. For our IgE-FcϵRI complex model, we use the average distance between the α-carbons of the flexible regions.

In the IgE complex, there are six flexible regions that connect molecular subunits. Each Fab arm is connected to the constant domain, resulting in two distance constraints. Fab arms are composed of two subunits connected together by two flexible regions, resulting in two distance constraints per arm. This results in six distance constraints total, outlined in Table 2.

Class	Subunits	Chain & Residues	Length
Arm Linker	Con 1, Fab 1C	H 248-255	8
Arm Linker	Con 1, Fab 2C	I 248-255	8
Fab Short	Fab 1C, Fab 1N	H 145-150	6
Fab Long	Fab 1C, Fab 1N	L 128-139	12
Fab Short	Fab 2C, Fab 2N	I 145-150	6
Fab Long	Fab 2C, Fab 2N	M 128-139	12

Table 2: Flexible regions between rigid antibody subunits. Con 1 represents the Constant subunit and Fab subunits are labeled "Fab *XY*", *X* being the arm index (1 or 2) and *Y* being the termini (N or C). Length is in number of residues in the flexible region.

3.2 GMM Refinement and Fitting

In this section we discuss how we refine GMM parameters to tailor our GMMs to a particular experiment (Algorithm 1). We then outline how we minimize GMM conformation to identify a best fit to the provided EM tilt series (Algorithm 2).

When tomograms and density maps are generated, various levels of image and post processing are applied to prepare the data for reconstruction/fitting. Depending on the parameterization of these processes, results can differ from experiment to experiment. This can alter the ability of the GMM to represent the data. In order to capture these differences, we introduce Algorithm 1 which performs parameter refinement, a method that produces a GMM that is tailored to a particular dataset that can also be used to fit other datasets from the same experiment.

Algorithm 1 GMM Parameter Refinement

Input. GMM model c_{init}, dataset projections P, parameter ranges R and maximum iteration count $iter_{max}$
Output. conformation c_{opt} optimized to match P

1: $c_{cur} \leftarrow c_{init}$
2: $o_{cur} \leftarrow EvalOverlap(c_{cur}, P)$
3: **for** $iter = 0; iter \leq iter_{max}; iter{+}{+}$ **do**
4: $\quad v_{cur} \leftarrow EvalParams(c_{cur}, R, o_{cur})$
5: $\quad c_{cur} \leftarrow UpdateParams(c_{cur}, v_{cur})$
6: $\quad o_{cur} \leftarrow EvalOverlap(c_{cur}, P)$
7: **end for**
8: $c_{opt} \leftarrow c_{cur}$

To generate a GMM specific to a given dataset, each parameter of the GMM needs to be optimized (positions and σ-scales) with respect to the input dataset. Parameter refinement begins with evaluating the initial GMM quality (line 2). To determine the quality of the match between a conformation and a dataset, we compare projections of the GMM and dataset from the same perspective. Projections are generated and compared as seen in Figure 2. Corresponding pixel values of the projections are evaluated using the OR and XOR operators. The OR operator provides us with the number of pixels in the projections occupied by both the GMM (u) and the part of the dataset visualized as the non-black pixels in the last row of Figure 2. The XOR operator provides us with the number of pixels that are occupied by either GMM or dataset but not both (x), representing the differences in projections (red and blue pixels in the last row of Figure 2). The overlap score is the proportion of pixels that overlap in the union, $((u - x) / u)$.

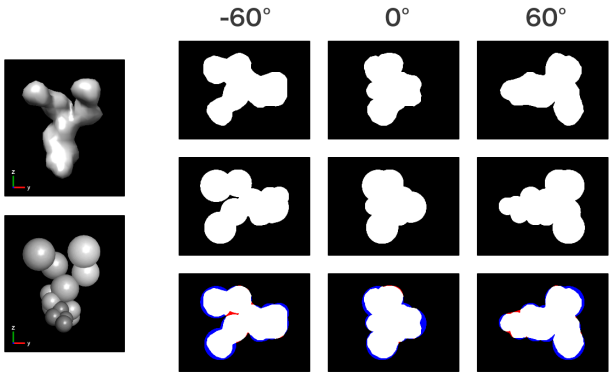


Figure 2: Comparison of tilt series to evaluate overlap score. Molecular data (*top left*) and GMM (*bottom left*) orthogonally projected at angles {-60,0,60} (*top/middle rows, right*). Projections from the same perspective are overlaid and evaluated for OR and XOR (*bottom row, right*).

To perform GMM parameter refinement, we use a binary search scheme over a range of values (R) centered at the initial parameter values of the input GMM (C_{init}). For each iteration, every parameter is evaluated for overlap score at 3 values, at the upper and lower bounds of the range and the middle (line 4). The overlap scores are compared and the parameter values with the highest overlap scores are used to update the saved values and the ranges for the next iteration (line 5). The process is iterated for a specified number of times ($iter_{max}$), and the resulting GMM parameter values are returned.

Algorithm 2 GMM conformation Fitting

Input. GMM model conformation c_{init} and tilt series projections t, maximum iteration count $iter_{max}$, window averaging width w and local minima detection interval i

Output. conformation c_{min} that is the best fit of t

1: $c_{cur} \leftarrow c_{init}$
2: $c_{best} \leftarrow c_{cur}$
3: $so_{cur} \leftarrow EvalSubOverlap(c_{cur})$
4: $so_{best} \leftarrow so_{cur}$
5: **for** $iter = 0; iter \leq iter_{max}; iter{+}{+}$ **do**
6: $c_{new} \leftarrow UpdateGMM(c_{cur}, t)$
7: $IsCollisionFree \leftarrow CollisionResolution(C_{new})$
8: $IsDCSatisfied \leftarrow DCSatisfaction(c_{new})$
9: **if** $InLocalMin(w, i)$ **then**
10: $c_{cur} \leftarrow Resample(c_{new})$
11: **else**
12: **if** $!IsCollisionFree \;||\; !IsDCSatisfied$ **then**
13: $c_{cur} \leftarrow Resample(c_{new})$
14: **else**
15: $c_{cur} = c_{new}$
16: **end if**
17: **end if**
18: $so_{cur} = EvalSubOverlap(c_{cur})$
19: **if** $EvalNewMax(so_{cur}, so_{best}) == True$ **then**
20: $c_{best} = c_{cur}$
21: $so_{best} \leftarrow so_{cur}$
22: **end if**
23: **end for**

At this point, we have a GMM for fitting a tilt series. Our procedure for conformational fitting is described in Algorithm 2. To begin fitting, we start by generating a GMM conformation to fit. Preparation of the initial GMM conformation for minimization can be performed using a variety of methods including rigid body fitting, user specified conformations, or using a sampling-based methods. For the IgE-FcϵRI-complex, we evaluate the procedure with an analysis of minimizing a wide variety of initial GMM conformations. We evaluate and score the fit of a given conformation directly to the tilt series by generating projections of our GMM conformation and individual subunits and compare the projections to their tilt series counterpart (line 3).

The overlap scores of the individual subunits are stored to a list used to monitor minimization progress. For each iteration, we evaluate changes in overlap score for each subunit. At a given positional resolution, all subunits are translated each axis in both directions (6 evaluations total) and overlap scores are calculated. If necessary, i.e., for evaluating a multi-component subunit, orientation is evaluated in a similar fashion to positions using rotation. The results of position and orientation evaluation are used to generate a new conformation c_{new} (line 6). The new conformation is resolved of any collisions and evaluated for distance constraint satisfaction.

To avoid getting stuck in local minima, we monitor the progress of the score throughout the minimization process. We do this by calculating the current average overlap score by averaging the last w overlap scores of the minimization. We evaluate progress by adding the current average overlap score to the list of previously averaged overlap scores and calculate the differences between consecutive average scores (deltas). If the most recent delta values of the local minima interval, i, are less than or equal to zero, the conformation is considered to be in a local minima (line 9) and is resampled. For resampling, individual subunit overlap scores are evaluated and the worst performing subunit is resampled. If the new conformation is not considered in a local minima, the new sample c_{new} is validated (line 12). If c_{new} is collision free and satisfies all distance constraints, it is assigned to c_{cur}. If not, the conformation is resampled.

Once a new conformation is generated, c_{cur} is updated and compared to the current best conformation and is stored if the overlap score is better (lines 18-21). The best scoring conformation is returned.

4 EXPERIMENTS

4.1 GMM Parameter Refinement

To evaluate whether or not our GMM model is robust enough to fit datasets of varying quality, we refine the parameters of our GMM to a variety of tilt series collected from simulated density maps of the native conformation of IgE-FcϵRI. Density maps were generated at three resolutions (20, 30, and 40 Å) using the molecular modeling software Chimera [18]. Density maps are rendered at a specified isolevel, producing an image of the surface of a density map that has been thresholded at the isolevel value. Higher isolevel values render surfaces that represent higher density volumes in the map. For lower isolevel values, the surface is typically larger since the lower threshold results in an increase of volume to render. As mentioned previously, values can vary between density maps so isolevel values for rendering are evaluated empirically on a per dataset basis. For our evaluation, the maps were rendered at two isolevel values (0.09 and 0.06), resulting in six datasets.

To generate projections for GMM parameter refinement, we rotate our sample molecule along the X and Y axis (dual-axis). Rotations about the X axis were generated over 180° in 10° intervals (18 angles total). Rotations about the Y axis were generated over 360° in 10° intervals (36 angles total). All combinations of X and Y angle values are evaluated, resulting in a total of 648 projections.

For our experiments, two rounds of ten iterations refinement were used (Algorithm 1). The first round started with an initial conformation, and the second started with the results of the first round (Figure 3). We observed that two rounds of refinement provided an exceptional increase in model quality, but more rounds did not significantly improve the result.

From the results, GMM is able to model a variety of density maps and isosurface values well. Overlap scores of the initial GMM range from ~70-85% to the different datasets. After refinement, our six different datasets all produced GMM parameters with overlap scores between 91% and 95%, Figure 3. This increase in overlap

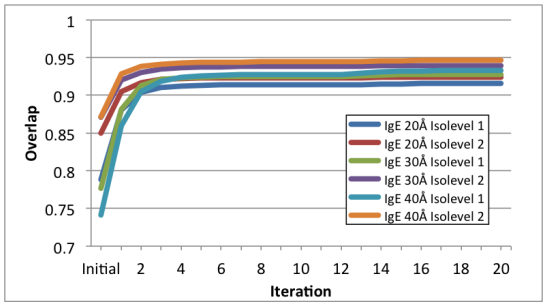


Figure 3: Refinement of the IgE-FcϵRI complex GMM parameters. Score (Y-axis) is shown over multiple iterations (X-axis) for two rounds of optimization.

shows how robust our GMM can be in dealing with various EM datasets and that we can produce GMMs tailored to fit particular datasets.

4.2 GMM Tilt Series Fitting

In this section we present results on tilt series fitting using our GMM and density map projections. We use single axis projection sets where rotation is performed about the Y-axis. We evaluate datasets generated from three angle ranges [-45°, 45°], [-60°, 60°] and [-90°, 90°]. These represent 90°, 120° and 180° of total range, respectively. We used the 30 Å density map of our default antibody conformation rendered at an isosurface value of 0.09 to generate our projections. Collision was detected using a σ-scale value of 1.5 and a distance of 3.5 Å was used as D, the distance-per-residue for distance constraints. Positional and orientation resolutions values were 1.0 Å and 1.0°, respectively. Fifty runs of each experiment were performed.

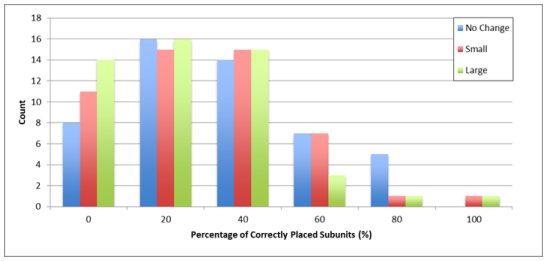
To evaluate our method, we fit GMM conformations to tilt series of the native state density map. We setup three conditions to model potential variation in the initial base position. Each condition allows a different amount of angle change in the base position. In the first condition, (No Change), we leave the IgE-FcϵRI complex base alone and just sample the Fab arms. This is considered the condition with the highest quality start states. In the second condition (Small), we randomly sample the IgE-FcϵRI complex base along the Y and Z axes within an angle range of [-30°, 30°] and then sample the Fab arms. The final condition (Large) is the same as the second but the angle range is [-60°, 60°] and models the condition with the lowest quality start states.

To determine the quality of the fitting, we used a scoring scheme to quantify the fit given subunit positions. The Root Mean Squared Distance (RMSD) of each GMM subunit is calculated, and a cutoff value of 8.0 Å (determined from subunit RMSD distributions) is used to determine if a given subunit is properly placed. Average RMSD values are reported in Table 3 and the number of properly located subunits is shown in Figure 4.

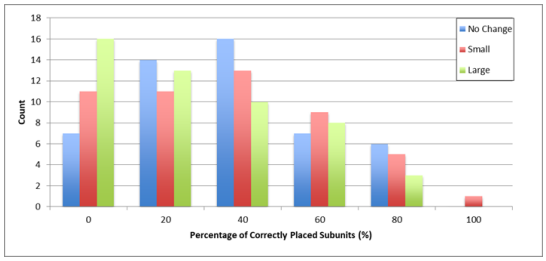
Table 3 demonstrates that as total angle range increases, RMSD decreases. Greater visibility, as provided by a larger number and range of tomograms, improves fitting quality. We see this trend in all conditions of the starting conformation. We also observe that starting condition does have an impact on fit quality. Sampling the

Tilt Series	No Change	Small	Large
90°	51.45 ± 22.26	57.56 ± 27.24	63.36 ± 27.41
120°	45.86 ± 19.69	54.83 ± 26.59	56.60 ± 23.81
180°	23.53 ± 12.27	26.58 ± 12.78	38.17 ± 22.36

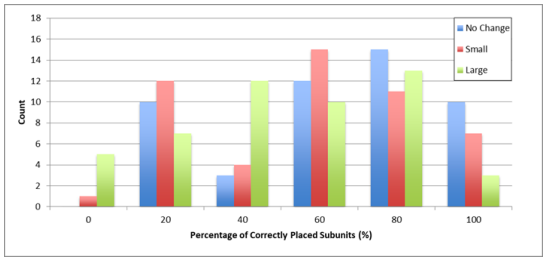
Table 3: Average RMSD (Å) with variance of the fitted GMM conformations to the native GMM conformation. Rows are angle range of the tilt series and the average RMSD values for different starting conditions.



(a) 90° Angle Range



(b) 120° Angle Range



(c) 180° Angle Range

Figure 4: Evaluation of correctly placed subunits for three experiments with different angle ranges. The number of generated conformations (Y-axis) is plotted versus percentage of correctly placed subunits (X-axis). No Change, Small, and Large represent the amount of variation in the position of the base.

base with no angle change performs the best and as base sampling range increases the fit is worse.

For the smallest total angle range of 90° (Fig 4a), we see that only a small number of conformations (~17%) produced place more than half of their subunits in the correct location. Figure 4b demonstrates that increasing the angle range to 120° increases overall

conformation quality. This is observed in the increase of conformations with correctly placed subunits, where ~26% of results placed more than half of the subunits correctly. This trend is continued in Figure 4c where the largest angle range of 180° produces the largest percentage of correctly positioned conformations (~64%). Of these correct conformations, about 13% have all subunits in the structure correctly positioned.

5 DISCUSSION AND CONCLUSION

In this paper, we presented a method for flexible fitting of GMM models to projections from tilt series. We performed this by generating a GMM representation of atomic models of the IgE-FcϵRI complex. We showed that our GMM can represent a range of tomogram resolutions. We also showed that the our GMM parameters can be refined to be compatible with a wide array of datasets. We then used our GMM to fit tilt series projections directly and evaluated our method's ability to perform flexible fitting. From our fitting results, we are able to see that the minimization scheme we implemented works better with more information and performs better when the initial conformation is closer to the conformation described by the tilt series. For future work, our goal is to work with our experimental collaborators who produce cryo EM tilt series and begin fitting their datasets.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) under Grant Numbers IIS-1528047, IIS-1553266 (Tapia, CAREER), and CCF-1518861 and by National Institutes of Health Grant P50GM085273 to the New Mexico SpatioTemporal Modeling Center. The authors would like to thank the University of New Mexico Center for Advanced Research Computing for the computational resources they provided, as well as Bridget Wilson and Niels Volkmann for their discussions and input.

REFERENCES

[1] Kamal Al Nasr and Jing He. 2009. An Effective Convergence Independent Loop Closure Method Using Forward Backward Cyclic Coordinate Descent. *International Journal of Data Mining and Bioinformatics* 3, 3 (2009), 346–361.
[2] J. Bennell, G. Scheithauer, Y. Stoyan, and T. Romanova. 2010. Tools of mathematical modeling of arbitrary object packing problems. *Annals of Operations Research* 179, 1 (2010), 343–368.
[3] L Bongini, D Fanelli, F Piazza, P De Los Rios, S Sandin, and U Skoglund. 2005. Dynamics of antibodies from cryo-electron tomography. *Biophysical Chemistry* 115, 2 (2005), 235–240.
[4] Pablo Chacón and Willy Wriggers. 2002. Multi-resolution contour-based fitting of macromolecular structures. *Journal of Molecular Biology* 317, 3 (2002), 375–384.
[5] Pilar Cossio and Gerhard Hummer. 2013. Bayesian analysis of individual electron microscopy images: Towards structures of dynamic and heterogeneous biomolecular assemblies. *Journal of Structural Biology* 184, 3 (2013), 427 – 437.
[6] Juan Esquivel-Rodríguez and Daisuke Kihara. 2012. Fitting Multimeric Protein Complexes into Electron Microscopy Maps Using 3D Zernike Descriptors. *The Journal of Physical Chemistry B* 116, 23 (2012), 6854–6861.
[7] Juan Esquivel-Rodríguez and Daisuke Kihara. 2013. Computational methods for constructing protein structure models from 3D electron microscopy maps. *Journal of Structural Biology* 184, 1 (2013), 93–102.
[8] Naomi Fox, Filip Jagodzinski, Yang Li, and Ileana Streinu. 2011. KINARI-Web: a server for protein rigidity analysis. *Nucleic acids research* 39, suppl 2 (2011), W177–W183.
[9] Richard Henderson, Andrej Sali, Matthew L Baker, Bridget Carragher, Batsal Devkota, Kenneth H Downing, Edward H Egelman, Zukang Feng, Joachim Frank, Nikolaus Grigorieff, and others. 2012. Outcome of the first electron microscopy validation task force meeting. *Structure* 20, 2 (2012), 205–214.
[10] Slavica Jonić and Carlos Óscar Sánchez Sorzano. 2016. Coarse-graining of volumes for modeling of structure and dynamics in electron microscopy: Algorithm to automatically control accuracy of approximation. *IEEE Journal of Selected Topics in Signal Processing* 10, 1 (2016), 161–173.
[11] Paul Joubert and Michael Habeck. 2015. Bayesian inference of initial models in cryo-electron microscopy using pseudo-atoms. *Biophysical Journal* 108, 5 (2015), 1165–1175.
[12] Takeshi Kawabata. 2008. Multiple Subunit Fitting into a Low-Resolution Density Map of a Macromolecular Complex Using a Gaussian Mixture Model. *Biophysical Journal* 95, 10 (2008), 4643 – 4658.
[13] Catherine L. Lawson, Matthew L. Baker, Christoph Best, Chunxiao Bi, Matthew Dougherty, Powei Feng, Glen van Ginkel, Batsal Devkota, Ingvar Lagerstedt, Steven J. Ludtke, Richard H. Newman, Tom J. Oldfield, Ian Rees, Gaurav Sahni, Raul Sala, Sameer Velankar, Joe Warren, John D. Westbrook, Kim Henrick, Gerard J. Kleywegt, Helen M. Berman, and Wah Chiu. 2011. EMDataBank.org: unified data resource for CryoEM. *Nucleic Acids Research* 39 (2011), D456 – D464.
[14] Steffen Lindert, Nathan Alexander, Nils Wötzel, Mert Karakaş, Phoebe L. Stewart, and Jens Meiler. 2012. EM-Fold: De Novo Atomic-Detail Protein Structure Determination from Medium-Resolution Density Maps. *Structure* 20, 3 (2012), 464 – 478.
[15] Vladan Lučić, Friedrich Förster, and Wolfgang Baumeister. 2005. Structural studies by electron tomography: from cells to molecules. *Annu. Rev. Biochem.* 74 (2005), 833–865.
[16] 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. 2014. Optimal Aggregation of FcϵRI with a Structurally Defined Trivalent Ligand Overrides Negative Regulation Driven by Phosphatases. *ACS Chemical Biology* 9, 7 (2014), 1508–1519.
[17] Niloy J. Mitra and Mark Pauly. 2009. Shadow Art. *ACM Trans. Graph.* 28, 5 (Dec. 2009), 156:1–156:7.
[18] Eric F. Pettersen, Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. 2004. UCSF Chimera–A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 13 (2004), 1605–1612.
[19] Alec Rivers, Frédo Durand, and Takeo Igarashi. 2010. 3D Modeling with Silhouettes. *ACM Trans. Graph.* 29, 4 (2010), 109:1–109:8.
[20] Sara Sandin, Lars-Göran Öfverstedt, Ann-Charlotte Wikström, Örjan Wrange, and Ulf Skoglund. 2004. Structure and flexibility of individual immunoglobulin G molecules in solution. *Structure* 12, 3 (2004), 409–415.
[21] Erica Ollmann Saphire, Robyn L Stanfield, MD Max Crispin, Paul WHI Parren, Pauline M Rudd, Raymond A Dwek, Dennis R Burton, and Ian A Wilson. 2002. Contrasting IgG structures reveal extreme asymmetry and flexibility. *Journal of Molecular Biology* 319, 1 (2002), 9–18.
[22] Sjors H.W. Scheres. 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* 180, 3 (2012), 519 – 530.
[23] Gunnar F. Schröder, Axel T. Brunger, and Michael Levitt. 2007. Combining Efficient Conformational Sampling with a Deformable Elastic Network Model Facilitates Structure Refinement at Low Resolution. *Structure* 15, 12 (2007), 1630 – 1641.
[24] Huimin Tong, Lei Zhang, Allan Kaspar, Matthew J Rames, Liqing Huang, Gary Woodnutt, and Gang Ren. 2013. Peptide-conjugation induced conformational changes in human IgG1 observed by optimized negative-staining and individual-particle electron tomography. *Scientific Reports* 3 (2013), 1089:1–1089:9.
[25] Maya Topf and Andrej Sali. 2005. Combining electron microscopy and comparative protein structure modeling. *Current Opinion in Structural Biology* 15, 5 (2005), 578 – 585.
[26] Elizabeth Villa and Keren Lasker. 2014. Finding the right fit: chiseling structures out of cryo-electron microscopy maps. *Current Opinion in Structural Biology* 25 (2014), 118–125.
[27] Jungdam Won and Jehee Lee. 2016. Shadow Theatre: Discovering Human Motion from a Sequence of Silhouettes. *ACM Trans. Graph.* 35, 4 (July 2016), 147:1–147:12.
[28] Lei Zhang and Gang Ren. 2012. IPET and FETR: Experimental Approach for Studying Molecular Structure Dynamics by Cryo-Electron Tomography of a Single-Molecule Structure. *PLOS ONE* 7, 1 (2012), 1–19.
[29] Xing Zhang, Lei Zhang, Huimin Tong, Bo Peng, Matthew J Rames, Shengli Zhang, and Gang Ren. 2015. 3D structural fluctuation of IgG1 antibody revealed by individual particle electron tomography. *Scientific Reports* 5 (2015), 9803:1–9803:13.