Biophysical Journal
Article

Biophysical Society

Structural Insights into Hearing Loss Genetics from Polarizable Protein Repacking

Mallory R. Tollefson,[1,3] Jacob M. Litman,[2] Guowei Qi,[2] Claire E. O'Connell,[1] Matthew J. Wipfler,[1] Robert J. Marini,[3] Hernan V. Bernabe,[1,3] William T. A. Tollefson,[1] Terry A. Braun,[1] Thomas L. Casavant,[1] Richard J. H. Smith,[3,*] and Michael J. Schnieders[1,2,*]

[1]Department of Biomedical Engineering and [2]Department of Biochemistry, University of Iowa, Iowa City, Iowa; and [3]Molecular Otolaryngology & Renal Research Laboratories, Department of Otolaryngology-Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa

ABSTRACT Hearing loss is associated with ~8100 mutations in 152 genes, and within the coding regions of these genes are over 60,000 missense variants. The majority of these variants are classified as "variants of uncertain significance" to reflect our inability to ascribe a phenotypic effect to the observed amino acid change. A promising source of pathogenicity information is biophysical simulation, although input protein structures often contain defects because of limitations in experimental data and/or only distant homology to a template. Here, we combine the polarizable atomic multipole optimized energetics for biomolecular applications force field, many-body optimization theory, and graphical processing unit acceleration to repack all deafness-associated proteins and thereby improve average structure MolProbity score from 2.2 to 1.0. We then used these optimized wild-type models to create over 60,000 structures for missense variants in the Deafness Variation Database, which are being incorporated into the Deafness Variation Database to inform deafness pathogenicity prediction. Finally, this work demonstrates that advanced polarizable atomic multipole force fields are efficient enough to repack the entire human proteome.

SIGNIFICANCE We are interrogating the genetics of deafness using a targeted sequencing panel (called OtoSCOPE) that includes 152 deafness-associated genes. OtoSCOPE enables us to identify an average of 545 variants per patient, which are curated in the deafness-specific database we purpose built called the Deafness Variation Database (http://deafnessvariationdatabase.org). To inform the interpretation of missense variants from a structural biology perspective, we describe new, to our knowledge, algorithms for repacking protein structures. Our approach, implemented in the publicly available software Force Field X (https://ffx.biochem.uiowa.edu), is used to generate 473 wild-type structures for OtoSCOPE genes. These protein models have been integrated into the Deafness Variation Database to inform classification of missense variants and form the foundation for downstream analyses of protein-protein binding and folding stability.

INTRODUCTION

As the most common human sensory deficit, deafness impacts an estimated 360 million people globally (World Health Organization data, http://www.who.int/pbd/deafness/estimates/en/index.html). Its cause is multifactorial, and with recent advances in the application of targeted genetic sequencing technology to clinical medicine, our understanding of genetic contributions to deafness has greatly advanced. The use of deafness-specific gene panels has changed the clinical paradigm in the evaluation of the deaf patient and is laying the foundation for personalized gene therapy to treat hearing loss.

The targeted genetic sequencing panel developed by our group, which we refer to as OtoSCOPE, includes 152 deafness-associated genes (1,2). Its use enables us to identify an average of 545 variants per patient, which are curated in the publicly available deafness-specific database we purpose built called the Deafness Variation Database (DVD, Fig. 1; Table S1; http://deafnessvariationdatabase.org)

Submitted February 20, 2019, and accepted for publication June 25, 2019.

*Correspondence: richard-smith@uiowa.edu or michael-schnieders@uiowa.edu

Mallory R. Tollefson and Jacob M. Litman contributed equally to this work.

Editor: Alan Grossfield.

https://doi.org/10.1016/j.bpj.2019.06.030

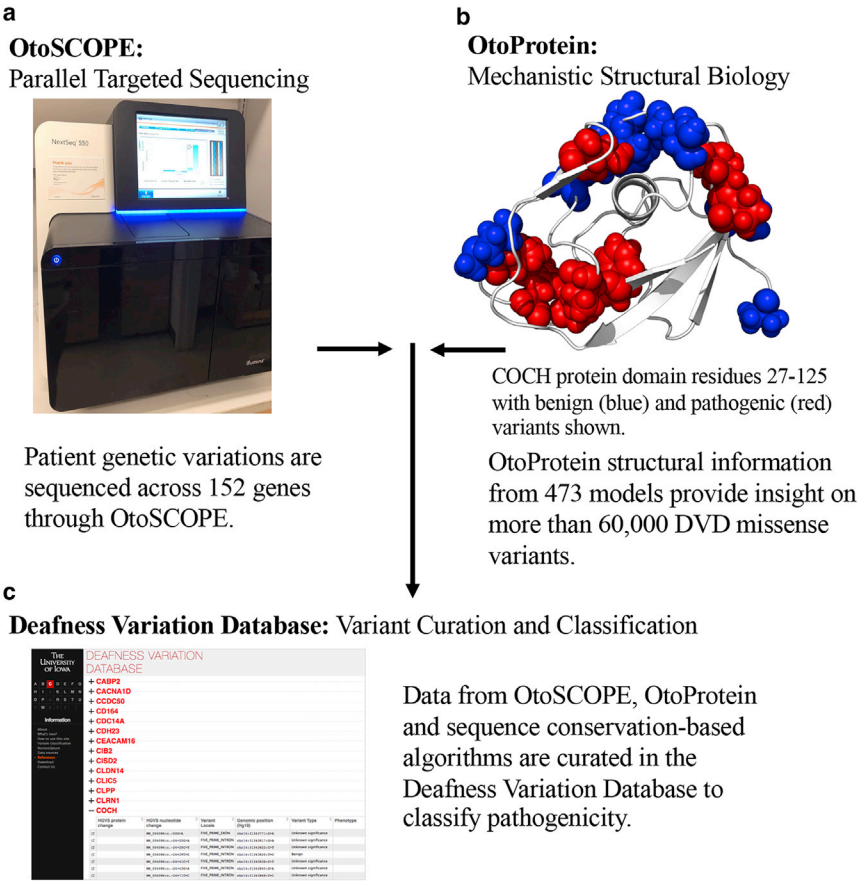

FIGURE 1 Incorporating structural biophysics into variant classification. (*a*) OtoSCOPE sequencing technology discovers 545 variants per patient on average, 71 of which are nonsynonymous coding, splice site, or indel variants. (*b*) Protein structural coverage for OtoScope genes is an important step toward identifying the molecular causation of disease-causing variants along with classifying VUSs. (*c*) Variants collected through OtoSCOPE sequencing are curated in the Deafness Variation Database (DVD). For each variant, the DVD combines minor allele frequency, experimental results, pathogenicity predictions from sequence conservation-based classifiers, and now, insights from protein structures (i.e., *OtoProtein*, described in this work). The pathogenicity for nearly 80% of the variants in the DVD remains unknown, which means they are placed into the variant of uncertain significance (VUS) category. To see this figure in color, go online.

(3,4). The DVD collates data from major public databases and uses criteria recommended by the American College of Medical Genetics and Genomics to classify every genetic variant as benign (B), likely benign (LB), variant of uncertain significance (VUS), likely pathogenic (LP), or pathogenic (P) based on collected evidence and curation by experts in genetic hearing loss. Of the ~800,000 variants in the genes included on OtoSCOPE that are listed in the DVD, more than 60,000 missense variations exist. Of these variants, ~4000 are LP/pathogenic, ~38,000 are VUSs, and ~18,000 are LB/B.

Many of the missense variations labeled as VUSs will ultimately be classified as LP/pathogenic, but we are currently relegated to classifying them as VUSs as a reflection of our inability to predict the phenotypic consequences of most genetic variations. We often lack variant-specific wet-lab-based functional evidence (5), and insights derived from atomic resolution simulations must continue to mature to reliably make meaningful genotype-phenotype correlations.

Atomic resolution simulation techniques such as molecular dynamics (MD) provide a promising first-principles approach for computationally predicting the potential impact of missense variants. However, its success is dependent, in part, on accurate protein structures. These structures are typically determined from an experimental method (i.e., x-ray crystallography, NMR, cryo-electron microscopy, etc.) or from homology modeling. The latter leverages existing protein structure(s) as a template from which to create the model of a homologous amino acid sequence. Homology modeling is most reliable when homologous sequences have at least 30% sequence identity, which typically indicates protein fold is conserved (6,7). To complement and enhance models available in databases such as ModBase (8) and SwissProt (9), dramatic improvements are possible by global optimization (i.e., repacking) of amino acid side chains using more advanced molecular physics than was originally available (or could be computationally afforded) at the time of their creation.

For example, most protein structures found in both the Protein Data Bank (10) and homology modeling databases (8,9) are based on refinement with pairwise potential energy functions (i.e., force fields) such as the fixed-charge Amber (11,12), CHARMM (13,14), and OPLS-AA (15,16) models (17). Over the past decade, more accurate polarizable force fields have emerged that overcome limitations in previous generation pairwise models (18), including both the Atomic Multipole Optimized Energetics for Biomolecular Applications (AMOEBA) force field (19,20) and the CHARMM

Drude (21) model. Structural optimization with these state-of-the-art energy functions, when used with continuum representations of solvation (22–24), can compensate for limitations in experimental data and improve homology models. However, multiple challenges must be overcome to realize the benefits of polarizable atomic multipole force fields, including mitigating their increased computational expense and overcoming the loss of convenient pairwise approximations that are widespread in structural biology software such as Modeler (25), Phenix (26), and Rosetta (27).

Our decision to use the AMOEBA force field for this work is based on a series of previous structural refinement studies performed over the last decade (28–35) that systematically demonstrate both improved agreement to x-ray and neutron diffraction data (i.e., lower R/R_{free}) and improved MolProbity metrics (36,37) compared to fixed-charge force fields. MolProbity identifies high-energy atomic clashes, unfavorable side-chain conformations, and polypeptide backbone conformations inconsistent with low-energy secondary structure. The algorithm is widely used by crystallographers to aid refinement of models by reporting structural features that are known to be unphysical. Lower MolProbity scores are consistent with higher-quality x-ray diffraction data (i.e., a score of 1.0 is calibrated to reflect 1.0 Å resolution data). For example, AMOEBA-assisted x-ray refinement on ultra-high-resolution (0.43–0.59 Å) peptide crystals (28) and high-resolution (0.65–0.89 Å) lysozyme, trypsin, and DNA data sets (29) demonstrated lower R/R_{free} values compared to conventional refinement using force fields without advanced electrostatics. Work on joint x-ray/neutron data sets for B-form DNA and Z-form DNA (31) demonstrated that AMOEBA-assisted refinement outperformed a variant of OPLS-AA in terms of both lower R/R_{free} values and improved water hydrogen bonding networks. Finally, more recent work on a series of proliferating cell nuclear antigen structures thoroughly compared AMOEBA-assisted x-ray refinement to OPLS-AA/L, including rotamer repacking in both cases, and AMOEBA again outperformed OPLS-AA/L in terms of both lower R/R_{free} and improved MolProbity scores (35).

Building on the promising refinement results from these previous studies, here we use refinement with AMOEBA to generate a family of deafness-related protein structures called *OtoProtein*. Our approach combines the AMOEBA potential energy function (19,20), many-body optimization theory (35), and GPU acceleration (38,39) to optimize all available deafness-associated protein models. To assess the resulting structures objectively, we evaluated overall quality with the MolProbity (36,37) algorithm. Correcting rotamer outliers often improves other metrics and permits further relaxation of the structure with local minimization, resulting in more realistic, lower-energy structures for downstream analysis (e.g., MD, alchemical free-energy simulations, or feature extraction for bioinformatics analysis).

As described in the Results, our mean postoptimization MolProbity score is consistent with near-atomic resolution. The structures have been integrated with the DVD to provide insight into the biophysical impacts of deafness-related genetic variations, which aids in predicting variant effect and pathogenicity. Our polarizable protein repacking algorithm is freely available in the open source software Force Field X (FFX, http://ffx.biochem.uiowa.edu) and may be useful to others in the community that are integrating structural biophysics into variant classification.

MATERIALS AND METHODS

Many-body energy expansion parallelization across GPUs

Generation and assessment of *OtoProtein* structures, as depicted in Fig. 1, will now be described in detail. Under a many-body potential, the total energy of a protein $E(\mathbf{r})$ can be defined to arbitrary precision using the expansion

$$E(\mathrm{r}) = E_{env} + \sum_i E_{self}(r_i) + \sum_i \sum_{j>i} E_2(r_i, r_j) + \sum_i \sum_{j>i} \sum_{k>j} E_3(r_i, r_j, r_k) + \ldots, \quad (1)$$

where E_{env} is the energy of the environment (i.e., the protein backbone and residues that are not being optimized); $E_{self}(r_i)$ is the self-energy of residue i that includes its intramolecular bonded energy terms and nonbonded interactions with the backbone; $E_2(r_i, r_j)$ is the two-body nonbonded interaction energy between residues i and j with other residues turned off; and $E_3(r_i, r_j, r_k)$ is the three-body nonbonded interaction energy between residues i, j, and k with other residues turned off. The self, two-body, and three-body energy terms are calculated as follows, where $E_{BB/SC}$ is the total energy of the backbone with the side chain(s) of the selected residue(s) included (shown graphically in Fig. 2 *a*).

$$E_{self}(r_i) = E_{BB/SC}(r_i) - E_{env} \quad (2)$$

$$E_2(r_i, r_j) = E_{BB/SC}(r_i, r_j) - E_{self}(r_i) - E_{self}(r_j) - E_{env} \quad (3)$$

$$E_3(r_i, r_j, r_k) = E_{BB/SC}(r_i, r_j, r_k) - E_{self}(r_i) - E_{self}(r_j) - E_{self}(r_k) - E_2(r_i, r_j) - E_2(r_i, r_k) - E_2(r_j, r_k) - E_{env} \quad (4)$$

Individual energy evaluations are calculated on graphical processing units (GPUs) via the CUDA kernels of OpenMM (39), and the evaluations are distributed over many GPUs, potentially over multiple nodes, using the PJ library (Fig. 2 *b*; (40)). Side-chain rotamer conformations that are not part of the optimum structure can be rigorously eliminated using mathematical expressions (Fig. 2 *c*; (35,41,42)).

Computing the self, two-body, and three-body energy terms as a function of rotamer conformation is computationally expensive. To address this challenge, our FFX program utilizes two complementary parallelization approaches, including 1) use of the Parallel Java (PJ) (40) message-passing interface library to distribute terms among multiple processes and 2) use of the OpenMM application programming interface (39) to perform force-field-energy evaluations on NVIDIA GPUs (Nvidia, Santa

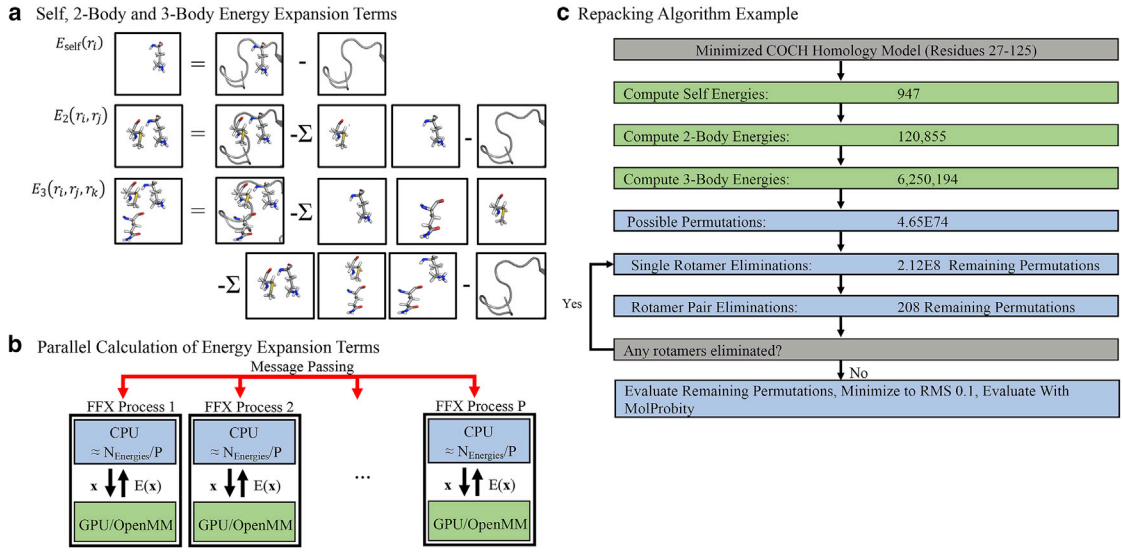

FIGURE 2 Overview of the protein repacking algorithm. (*a*) Depictions of rotamer self, two-body, and three-body energy terms are given. (*b*) Parallel computation of energy terms across processes and GPUs is shown. Processes (*blue boxes*) are each assigned a group of self-energies to calculate, where $N_{self} = \sum_{i=1}^{N_{residues}} n_i$ is the sum of rotamers across all residues to give N_{self}/P evaluations per process. Processes compute energy values by sending a conformation $\mathbf{x}$ to a GPU (*green box*) for evaluation using the OpenMM application programming interface, followed by return of the energy $E(\mathbf{x})$ and its communication to all processes using the Java message-passing interface (*red arrows*). The two-body and three-body energies are parallelized in a similar fashion. (*c*) The number of side-chain energies and conformational permutations for a 98-residue COCH protein domain are shown as an example. After all energy terms have been calculated (*green rectangles*), the combinatorial side-chain conformational space is reduced using many-body Goldstein rotamer and rotamer pair elimination criteria (see Materials and Methods) to achieve a tractable number of permutations to evaluate. Before eliminations, 4.65×10^{74} side-chain permutations exist, but only 208 permutations remain to be evaluated after eliminations. To see this figure in color, go online.

Clara, CA) via CUDA kernels. FFX uses PJ to divide each shared memory node of a multiple node compute cluster into one or more processes (Fig. 2 *b*). Energy terms are then assigned to processes, evaluated, and globally communicated across all processes using PJ message passing, with synchronization steps between calculation of the self, two-body, and three-body energy terms (i.e., two-body terms depend on self-terms as shown in Eq. 3, and thus must be calculated after self-energies are completed and before three-body energies). The FFX-OpenMM interface (based on Java Native Access wrappers to the OpenMM C++ application programming interface) is used to offload energy evaluations from FFX, which executes on central processing units (CPUs), to OpenMM on a GPU. Once all energy terms are calculated, side-chain rotamers and rotamer pairs are eliminated by lower-energy alternatives based on rigorous mathematical inequalities that have been described for pairwise force fields (e.g., dead-end elimination (41) and Goldstein elimination (42)) and more recently generalized to include three-body terms for use with many-body force fields (35) such as the polarizable AMOEBA model (20,30). The many-body Goldstein criteria for rotamer elimination (35), truncated at three-body terms, is given by

$$E_{self}\left(r_i^{\alpha}\right) - E_{self}\left(r_i^{\beta}\right) + \sum_{j'} \min_{\gamma} \left\{ E_2\left(r_i^{\alpha}, r_j^{\gamma}\right) - E_2\left(r_i^{\beta}, r_j^{\gamma}\right) + \sum_{k'} \min_{\delta} \left[E_3\left(r_i^{\alpha}, r_j^{\gamma}, r_k^{\delta}\right) - E_3\left(r_i^{\beta}, r_j^{\gamma}, r_k^{\delta}\right) \ldots \right] \right\} > 0 \tag{5}$$

and, if satisfied, indicates rotamer α of residue i is eliminated by rotamer β (the ellipses signify the presence of further higher-order terms). The expression for rotamer pair elimination is given by

$$E_{pair}\left(r_i^{\alpha}, r_j^{\beta}\right) - E_{pair}\left(r_i^{\gamma}, r_j^{\delta}\right) + \sum_{k'} \min_{\varepsilon} \left\{ E_2\left(r_i^{\alpha}, r_k^{\varepsilon}\right) + E_2\left(r_j^{\beta}, r_k^{\varepsilon}\right) + E_3\left(r_i^{\alpha}, r_j^{\beta}, r_k^{\varepsilon}\right) - E_2\left(r_i^{\gamma}, r_k^{\varepsilon}\right) - E_2\left(r_j^{\delta}, r_k^{\varepsilon}\right) - E_3\left(r_i^{\gamma}, r_j^{\delta}, r_k^{\varepsilon}\right) + \sum_{l'} \min_{\eta} \left[E_3\left(r_i^{\alpha}, r_k^{\varepsilon}, r_l^{\eta}\right) + E_3\left(r_j^{\beta}, r_k^{\varepsilon}, r_l^{\eta}\right) - E_3\left(r_i^{\gamma}, r_k^{\varepsilon}, r_l^{\eta}\right) - E_3\left(r_i^{\delta}, r_k^{\varepsilon}, r_l^{\eta}\right) + \ldots \right] \right\} > 0 \tag{6}$$

and, if satisfied, indicates that the rotamer pair $(r_i^{\alpha}, r_j^{\beta})$ for residues i and j is eliminated by rotamer pair $(r_i^{\gamma}, r_j^{\delta})$.

Four approximations to rigorous use of the many-body Goldstein inequalities given above were explored, each of which is summarized here and described more fully in the Results. First, it was determined that the expansion could be truncated at pairwise terms because of damping of three-body and higher-order terms by the generalized Kirkwood implicit solvent. However, in the absence of implicit solvent, previous work demonstrated that inclusion of three-body terms is sometimes necessary (35). The second approximation was a distance cutoff; if the closest rotamers for a residue pair or triple are more than 2 Å apart, the interaction energy is set to 0. Third, pruning was utilized to remove rotamers with self-energies 25 kcal/mol or more above the lowest self-energy of a residue before calculation of two-body energies. This criterion is based on the heuristic observation that rotamers with such an unfavorable self-energy (e.g., due to an atomic clash with backbone atoms) are not found in well-packed structures. However, for structures with significant backbone flaws, this approximation

must be used with care because it can incorrectly eliminate the "least bad" rotamer that is actually part of the global minimal conformation. Our final approximation involved imposing a three-dimensional grid over the protein, followed by optimization within each subdomain (cube) of the grid, rather than including all protein residues simultaneously. Although the repacking algorithm is a provable global optimizer within a single subdomain of the grid, it is not for the protein grid as a whole because coordinated changes between subdomains are neglected.

The *OtoProtein* Structure Database

Comparative protein modeling provides a means to predict the structure of a protein whose atomic coordinates have not been solved experimentally by crystallography, NMR, etc. (43). Many human genes implicated in hearing loss have not been studied experimentally, so computational approaches are necessary to generate high-quality protein structures. Comparative protein modeling begins from an experimental structure for an evolutionarily related protein, which is used as a template for the target sequence (10,44). The percent sequence identity between the homologs provides an estimate of model reliability (45). Comparative protein models are conducive to the study of protein function, dynamics, and interactions with other molecules such as ligands, DNA, RNA, or other proteins. Homology models can also be used to study missense variants, providing a promising basis for understanding the role of protein phenotypes in heterogenic diseases like hearing loss.

However, comparative protein models from leading databases often include defects directly related to approximations in the methods used for their generation (e.g., pairwise force fields, local rather than global optimization, etc.). We sought to improve comparative protein models from SwissProt (44) and ModBase (8) for 152 genes included in the OtoSCOPE platform. Although using homology models based on a sequence identity of 30% or greater generally gives confidence that the protein backbone fold has been evolutionarily conserved (45), this work includes all publicly available models (the average sequence identity was 41.7% for all 473 structural models). Both SwissProt and ModBase strive to provide structural coverage for the largest portion of the human proteome possible; however, this limits their ability to explore the use of advanced many-body force fields. Here, we show that use of the polarizable AMOEBA force field in tandem with global optimization of amino acid side chains (35) can significantly improve the quality of SwissProt or ModBase structures as assessed by tools like MolProbity (36,37). High-quality protein structural models, in turn, provide optimal starting points for downstream MD simulations that can be used to analyze missense variations (i.e., calculation of folding and/or binding free-energy differences). The parallelized repacking algorithm described here demonstrates that it is now feasible to refine large databases of homology models using advanced polarizable atomic multipole force fields.

All homology models were refined using the 2018 AMOEBA protein force field (20,46) with generalized Kirkwood implicit solvent (23). The input homology models were first locally optimized using the low-memory Broyden-Fletcher-Goldfarb-Shanno algorithm to a root mean square (RMS) gradient convergence criterion of 0.8 kcal/mol/Å. The rationale for minimizing with a relatively loose convergence criterion before rotamer optimization was to relax the backbone conformation without excessively favoring the starting conformation over alternative rotamers. Locally optimizing to a tighter convergence criterion before side-chain optimization resulted in higher-energy, less favorable structures because of overstabilizing starting rotamers. Next, the side-chain repacking algorithm was applied, followed by a final local low-memory Broyden-Fletcher-Goldfarb-Shanno minimization to an RMS gradient convergence criterion of 0.1 kcal/mol/Å. The resulting protein structures and original homology models were then evaluated and compared using both the MolProbity assessment tool and AMOEBA/GK energies. These optimized wild-type structures were used as input for creating more than 60,000 variant structures for missense variations in the DVD. For each missense variant in the DVD with structural coverage available in the wild-type *OtoProtein* data set, the corresponding structure(s) were mutated. We then locally repacked all residues within 2 Å of the missense variant (i.e., based on Eq. 7 below) to correct any atomic clashes introduced by the variant amino acid. Many proteins in our data set have more than one homology model available, often covering similar residue ranges. For each missense variant, a locally repacked structure was created for every available wild-type model, which yielded multiple structures for many missense variants in our data set.

RESULTS

Polarizable protein repacking algorithm using GPUs

To benefit fully from the emergence of polarizable force fields in the context of protein structure prediction and repacking, the theory that underlies established algorithms must be revisited to incorporate many-body electronic polarization and to optimize performance across GPUs. We examined four approximations to our many-body protein repacking algorithm to enhance efficiency while maintaining structural quality. The approximations are illustrated using a 98-residue COCH protein (residues 27–125), which was chosen based on its high sequence identity to an experimental NMR template (98% identity) and modest size. Previous work showed that truncating the energy expansion at three-body terms resulted in accurate side-chain positions being identified in the context of real-space x-ray refinement (35). However, when using the native environment approximation (47) in combination with the AMOEBA/GK implicit solvent (23), we found that the contribution of energy terms within the energy expansion decays quickly. The magnitude of each term in Eq. 1 dampens significantly enough that truncation at two-body terms is sufficiently accurate for repacking in implicit solvent (Table 1). This damping manifests as smaller two-body and three-body contributions when GK is enabled (Table S2). In fact, the magnitude of three-body interactions is reduced to such an extent that they generally do not affect side-chain rotamer eliminations (whereas our prior crystal refinement work did not employ an implicit solvation model). Truncation at two-body energy terms results in a nearly 52× speed-up (Table S3) as compared to the original rotamer optimization protocol (35) without any rotamer changes compared to including three-body terms. In future work, we plan to additionally optimize the protonation states of all titratable residues, which will necessitate a fresh appraisal of the impact of three-body energies because of the formal charge of residues changing.

The second approximation applies a distance-based cutoff between residues, which results in the interaction energy of two or more side chains being set to 0 if the minimal atomic distance between rotamer permutations is above a defined cutoff. The minimal distance $d_{\min}$ between two residues i and j is calculated using the expression

$$d_{min}(i,j) = \min_{\{\alpha=1..n_i,\beta=1..n_j\}} \left[dist\left(r_i^{\alpha}, r_j^{\beta}\right)\right], \qquad (7)$$

where the min operation is over the set of all rotamer permutations (i.e., residues i and j have n_i and n_j rotamers,

TABLE 1 Adjustable Repacking Parameters Are Examined in the Context of Computational Expense and Structural Quality

Truncation of the Energy Expansion at Either Two-Body or Three-Body Terms				
Expansion Truncation	**Relative Energy**	**Time**	**Speed-Up**	–
Two-body*	0.0	847	51.7×	–
Three-body	0.0	43,850	1.0×	–
Based on Truncation at Two-Body terms, Residue Distance Cutoffs from 1 to 6 Å Are Evaluated				
Residue Cutoff (Å)	**Relative Energy**	**Time**	**Speed-op**	**Overall**
1	32.6	137	5.9×	320.0×
2*	0.2	304	2.7×	144.2×
3	0.0	420	1.9×	104.4×
6	0.0	813	1.0×	53.9×
Based on Truncation at Two-Body Energy Terms and a 2 Å Residue Cutoff, Pruning Thresholds Are Evaluated				
Pruning Threshold	**Relative Energy**	**Time**	**Speed-Up**	**Overall**
5	0.2	43	7.1×	1019.8×
15	0.2	113	2.7×	388.1×
25*	0.2	142	2.1×	308.8×
No pruning	0.0	304	1.0×	144.2×
Based on Truncation at Two-Body Energy Terms, a 2 Å Residue Cutoff, and 25 kcal/mol Pruning Threshold, Cube Edge Lengths from 10 to 30 Å Are Evaluated for Cube Optimization				
Cube Size (Å)	**Relative Energy**	**Time**	**Speed-Up**	**Overall**
10*	0.2	40	3.6×	1096.3×
20	0.9	94	1.5×	466.5×
30	0.2	142	1.0×	308.8×

All tests used residues 27–125 of isoform 1 of the COCH protein. All relative potential energies (in kcal/mol) are compared to the global rotamer minimum of the COCH protein as calculated when using no approximations. Times are wall clock times in seconds using a node with four GPUs. The individual and overall speed-ups for each approximation are given. Criteria for evaluating residue distance cutoffs are described in the main text.
*The recommended choice for use with AMOEBA and the GK implicit solvent for each adjustable parameter.

respectively, to give $n_i \times n_j$ permutations), and the distance function (*dist*) returns the minimal pairwise atomic distance given rotamer conformations r_i^{α} and r_j^{β}. We tested a range of cutoffs and found that 2 Å, when combined with truncation at two-body energies, provides a 144.2× speed-up compared to the original protocol while only increasing the energy relative to the global minimum by 0.2 kcal/mol (i.e., two solvent-exposed side chains had different conformations) (Table 1). Although 2 Å appears to be an overly aggressive cutoff at first glance, evidence from our data set of protein models (Table S4) shows structures still closely approach the global rotamer minimum. We emphasize that AMOEBA/GK force-field energetics are still evaluated with a typical 12 Å pairwise atomic cutoff, whereas the 2 Å residue cutoff is applied only in the context of using Eq. 7 to define interacting residues of the many-body energy expansion (see Figs. S1 and S2 for examples of applying Eq. 7).

The third approximation prunes rotamers and/or rotamer pairs if their conformation is higher than the lowest energy alternative plus a threshold

$$E_{self}\left(r_i^{\alpha}\right) > E_{self}\left(r_i^{\beta}\right) + threshold. \quad (8)$$

A pruning threshold of 25 kcal/mol results in further speed-up without compromising the quality of output structures (Table 1). Although pruning inequalities are not rigorous, unlike the mathematically proven Goldstein eliminations (see Materials and Methods), they obviate calculating many pair energies to yield over a 3× speed-up. Pruning did not result in any additional changes to rotamer conformations as compared to the global minimum found when using a two-body expansion and cutoff of 2 Å.

The final approximation uses a series of cube-shaped domains defined by imposing a three-dimensional grid over the protein, followed by sequential optimization of each cube of the grid. This approximation is especially useful for large protein domains that have an intractable number of energetically closely spaced permutations even after application of elimination criteria. By varying cube size and cube overlap, we determined that a cube edge length of 10 Å with no overlap optimized performance without degrading quality (Table 1). Cube optimization results in no additional change in energy relative to the global minimum found when using a two-body expansion and residue cutoff of 2 Å (note that a 30 Å cube contains the whole COCH domain and is a global optimization). Combining all four optimal approximations results in a total speed-up of ∼3 orders of magnitude.

We next implemented a parallelization approach that combines PJ with GPU acceleration. As the number of nodes is increased, our PJ message-passing parallelization algorithm achieved a near-linear speed-up (Tables 2 and S5). Offloading energy evaluations to OpenMM on a single node equipped with a GPU (two Intel Xeon E5-2680v4 CPUs

TABLE 2 Energy Evaluation Timings for Global Side-Chain Optimization of ACTG1 Residues 6–375 and COCH Residues 27–125 Using a Varying Number of GPUs

		Time for Energies (sec)		Speed-Up (Relative to Using All CPU Cores)	
Number of Nodes	Number of GPUs	ACTG1	COCH	ACTG1	COCH
1	0 (CPUs only)	33,126	5505	1.0×	1.0×
1	1	2576	479	12.9×	11.5×
1	2	1277	251	25.9×	21.9×
1	4	656	142	50.5×	38.8×
2	8	336	76	98.6×	72.4×
4	16	175	43	189.3×	128.0×

Each node contains two Intel Xeon E5-2680v4 CPUs and four NVIDIA GTX 1080 TI GPUs.

(Intel, Santa Clara, CA) and one NVIDIA GTX 1080 TI GPU [Nvidia]) resulted in a 11.5-fold speed-up compared to using the same node with no GPU (i.e., a single GPU was 11.5× faster than parallelization over all 28 Intel CPU cores) on the COCH protein domain. Testing parallelization on a larger protein domain such as ACTG1 residues 6–375 demonstrated greater speed-ups relative to the smaller COCH domain (e.g., a 189.3× speed-up for ACTG1 using 16 GPUs, compared to 128.0× for COCH). Our original CPU parallelized Java implementation of the algorithm with no approximations (COCH protein domain run on two Intel Xeon E5-2680v4 CPUs (Intel), three-body expansion, 6 Å cutoff, and no pruning) required calculation of over 6 million energy terms and consumed 16.5 compute days on a node. By combining algorithm approximations, parallelization across four processes on one node (i.e., PJ message passing) and GPU acceleration (1 GPU per process, four GPUs total), our algorithm executes the 20,232 AMOEBA/GK energy terms in only 142 s.

Comparison of Amber99sb/GB and AMOEBA/GK protein repacking

In ongoing work, we are performing AMOEBA free-energy simulations of DVD variants to provide insights into their pathogenic mechanisms. To illustrate the impact of optimizing structures using AMOEBA/GK compared to Amber99sb/generalized born (GB) before MD simulations, we optimized a model of CDH23 containing residues 887–1408 in both the Amber99sb/GB and AMOEBA/GK force fields. We chose to analyze CDH23 887–1408 because of its prevalence in hearing loss; CDH23 has more than 2000 missense variants in the DVD. Optimizing the CDH23 model with Amber99sb/GB resulted in 31 side chains changing from their original conformation, whereas optimizing with AMOEBA/GK resulted in 54 rotamer changes (Fig. 3). The extensive discrepancies between CDH23 conformations that result from repacking under AMOEBA/GK and Amber99sb/GB are consistent with our extensive prior structural refinement comparisons (28–35). Thus, because our focus is on using AMOEBA for downstream free-energy simulations, further Amber99sb/GB repacking was not justified.

The *OtoProtein* Structure Database

We applied our accelerated repacking algorithm to a set of 473 deafness-associated protein models. For both starting homology models and refined structures, quality was assessed using the heuristic MolProbity algorithm, which examines steric clashes, poor side-chain rotamers, and amino acid backbone favorability (e.g., ϕ/ψ dihedral angle combinations). The MolProbity score is calibrated to predict the quality of x-ray diffraction data that is expected to have produced the assessed structure (i.e., a MolProbity score of 1.5 corresponds to an expected x-ray resolution of 1.5 Å, where lower values indicate higher quality). On average, we reduced steric clashes per 1000 atoms from 25.1 to 0.03, decreased Ramachandran outliers from 2.03 to 0.94%, and decreased the percentage of poor side-chain rotamers from 2.3 to 1.6% (Table 3). Overall, the repacking protocol improved the mean MolProbity score from 2.16 to 1.04, demonstrating that our structures are consistent with protein structural models near-atomic resolution (Fig. 4). The average AMOEBA force-field energy for the data set when locally optimized to RMS gradient convergence criteria of 0.8 kcal/mol/Å was −15,342 kcal/mol. After global side-chain optimization, the average AMOEBA energy for the data set was reduced to −16,287 kcal/mol, a reduction of 945 kcal/mol from the structures that were minimized to an RMS gradient criterion of 0.8 kcal/mol/Å without rotamer optimization. Although local minimization without rotamer optimization dramatically reduces atomic clashes, the number of poor rotamers increased from 2.3 to 2.9% and motivates the need for side-chain repacking. The overall repacking procedure required just 71 GPU days for all 473 *OtoProtein* structures. The complete list of statistics for each model is available in Table S6. Based on these results, GPU-accelerated repacking with the polarizable AMOEBA force field could potentially be used to improve the quality of large protein structure databases with only a modest investment in hardware.

To assess the impact of our optimization algorithm as a function of protein features, we compared both final MolProbity score and MolProbity improvement (i.e., the change in MolProbity score due to refinement) with sequence identity to the homology template using linear

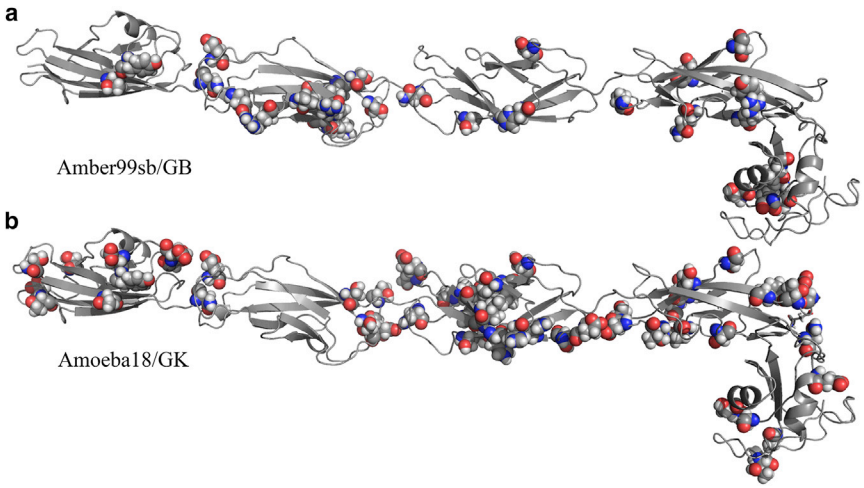

FIGURE 3 CDH23 residues 887–1408 optimized using (*a*) Amber99sb/GB and (*b*) AMOEBA/GK. Side chains with changes in conformation relative to the original homology model are shown as space-filling spheres. When optimized with the fixed-charge Amber99sb/GB force field, 31 side-chain rotamers changed conformation, whereas repacking with AMOEBA/GK resulted in 54 side-chain rotamers changing. To see this figure in color, go online.

regression (Figs. S3 and S4). With R^2 values of only 0.0035 and 0.0256, respectively, neither the MolProbity score nor MolProbity improvement were strongly correlated with sequence identity to the homology model. Similarly, neither MolProbity score nor MolProbity improvement were strongly correlated with the number of residues in the protein structure based on R^2 values of only 0.0119 and 0.0449, respectively (Figs. S5 and S6). We conclude that the favorable improvements afforded by the repacking approach described here are largely independent of both sequence identity to the homology model and protein size.

We demonstrate the impact of repacking protein models on variant interpretation through modeling of a pathogenic variant that causes Usher syndrome. Buried in a domain of the CDH23 structure shown in Fig. 3 is valine residue 1090, which causes Usher syndrome when mutated to isoleucine (Fig. 5 *a*). In the unrefined model of CDH23, a neighboring valine (position 1039) is oriented away from residue 1090 (Fig. 5 *b*). Optimization using the repacking approach presented here results in the β sheets surrounding Val1090 lengthening and the Val1039 rotamer closely approaching Val1090 (Fig. 5 *c*). When isoleucine is introduced at position 1090 in the unoptimized structure, Ile1090 and Val1039 do not clash, and accommodation of the variant appears possible. However, when Ile1090 is introduced in the AMOEBA/GK repacked structure, Val1039 and Ile1090 clash, indicating that the variant (known to be pathogenic) is consistent with destabilization of the CDH23 fold. Despite the qualitative nature of this analysis, it illustrates that without repacking, downstream variant free-energy simulations based on homology models will generally need to reach longer timescales to accommodate structural relaxations.

The *OtoProtein* Structure Database has been incorporated into the DVD to provide public availability of the models in combination with the exhaustive DVD genetic information. The combination of *OtoProtein* structural information with existing DVD data (e.g., minor allele frequency, pathogenicity assessment, etc.) provides a powerful platform for the auditory research community. For example, it is now possible to visualize clustering of pathogenic variations in specific domains of a protein and to examine structural features that correlate with pathogenicity (Fig. 6). Additionally, more than 60,000 variant-specific structures are available publicly (https://github.com/mrkeeney/deafness-variant-structures) and are currently being incorporated into the DVD.

DISCUSSION

Structural coverage of the human proteome has increased rapidly since the early 1990s, with ~40% of the human proteome now having comparative models based on templates with a sequence identity of at least 30% (9). Here, we applied a GPU-accelerated polarizable protein repacking algorithm to the deafness-associated proteome defined by homology models of any sequence identity (average sequence identity of the data set is 41.7%). We found that 38.8% of the deafness-associated proteome could be

TABLE 3 Average Refinement Statistics for the *OtoProtein* Structure Database

Database	Clash Score	Poor Rotamers	Ramachandran Favored	Ramachandran Outliers	MolProbity Score
Homology	25.09	2.33%	91.95%	2.03	2.16
Minimization	2.75	2.92%	91.85%	1.87	1.66
OtoProtein	0.03	1.60%	93.48%	0.94	1.04

The data set contains 473 structures.

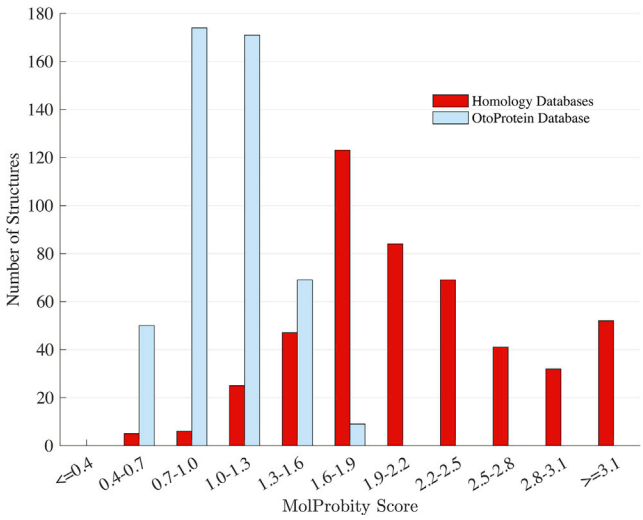

FIGURE 4 Histogram of MolProbity scores for the *OtoProtein* Structure Database before and after optimization. Before optimization (*red*), the 473 structures have an average MolProbity score of 2.16, whereas after optimization (*blue*), the data set has an average MolProbity score of 1.04 (i.e., approaching the quality expected of atomic resolution x-ray structures). To see this figure in color, go online.

modeled structurally, comparable to structural coverage of the entire human proteome. The 473 structural models we collected and optimized span 145 deafness-associated genes. These structures had an initial average MolProbity score of 2.16, but after repacking, the average score improved to approximately atomic resolution at 1.04. These calculations required just 71 GPU days. In addition to covering nearly 40% of OtoSCOPE with atomic resolution structural models, our *OtoProtein* database provides structural coverage for 22,809 of the 61,971 missense variations in the DVD (16,203 are VUSs, 1931 are LP/pathogenic, and 4675 are B/LB). These models are publicly available in the DVD through the NGL protein viewer (48). When integrated with information on patient missense variations available through the DVD, the *OtoProtein* database represents a unique tool for understanding deafness genetics from a structural perspective. Building on the *OtoProtein* structural platform, we have created more than 60,000 models of DVD missense variations, which are publicly available. Future work will simulate these missense variations to quantify thermodynamic free-energy differences and thereby provide insight into how they disrupt protein folding and/or alter protein-protein interactions.

The GPU-accelerated protein repacking algorithm is freely available to the research community through the FFX program, which may be useful to refine other structural data sets outside of the deafness domain. The algorithm is designed for use with advanced polarizable force fields and features an energy expansion up to three-body interactions. Computational speed is achieved using an architecture based on parallelization across an arbitrary number of compute nodes and GPUs and, together with algorithm optimizations, provides multiple orders of magnitude speed-up without compromising structural quality. Although polarizable repacking algorithms were previously not efficient enough to apply to large-scale data sets, this work opens the door to their application to all protein structures in the human proteome. For example, the Swiss Model Repository (SMR) lists 45,083 homology models with an average residue length of 232 amino acids (9). Structures of this size (i.e., ~230 residues) require only ~260 s to repack using our algorithm on a node with four GPUs (e.g., repacking our DSSP 88–318 model of 230 residues took 262 s of wall clock time). Based on the average model size in the SMR, we estimate that repacking all SMR human proteins

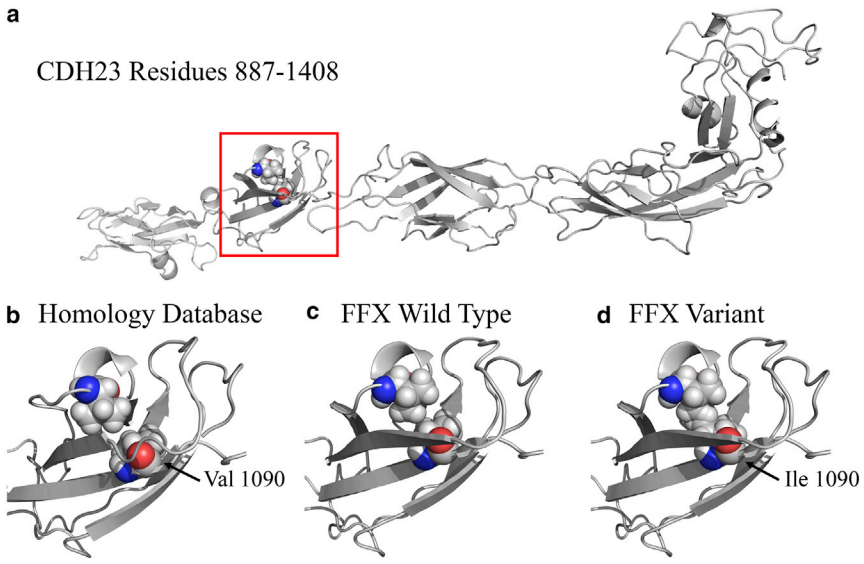

FIGURE 5 CDH23 structure showing the pathogenic DVD variant Val 1090 Ile. (*a*) The wild-type CDH23 model for residues 887–1408 is shown, with Val 1090 highlighted by a red box. (*b*) Shown in spacefill format are Val 1039 and Val 1090. (*c*) After AMOEBA/GK repacking, the conformation of Val 1039 changes relative to the variant. (*d*) A clash is present between Val 1039 and the Ile 1090 variant, which is consistent with altered folding stability and classification as disease causing. This example illustrates that without repacking, downstream free-energy simulation timescales must (in general) increase to allow relaxation of nonoptimal side-chain conformations. To see this figure in color, go online.

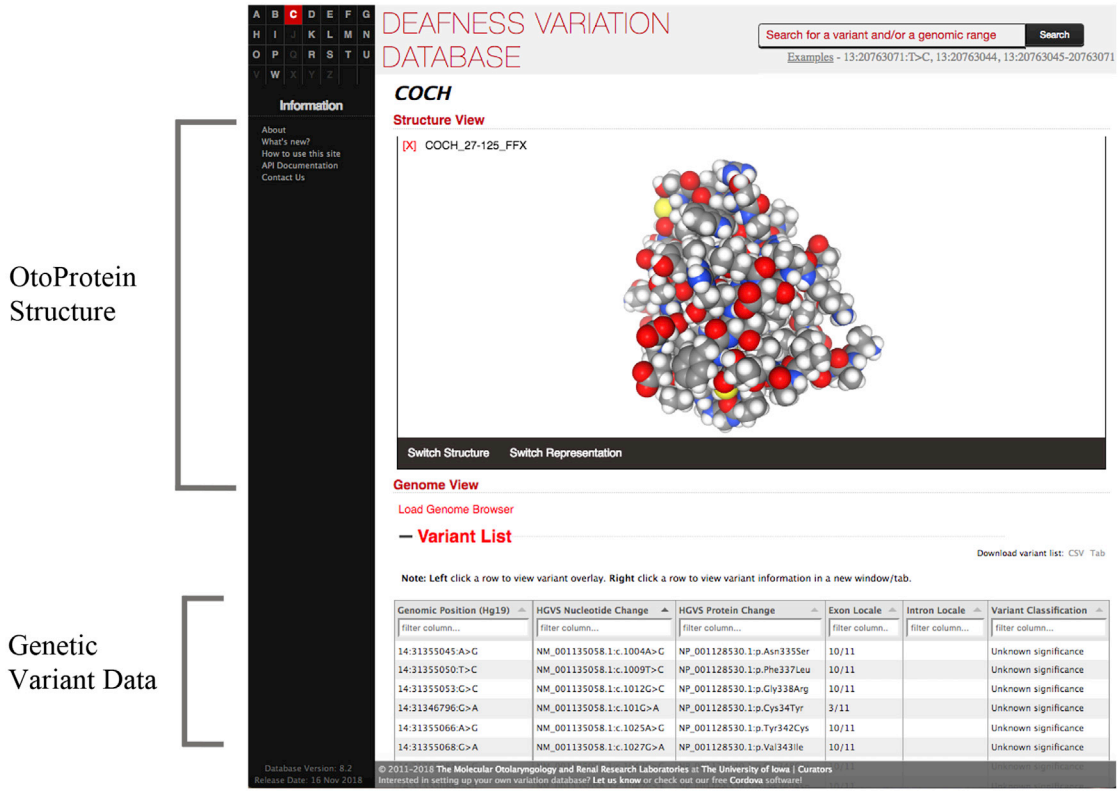

FIGURE 6 Incorporation of *OtoProtein* structures into the DVD. All models developed with the GPU-accelerated AMOEBA/GK protein repacking algorithm are publicly available in the DVD, where they can be viewed in combination with genomic and variant data. To see this figure in color, go online.

would require only ~140 days on a node equipped with four GPUs (i.e., ~2 weeks on our compute cluster, which has 10 such nodes).

A limitation of this repacking algorithm is its reliance on existing homology models to serve as initial coordinates. Although we have demonstrated this improves the quality of existing structural models, it does not provide coverage of proteins through ab initio or de novo techniques. This limitation is the subject of ongoing work based on GPU-accelerated biased sampling methods, which we are using to expand structural coverage of the OtoSCOPE proteome. Despite this limitation, the *OtoProtein* structural information is already being used to gain insight into the protein phenotype of missense variants associated with deafness.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.06.030.

AUTHOR CONTRIBUTIONS

Conceived the theory, M.R.T., J.M.L., and M.J.S.; performed the experiments, M.R.T., J.M.L., G.Q., C.E.O'C., and W.T.A.T.; analyzed the data, M.R.T., J.M.L., R.J.H.S., and M.J.S.; contributed code/tools/structures, M.R.T., J.M.L., G.Q., C.E.O'C., R.J.M., H.V.B., W.T.A.T., T.A.B., T.L.C., and M.J.S.; wrote the manuscript, M.R.T., J.M.L., W.T.A.T., T.A.B., T.L.C., R.J.H.S., and M.J.S.

ACKNOWLEDGMENTS

All computations were performed on The University of Iowa Argon cluster with support and guidance from Glenn Johnson, Ben Rogers, and Brenna Miller.

This material is based upon work supported by the NSF (National Science Foundation) Graduate Research Fellowship under Grant No. 000390183 to M.R.T. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the NSF. J.M.L. was supported by NIH/Center for Biocatalysis and Bioprocessing Award T32 GM0008365 and the University of Iowa Presidential Graduate Research Fellowship. M.J.S. was supported by NSF grant CHE-1751688. R.J.H.S., T.A.B., T.L.C., and M.J.S. were supported in part by NIH grant DC012049. G.Q. and C.E.O'C. were supported by fellowships from The University of Iowa Center for Research by Undergraduates.

REFERENCES

1. Shearer, A. E., and R. J. Smith. 2012. Genetics: advances in genetic testing for deafness. *Curr. Opin. Pediatr.* 24:679–686.

2. Shearer, A. E., E. A. Black-Ziegelbein, …, R. J. Smith. 2013. Advancing genetic testing for deafness with genomic technology. *J. Med. Genet.* 50:627–634.

3. Ephraim, S. S., N. Anand, …, T. A. Braun. 2014. Cordova: web-based management of genetic variation data. *Bioinformatics.* 30:3438–3439.
4. Shearer, A. E., R. W. Eppsteiner, …, R. J. Smith. 2014. Utilizing ethnic-specific differences in minor allele frequency to recategorize reported pathogenic deafness variants. *Am. J. Hum. Genet.* 95:445–453.
5. Richards, S., N. Aziz, …, H. L. Rehm; ACMG Laboratory Quality Assurance Committee. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the association for molecular pathology. *Genet. Med.* 17:405–424.
6. Rost, B. 1999. Twilight zone of protein sequence alignments. *Protein Eng.* 12:85–94.
7. Chen, M., X. Lin, …, P. G. Wolynes. 2018. Template-guided protein structure prediction and refinement using optimized folding landscape force fields. *J. Chem. Theory Comput.* 14:6102–6116.
8. Pieper, U., B. M. Webb, …, A. Sali. 2014. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* 42:D336–D346.
9. Bienert, S., A. Waterhouse, …, T. Schwede. 2017. The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res.* 45:D313–D319.
10. Berman, H. M., J. Westbrook, …, P. E. Bourne. 2000. The protein data bank. *Nucleic Acids Res.* 28:235–242.
11. Case, D. A., T. E. Cheatham, III, …, R. J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.
12. Hornak, V., R. Abel, …, C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
13. MacKerell, A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
14. Brooks, B. R., C. L. Brooks, III, …, M. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.
15. Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.
16. Kaminski, G. A., R. A. Friesner, …, W. L. Jorgensen. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B.* 105:6474–6487.
17. Ponder, J. W., and D. A. Case. 2003. Force fields for protein simulations. *Adv. Protein Chem.* 66:27–85.
18. Shi, Y., P. Ren, …, J.-P. Piquemal. 2015. Polarizable force fields for biomolecular modeling. *In* Reviews in Computational Chemistry. K. B. Lipkowitz, ed. John Wiley & Sons, Inc, pp. 51–86.
19. Ponder, J. W., C. Wu, …, T. Head-Gordon. 2010. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B.* 114:2549–2564.
20. Shi, Y., Z. Xia, …, P. Ren. 2013. The polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* 9:4046–4063.
21. Lemkul, J. A., J. Huang, …, A. D. MacKerell, Jr. 2016. An empirical polarizable force field based on the classical Drude oscillator model: development history and recent applications. *Chem. Rev.* 116:4983–5013.
22. Schnieders, M. J., N. A. Baker, …, J. W. Ponder. 2007. Polarizable atomic multipole solutes in a Poisson-Boltzmann continuum. *J. Chem. Phys.* 126:124114.
23. Schnieders, M. J., and J. W. Ponder. 2007. Polarizable atomic multipole solutes in a generalized Kirkwood continuum. *J. Chem. Theory Comput.* 3:2083–2097.
24. Aleksandrov, A., F. Y. Lin, …, A. D. MacKerell, Jr. 2018. Combining the polarizable Drude force field with a continuum electrostatic Poisson-Boltzmann implicit solvation model. *J. Comput. Chem.* 39:1707–1719.
25. Webb, B., and A. Sali. 2016. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics.* 54:5.6.1–5.6.37.
26. Adams, P. D., P. V. Afonine, …, P. H. Zwart. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66:213–221.
27. Park, H., S. Ovchinnikov, …, D. Baker. 2018. Protein homology model refinement by large-scale energy optimization. *Proc. Natl. Acad. Sci. USA.* 115:3054–3059.
28. Schnieders, M. J., T. D. Fenn, …, A. T. Brunger. 2009. Polarizable atomic multipole X-ray refinement: application to peptide crystals. *Acta Crystallogr. D Biol. Crystallogr.* 65:952–965.
29. Fenn, T. D., M. J. Schnieders, …, V. S. Pande. 2010. Polarizable atomic multipole X-ray refinement: hydration geometry and application to macromolecules. *Biophys. J.* 98:2984–2992.
30. Schnieders, M. J., T. D. Fenn, and V. S. Pande. 2011. Polarizable atomic multipole X-ray refinement: particle Mesh Ewald electrostatics for macromolecular crystals. *J. Chem. Theory Comput.* 7:1141–1156.
31. Fenn, T. D., M. J. Schnieders, …, A. T. Brunger. 2011. Reintroducing electrostatics into macromolecular crystallographic refinement: application to neutron crystallography and DNA hydration. *Structure.* 19:523–533.
32. Fenn, T. D., and M. J. Schnieders. 2011. Polarizable atomic multipole X-ray refinement: weighting schemes for macromolecular diffraction. *Acta Crystallogr. D Biol. Crystallogr.* 67:957–965.
33. Schnieders, M. J., T. S. Kaoud, …, P. Ren. 2012. Computational insights for the discovery of non-ATP competitive inhibitors of MAP kinases. *Curr. Pharm. Des.* 18:1173–1185.
34. Ren, P., J. Chun, …, N. A. Baker. 2012. Biomolecular electrostatics and solvation: a computational perspective. *Q. Rev. Biophys.* 45:427–491.
35. LuCore, S. D., J. M. Litman, …, M. J. Schnieders. 2015. Dead-end elimination with a polarizable force field repacks PCNA structures. *Biophys. J.* 109:816–826.
36. Davis, I. W., A. Leaver-Fay, …, D. C. Richardson. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35:W375–W383.
37. Chen, V. B., W. B. Arendall, III, …, D. C. Richardson. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66:12–21.
38. Friedrichs, M. S., P. Eastman, …, V. S. Pande. 2009. Accelerating molecular dynamic simulation on graphics processing units. *J. Comput. Chem.* 30:864–872.
39. Eastman, P., J. Swails, …, V. S. Pande. 2017. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* 13:e1005659.
40. Kaminsky, A. 2007. Parallel Java: a unified API for shared memory and cluster parallel programming in 100% Java. In 2007 IEEE International Parallel and Distributed Processing Symposium (IEEE), pp. 1–8.
41. Desmet, J., M. De Maeyer, …, I. Lasters. 1992. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature.* 356:539–542.
42. Goldstein, R. F. 1994. Efficient rotamer elimination applied to protein side-chains and related spin glasses. *Biophys. J.* 66:1335–1340.
43. Fiser, A., M. Feig, …, A. Sali. 2002. Evolution and physics in comparative protein structure modeling. *Acc. Chem. Res.* 35:413–421.
44. Kiefer, F., K. Arnold, …, T. Schwede. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* 37:D387–D392.
45. Eramian, D., N. Eswar, …, A. Sali. 2008. How well can the accuracy of comparative protein structure models be predicted? *Protein Sci.* 17:1881–1893.
46. Jing, Z., C. Liu, …, P. Ren. 2018. Many-body effect determines the selectivity for Ca^{2+} and Mg^{2+} in proteins. *Proc. Natl. Acad. Sci. USA.* 115:E7495–E7501.
47. Polydorides, S., and T. Simonson. 2013. Monte Carlo simulations of proteins at constant pH with generalized Born solvent, flexible sidechains, and an effective dielectric boundary. *J. Comput. Chem.* 34:2742–2756.
48. Rose, A. S., A. R. Bradley, …, P. W. Rose. 2018. NGL viewer: web-based molecular graphics for large complexes. *Bioinformatics.* 34:3755–3758.

Biophysical Journal, Volume *117*

Supplemental Information

Structural Insights into Hearing Loss Genetics from Polarizable Protein Repacking

Mallory R. Tollefson, Jacob M. Litman, Guowei Qi, Claire E. O'Connell, Matthew J. Wipfler, Robert J. Marini, Hernan V. Bernabe, William T.A. Tollefson, Terry A. Braun, Thomas L. Casavant, Richard J.H. Smith, and Michael J. Schnieders

Supplementary Information

Table S1. This list includes all genes published in version 8 of the Deafness Variation Database (http://deafnessvariationdatabase.org/).

ACTG1	COL2A1	FAM65B (RIPOR2)	KITLG	OTOA	SLC17A8	TWNK
ADCY1	COL4A3	FGF3	LARS2	OTOF	SLC22A4	USH1C
AIFM1	COL4A4	FGFR1	LHFPL5	OTOG	SLC26A4	USH1G
ALMS1	COL4A5	FGFR2	LOXHD1	OTOGL	SLC26A5	USH2A
ATP2B2	COL4A6	FOXI1	LOXL3	P2RX2	SLITRK6	WFS1
ATP6V1B1	COL9A1	GATA3	LRTOMT	PAX3	SMPX	WHRN
BDP1	COL9A2	GIPC3	MARVELD2	PCDH15	SNAI2	
BSND	CRYM	GJB2	MCM2	PDZD7	SOX10	
CABP2	DCDC2	GJB3	MET	PEX1	STRC	
CACNA1D	DFNA5 (GSDME)	GJB6	MIR-96	PEX6	SYNE4	
CCCP110	DFNB31	GPR98 (ADGRV1)	MITF	PJVK	TBC1D24	
CCDC50	DFNB59	GPSM2	MSRB3	PNPT1	TBX1	
CD164	DIABLO	GRHL2	MT-TL1	POLR1C	TCOF1	
CDC14A	DIAPH1	GRXCR1	MT-TS1	POLR1D	TECTA	
CDH23	DIAPH3	GRXCR2	MT-RNR1	POU3F4	TECTB	
CEACAM16	DSPP	HARS2	MYH14	POU4F3	TIMM8A	
CIB2	EDN3	HGF	MYH9	PRPS1	TJP2	
CISD2	EDNRB	HOMER2	MYO15A	PTPRQ	TMC1	
CLDN14	ELMOD3	HSD17B4	MYO3A	RDX	TMEM132E	
CLIC5	EPS8	ILDR1	MYO6	ROR1	TMIE	
CLPP	EPS8L2	KARS	MYO7A	S1PR2	TMPRSS3	
CLRN1	ESPN	KCNE1	NARS2	SCN7A	TNC	
COCH	ESRRB	KCNJ10	NLRP3	SERPINB6	TPRN	
COL11A1	EYA1	KCNQ1	OPA1	SIX1	TRIOBP	
COL11A2	EYA4	KCNQ4	OSBPL2	SIX5	TSPEAR	

Table S2. Related to Table 1. Decomposition of the many-body energy expansion (up to 3-body terms) for COCH domain residues 27-125 as function of generalized Kirkwood (GK) implicit solvent use and residue cutoff. Without GK, the energy expansion decays slowly, and the neglected energy terms sum to hundreds of kcal/mol (i.e. the neglected energy is the known Total AMOEBA/GK potential energy less explicitly calculated terms in the expansion). With GK, the energy expansion decays quickly, and neglected energy is on the order of only a few kcal/mol. The sum of 3-body energy terms when using GK is always less than 1 kcal/mol, but reaches 3.9 kcal/mol without GK.

GK Solvent	Cutoff (Å)	Backbone (kcal/mol)	Self-Energy Sum (kcal/mol)	2-Body Energy Sum (kcal/mol)	3-Body Energy Sum (kcal/mol)	Neglected Sum (kcal/mol)	Total Potential (kcal/mol)
No	2	-953.34	-1215.03	18.17	-0.19	302.76	-1847.63
No	3			-3.59	3.89	320.44	
No	6			163.77	1.77	155.19	
No	9			374.22	0.06	-53.54	
Yes	2	-2039.14	-1508.40	-99.69	-0.06	-11.88	-3659.19
Yes	3			-110.10	-0.13	-1.41	
Yes	6			-112.61	-0.30	1.27	
Yes	9			-110.88	-0.57	-0.19	

Table S3. Related to Table 1. Energy timings for a three-body expansion of the COCH domain residues 27-125.

Term	Number of Energies	% of Total Energies	Seconds
Self	947	0.01	262
Two-body	120,855	1.91	32,940
Three-body*	6,214,642	98.08	1,624,625
Total	6,336,444	100.00	1,657,827

*indicates extrapolation from finishing 3% of 3-body energies

Table S4. Related to Table 1. Residue cutoffs of 2, 3 and 6 Å are examined for three different structures. A comparison of 2 and 3 Å residue cutoffs (applied to residue pair distances from Eq. 7) shows no total potential energy differences greater than 2.5 kcal/mol and MolProbity scores do not vary by more than 0.04 Å.

Residue Cutoff (Å)	Gene and Residue Range	Total Potential	MolProbity
2	ACTG1 2-375	-16919.89	0.71
3	ACTG1 2-375	-16922.21	0.71
6	ACTG1 2-375	-16920.71	0.93
2	COCH 27-125	-3549.42	1.78
3	COCH 27-125	-3549.99	1.78
6	COCH 27-125	-3550.16	1.74
2	LRTOMT 118-287	-7112.79	1.40
3	LRTOMT 118-287	-7112.96	1.40
6	LRTOMT 118-287	-7117.19	1.42

Table S5. Related to Table 2. Energy evaluation timings for global side-chain optimization of residues 27-125 for the COCH domain using a varying number of GPUs (each node contains two Intel Xeon E5-2680v4 CPUs and one or more NVIDIA GTX 1080 TI GPUs) and the AMBER99sb/GB force field.

Computing Unit (# GPUs)	Time For Energies (sec)	Speed-Up (Relative to Using all CPU Cores and no GPU)
0 (CPUs only)	1034.0	1.0x
1	216.5	4.8x
2	124.5	8.3x
4	69.4	14.9x
8	43.5	23.8x
16	28.0	36.9x

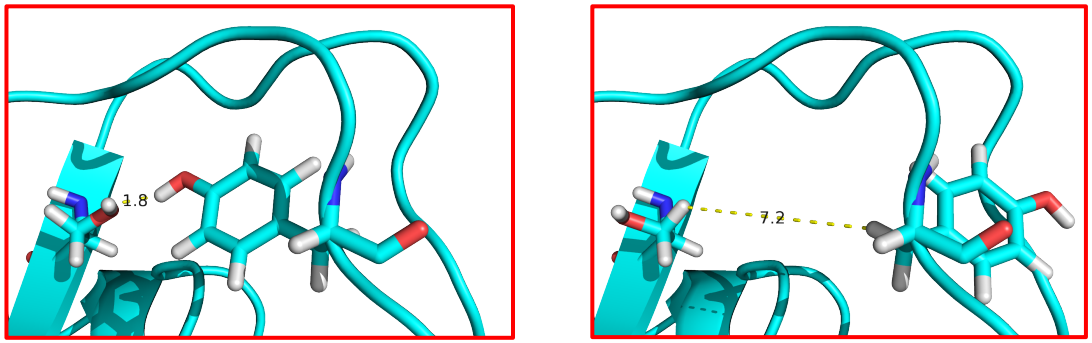

Figure S1. Related to Table 1B. Side-chain repacking distance cutoffs. The distances between two (or more) side-chains can vary widely as a function of the currently applied rotamers. The interaction above shows one rotamer pair with a 7.2 Å separation distance, while those same two side-chains have a separation distance of only 1.8 Å for a 2nd rotamer pair. The minimum distance between the two residues (i.e. from applying Equation 7) in this case is 1.8 Å, which is within a 2 Å cutoff.

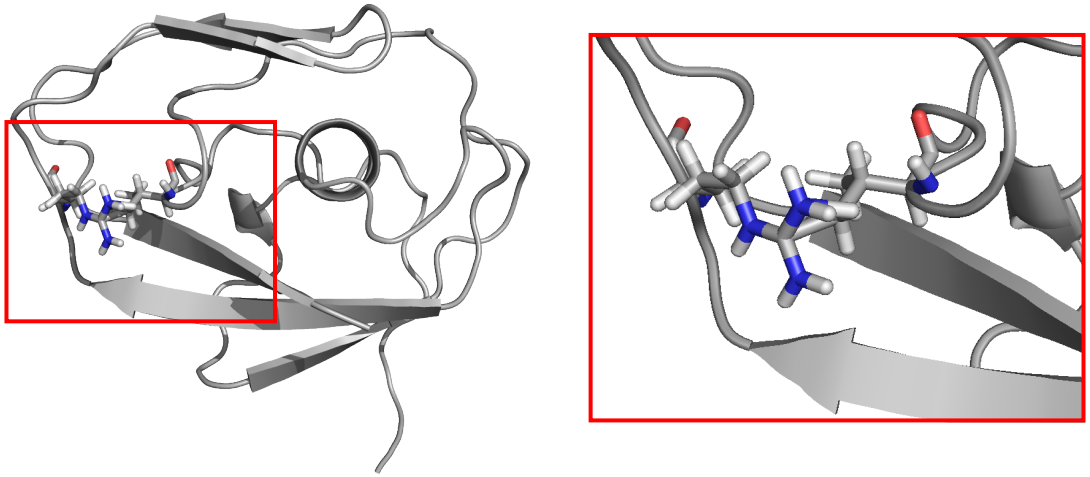

Figure S2. Related to Table 1B. Side-chain repacking clashes. Two interacting residues often have a clash for some rotamer permutations for one or more pairs of atoms. These clashes lead to residue separation distance of ~0 Å when applying Equation 7, which results in 2-body energies being calculated for all rotamer permutations.

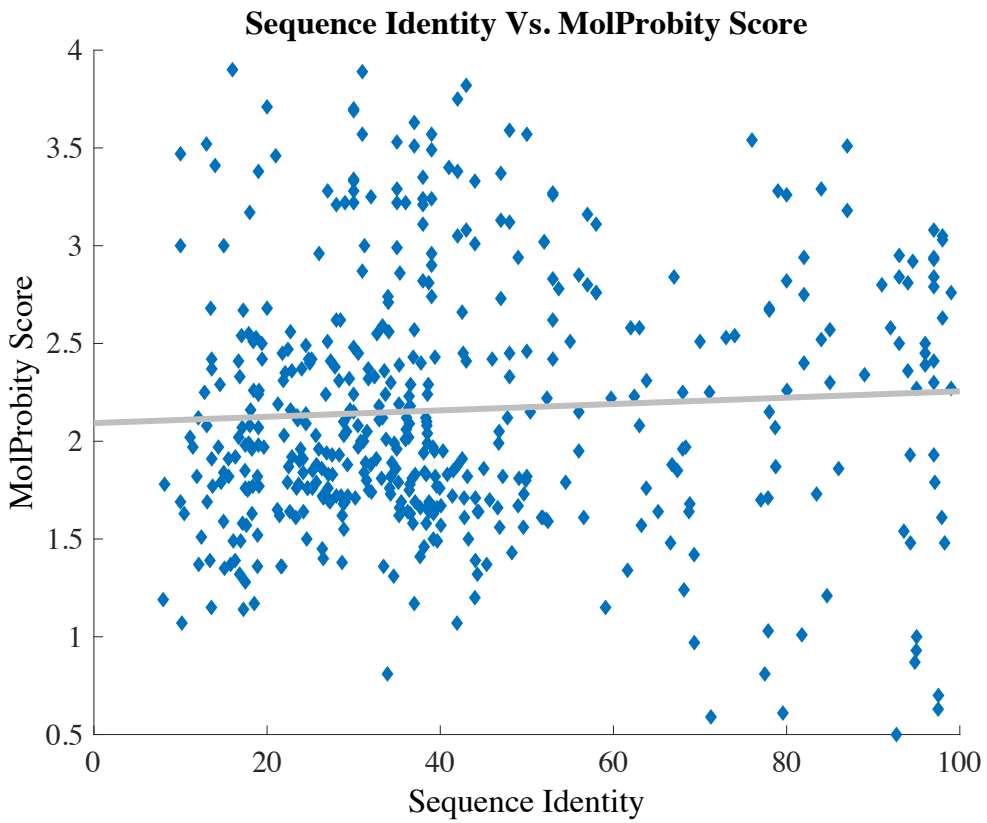

Figure S3. Related to Figure 3. Scatter plot of sequence identity vs MolProbity score for the 473 wildtype protein structures results in a linear regression R^2 of only 0.0035.

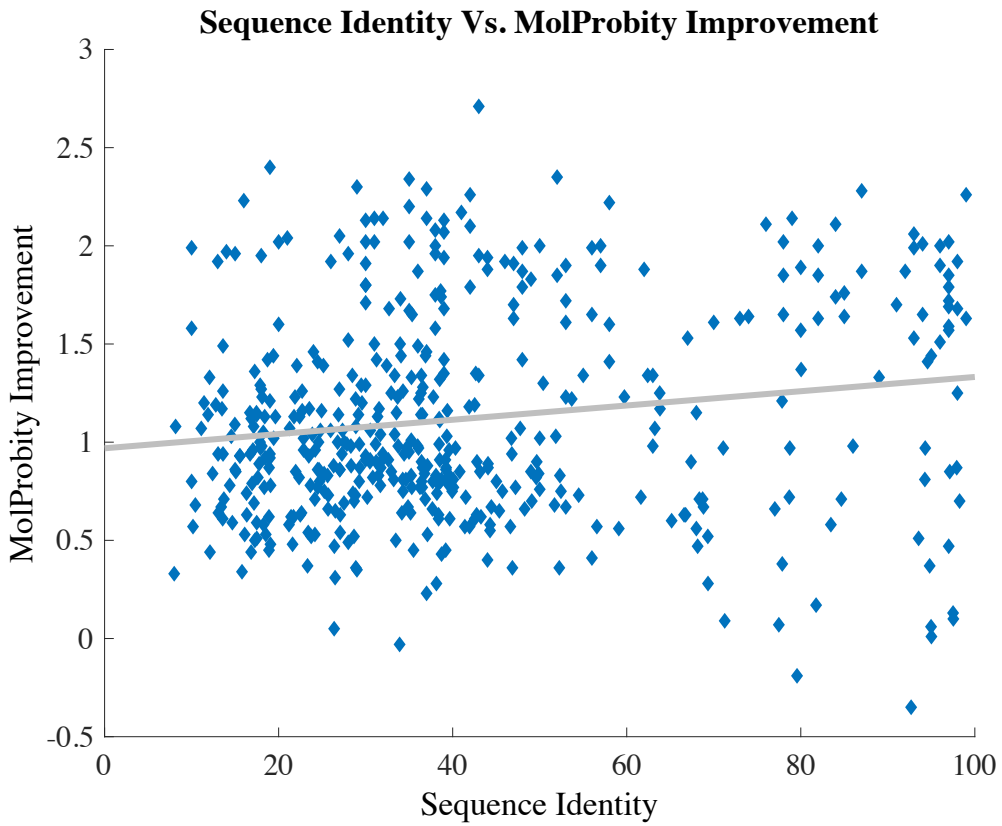

Figure S4. Related to Figure 3. Scatter plot of sequence identity vs. MolProbity improvement (*i.e.* the difference between original homology model MolProbity score compared to after AMOEBA/GK repacking) for the 473 wildtype protein structures results in a linear regression R^2 of only 0.0256.

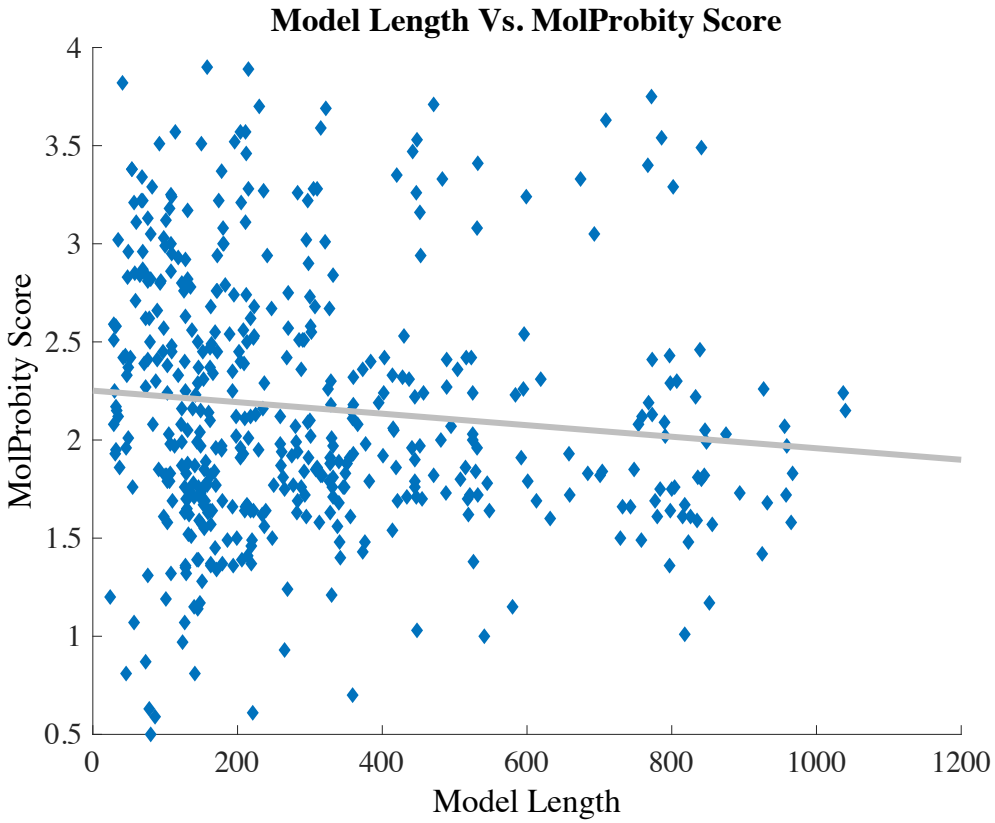

Figure S5. Related to Figure 3. Scatter plot of model length vs. MolProbity score for the 473 wildtype protein structures results in a linear regression R^2 of only 0.0119.

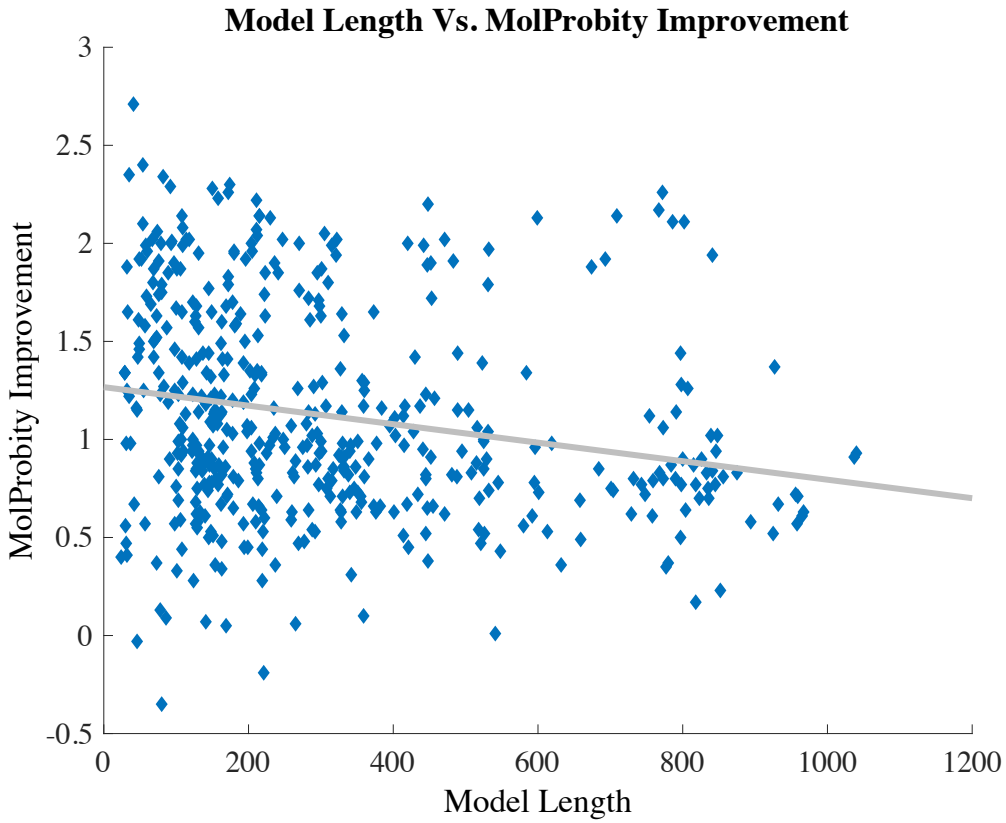

Figure S6. Related to Figure 3. Scatter plot of model length vs. MolProbity improvement for the 473 wildtype protein structures results in a linear regression R^2 of only 0.0449.

Table S6. Refinement statistics for all of the homology models analyzed in this work. The gene and residue ranges being modeled are shown in the two left-most columns, followed by the sequence identity to the template experimental structure the model was based off of, and then output from the MolProbity scoring algorithm. Table S6 is available in CSV format upon request.

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
ACTG1	2-375	94.0	Original	2.21	41.67	0.27	99.46	2.36	-16095.87
			FFX	1.89	0.00	0.00	98.92	0.71	-16919.89
ADCY1	288-477	74.0	Original	2.44	55.59	0.53	97.87	2.54	-8160.62
			FFX	1.22	0.00	0.00	95.21	0.90	-8758.46
	855-1057	42.7	Original	4.00	8.13	4.48	90.05	2.45	-8507.87
			FFX	1.71	0.00	1.49	93.53	1.10	-9004.61
AIFM1	122-611	95.0	Original	1.74	40.78	0.20	98.36	2.27	-20925.26
			FFX	0.74	0.00	0.20	95.29	0.83	-22227.11
ALMS1	3992-4150	16.0	Original	6.21	201.82	2.55	90.45	3.90	-7731.67
			FFX	5.52	0.00	2.55	87.90	1.67	-8383.55
ATP2B2	364-948	36.4	Original	2.17	7.71	1.54	89.88	2.23	-23907.04
			FFX	0.39	0.00	0.69	94.17	0.89	-25502.09
	49-1086	30.0	Original	3.42	4.59	3.47	87.84	2.24	-41790.12
			FFX	2.51	0.06	1.25	91.70	1.33	-45041.74
	49-1089	30.0	Original	3.19	4.26	2.41	89.70	2.15	-42612.58
			FFX	1.82	0.00	1.54	90.95	1.22	-45130.05
ATP6V1B1	42-497	77.0	Original	0.52	6.18	0.66	94.71	1.70	-19376.71
			FFX	1.29	0.28	0.44	94.93	1.04	-20325.82
ATP6V1B1	49-501	57.0	Original	3.65	87.65	1.11	95.34	3.16	-19140.93
			FFX	2.60	0.00	1.55	93.13	1.26	-20274.24
BDP1	247-376	37.1	Original	0.85	4.21	2.34	86.72	1.83	-6161.71
			FFX	0.85	0.00	0.00	92.97	0.95	-6439.74
	864-1069	28.0	Original	5.43	124.33	1.47	97.55	3.21	-11302.65
			FFX	3.80	0.00	0.00	95.59	1.25	-12098.05
BSND	240-313	28.0	Original	1.61	71.11	1.39	97.22	2.62	-4105.85
			FFX	1.61	0.00	0.00	93.06	1.10	-4305.94
CABP2	71-220	78.0	Original	0.00	46.55	0.00	98.65	2.15	-7773.50
			FFX	0.78	0.00	0.00	98.65	0.50	-8105.86
	72-220	77.9	Original	6.30	2.99	0.00	97.96	1.71	-7879.95
			FFX	0.00	0.00	0.00	98.64	0.50	-8130.21
	73-218	38.6	Original	8.87	5.22	1.39	95.83	2.29	-7431.79
			FFX	0.81	0.00	0.00	97.92	0.52	-7805.44

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	73-219	44.1	Original	1.60	1.30	1.38	94.48	1.39	-7642.12
			FFX	0.00	0.00	0.00	97.93	0.52	-7839.68
	76-216	44.1	Original	1.68	2.72	2.16	92.09	1.71	-7306.47
			FFX	0.00	0.45	1.44	97.12	0.82	-7534.06
CACNA1D	1548-1622	93.0	Original	3.08	50.75	0.00	95.89	2.84	-3392.78
			FFX	0.00	0.00	0.00	95.89	0.78	-3544.73
	1549-1622	94.8	Original	3.08	0.00	0.00	100.00	0.87	-3418.61
			FFX	0.00	0.00	0.00	100.00	0.50	-3535.79
CCDC50	77-159	35.0	Original	5.48	82.14	0.00	95.06	3.29	-4932.08
			FFX	1.37	0.00	0.00	95.06	0.95	-5165.75
CD164	56-105	26.0	Original	0.00	121.64	2.08	93.75	2.96	-2036.11
			FFX	2.13	0.00	2.08	95.83	1.04	-2159.47
CDC14A	11-343	67.0	Original	2.08	72.05	0.91	96.07	2.84	-14382.76
			FFX	3.82	0.00	0.91	94.56	1.31	-15526.89
	15-343	66.8	Original	2.11	4.29	1.22	93.58	1.88	-14707.84
			FFX	2.81	0.00	0.61	93.88	1.25	-15318.90
CDH23	1107-1626	28.7	Original	0.00	4.16	0.58	93.63	1.62	-24391.45
			FFX	0.45	0.00	0.39	93.63	0.92	-25458.39
	1321-1846	31.1	Original	0.44	7.97	2.48	89.50	2.00	-24475.87
			FFX	0.66	0.00	1.15	91.41	1.01	-25750.00
	1424-1953	31.2	Original	0.65	6.28	0.95	91.86	1.84	-25304.09
			FFX	0.43	0.00	1.52	93.18	0.94	-26868.44
	2211-2699	34.3	Original	0.24	5.37	0.82	93.22	1.73	-22811.90
			FFX	0.24	0.00	0.82	93.63	0.92	-23959.17
	240-772	28.5	Original	0.43	4.51	2.45	91.71	1.72	-23024.64
			FFX	1.07	0.00	1.13	92.66	0.98	-24278.44
	2503-2603	35.0	Original	1.14	91.82	2.02	90.91	2.99	-4297.13
			FFX	2.27	0.00	1.01	89.90	1.32	-4518.60
	2506-3051	23.7	Original	0.42	5.75	1.65	92.65	1.78	-23327.99
			FFX	1.26	0.00	1.29	93.57	1.00	-24699.75
	375-879	33.5	Original	2.68	7.97	2.78	87.87	2.36	-22035.26
			FFX	1.79	0.00	1.19	91.05	1.21	-23748.16
	782-1308	28.7	Original	1.52	1.12	1.14	93.52	1.38	-22585.79
			FFX	1.08	0.00	0.38	95.24	0.86	-23810.21
	887-1408	26.4	Original	3.79	0.50	2.88	87.88	1.72	-21387.14
			FFX	2.23	0.00	1.92	92.12	1.25	-22569.53
	894-1417	25.2	Original	3.33	7.86	3.45	88.12	2.42	-21311.52

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.67	0.00	0.96	90.61	1.03	-22767.11
	905-1413	31.5	Original	0.69	5.52	2.17	91.52	1.80	-22301.54
			FFX	0.92	0.00	0.59	92.50	0.97	-23487.82
	997-1528	23.9	Original	1.10	7.75	3.77	91.32	1.96	-22806.57
			FFX	0.66	0.00	0.94	93.58	0.92	-23896.39
CEACAM16	20-118	37.0	Original	1.22	55.06	0.00	95.88	2.57	-3994.18
			FFX	3.66	0.00	0.00	96.91	1.11	-4163.72
	20-421	16.3	Original	1.47	4.58	3.50	89.50	1.92	-14857.10
			FFX	2.36	0.00	2.50	91.50	1.29	-15858.62
CEACAM16-Dimer	19-422	19.4	Original	1.77	12.57	4.10	87.19	2.42	-29164.80
			FFX	2.95	0.08	1.87	91.42	1.40	-31738.99
	27-422	21.3	Original	1.20	10.73	3.30	88.83	2.19	-28921.58
			FFX	1.35	0.17	1.40	92.77	1.12	-31200.05
CIB2	139-174	52.0	Original	3.33	98.11	0.00	97.06	3.02	-2189.09
			FFX	0.00	0.00	0.00	97.06	0.67	-2239.22
	8-187	38.7	Original	1.22	3.12	3.93	91.01	1.69	-8802.87
			FFX	1.22	0.00	1.69	95.51	0.88	-9158.32
	9-187	39.3	Original	3.07	3.84	1.69	93.79	1.95	-8757.11
			FFX	1.84	0.00	1.13	96.61	0.92	-9142.42
CISD2	68-133	97.0	Original	3.33	29.92	1.56	92.19	2.84	-3208.38
			FFX	1.67	0.00	0.00	92.19	1.15	-3311.94
CLDN14	4-183	45.4	Original	1.36	4.34	1.12	97.75	1.37	-5385.07
			FFX	0.68	0.00	0.56	96.63	0.72	-5711.90
	4-184	43.0	Original	2.70	85.67	3.35	94.97	3.08	-5109.82
			FFX	0.00	0.36	3.35	91.62	1.13	-5632.88
CLIC5	162-409	78.0	Original	1.42	63.67	0.81	95.93	2.67	-11551.14
			FFX	0.94	0.00	0.41	97.15	0.65	-12269.96
CLPP	74-245	58.0	Original	2.76	76.35	0.00	97.65	2.76	-6754.55
			FFX	4.83	0.00	0.00	95.29	1.35	-7103.42
CLRN1	99-213	31.0	Original	4.00	120.43	0.88	89.38	3.57	-3651.63
			FFX	4.00	0.00	0.00	88.50	1.55	-4016.48
COCH	162-387	23.5	Original	0.52	10.24	4.46	87.95	2.13	-9177.87
			FFX	2.09	0.00	1.79	92.86	1.20	-9645.22
	162-544	25.7	Original	1.87	2.37	2.36	89.50	1.79	-15705.36
			FFX	1.56	0.00	1.57	92.13	1.13	-16649.60
	266-503	14.6	Original	2.46	11.51	2.12	93.64	2.29	-9283.22

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	2.96	0.00	2.54	94.07	1.26	-9764.98
	27-125	98.0	Original	2.50	70.10	1.03	93.81	3.03	-3659.32
			FFX	5.00	0.69	0.00	90.72	1.78	-3868.52
COL11A1	1500-1532	62.0	Original	4.76	38.65	0.00	100.00	2.58	-1222.53
			FFX	0.00	0.00	0.00	96.77	0.70	-1271.96
	67-229	28.9	Original	0.00	10.88	2.48	90.06	2.10	-7224.92
			FFX	0.00	0.00	0.62	94.41	0.88	-7500.58
COL11A1-Trimer	1595-1805	44.3	Original	0.18	3.18	2.72	90.72	1.64	-29011.08
			FFX	1.07	0.00	1.28	90.56	1.06	-30608.28
COL11A2	46-192	31.7	Original	3.31	6.87	5.52	88.28	2.37	-5528.01
			FFX	4.13	0.00	4.83	86.90	1.59	-5783.03
	919-950	63.0	Original	0.00	46.39	0.00	93.33	2.58	-1201.33
			FFX	5.26	0.00	0.00	90.00	1.60	-1284.63
COL11A2-Trimer	1542-1735	46.6	Original	1.00	3.65	2.26	91.49	1.66	-26654.76
			FFX	1.20	0.00	1.56	90.80	1.09	-28017.99
COL2A1	969-1000	97.0	Original	0.00	27.59	0.00	100.00	1.93	-1166.62
			FFX	5.00	0.00	0.00	93.33	1.46	-1247.06
COL2A1-Trimer	1248-1487	68.7	Original	1.78	2.80	0.99	94.51	1.64	-33839.58
			FFX	1.13	0.00	0.70	94.23	0.93	-35383.79
COL4A3	637-674	86.0	Original	0.00	23.31	0.00	100.00	1.86	-1333.20
			FFX	0.00	0.00	0.00	94.44	0.88	-1375.13
COL4A4	1465-1688	73.0	Original	0.53	50.91	0.90	95.05	2.53	-8666.36
			FFX	0.53	0.00	0.90	94.14	0.90	-9175.96
COL4A5	1461-1683	84.0	Original	2.63	38.68	0.45	97.29	2.52	-8494.09
			FFX	0.53	0.00	0.00	95.93	0.78	-9001.87
COL4A6	1466-1689	78.0	Original	2.07	41.50	0.45	95.05	2.68	-8416.80
			FFX	1.04	0.00	0.45	95.50	0.83	-8987.94
COL4A6-Hexamer	1467-1689	65.2	Original	1.04	4.50	0.90	94.12	1.64	-52381.81
			FFX	1.56	0.05	1.28	94.49	1.04	-54965.54
COL9A1	363-392	63.0	Original	5.26	10.26	0.00	100.00	2.08	-1085.20
			FFX	0.00	0.00	0.00	96.43	0.74	-1140.12
COL9A2	33-79	68.0	Original	0.00	26.55	0.00	97.78	1.96	-2055.17
			FFX	0.00	0.00	0.00	95.56	0.81	-2123.72
CRYM	4-314	30.0	Original	2.36	0.00	1.29	93.20	3.28	-12367.56
			FFX	4.33	0.00	0.65	91.59	1.48	-13220.35

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
DCDC2	129-226	96.0	Original	1.16	54.04	0.00	96.88	2.45	-4310.76
			FFX	1.16	0.00	0.00	100.00	0.55	-4545.30
	15-123	31.3	Original	5.38	18.22	3.74	83.18	3.00	-4650.71
			FFX	4.30	0.00	3.74	87.85	1.58	-4906.98
DIABLO	67-239	99.0	Original	3.38	44.82	0.58	96.49	2.76	-8041.05
			FFX	0.68	0.00	0.00	98.25	0.50	-8591.04
DIAPH1	762-1215	97.0	Original	4.13	44.35	1.33	94.91	2.94	-22098.48
			FFX	3.64	0.00	0.22	95.80	1.22	-23358.68
	92-452	94.3	Original	10.67	0.68	0.56	93.59	1.93	-16982.34
			FFX	2.44	0.00	0.28	95.26	1.12	-17823.43
DIAPH1-Dimer	142-483	94.3	Original	5.86	0.82	0.15	97.21	1.48	-33400.71
			FFX	0.81	0.00	0.00	97.94	0.51	-34668.15
DIAPH3	122-479	50.4	Original	4.55	4.10	1.69	92.70	2.15	-16786.33
			FFX	1.52	0.00	0.28	96.63	0.85	-17635.74
	633-1078	59.7	Original	4.06	4.69	2.25	91.22	2.22	-21919.74
			FFX	1.43	0.00	0.90	94.59	0.99	-23168.42
	916-1096	15.0	Original	2.99	91.18	0.56	96.65	3.00	-9348.95
			FFX	1.80	0.00	0.56	94.97	1.04	-9999.64
DIAPH3-Dimer	173-511	49.6	Original	2.88	1.18	0.74	94.51	1.56	-32558.65
			FFX	0.80	0.00	0.30	96.44	0.74	-33949.54
	637-1053	31.5	Original	2.27	6.40	1.69	93.37	2.05	-41250.90
			FFX	0.50	0.14	0.48	95.30	0.88	-43599.46
DSPP	88-318	30.0	Original	3.26	140.58	6.11	82.53	3.70	-12328.88
			FFX	2.72	0.00	3.06	79.48	1.57	-13123.69
EDN3	42-91	24.0	Original	0.00	24.19	2.08	91.67	2.37	-2258.58
			FFX	0.00	0.00	0.00	93.75	0.91	-2367.87
	97-121	44.0	Original	0.00	0.00	4.35	82.61	1.20	-1098.66
			FFX	0.00	0.00	0.00	95.65	0.80	-1156.49
EDNRB	105-402	30.0	Original	1.88	130.30	3.04	93.24	3.22	-9457.52
			FFX	4.51	0.00	1.69	91.22	1.51	-10589.24
ELMOD3	277-312	36.1	Original	6.25	1.69	0.00	88.24	2.12	-1423.19
			FFX	0.00	0.00	0.00	94.12	0.90	-1483.05
EPS8	699-788	97.0	Original	0.00	51.58	1.14	96.59	2.41	-3756.63
			FFX	0.00	0.00	0.00	97.73	0.56	-3909.23
EPS8-Dimer	59-188	44.3	Original	0.83	3.40	0.78	96.88	1.32	-12611.43
			FFX	0.00	0.00	0.00	96.09	0.77	-13017.20

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
EPS8L2	611-698	97.0	Original	1.28	47.01	0.00	97.67	2.30	-4135.37
			FFX	0.00	0.00	0.00	96.51	0.73	-4343.06
EPS8L2-Dimer	46-175	45.7	Original	3.23	2.16	0.78	95.31	1.70	-12837.81
			FFX	0.81	0.00	0.00	92.97	0.95	-13329.34
	495-550	63.8	Original	5.21	4.36	0.00	99.07	1.76	-5608.72
			FFX	1.04	0.00	0.00	99.07	0.51	-5800.21
ESPN	1-330	85.0	Original	2.00	38.75	0.00	98.17	2.30	-14391.48
			FFX	1.60	0.00	0.00	99.09	0.66	-14931.84
	1-331	84.7	Original	1.20	2.06	0.30	96.96	1.21	-14472.53
			FFX	0.00	0.00	0.00	98.18	0.50	-14916.48
	75-325	39.6	Original	1.03	6.54	0.00	93.98	1.77	-10938.03
			FFX	0.00	0.00	0.00	95.58	0.81	-11336.42
	90-324	35.2	Original	0.55	5.84	2.15	95.71	1.62	-10238.66
			FFX	0.55	0.29	0.43	98.28	0.61	-10659.53
ESRRB	100-429	36.5	Original	2.08	6.33	3.66	88.41	2.18	-14868.95
			FFX	1.04	0.00	1.22	90.85	1.04	-15556.17
	100-432	36.8	Original	1.03	5.71	2.42	91.84	1.81	-14774.04
			FFX	0.69	0.00	0.91	92.45	0.97	-15497.02
	101-431	36.0	Original	1.38	7.62	1.52	92.10	2.01	-14899.18
			FFX	1.04	0.19	1.22	93.01	1.03	-15734.16
	102-429	35.2	Original	1.74	3.00	2.76	91.10	1.79	-14802.81
			FFX	1.74	0.00	1.84	92.64	1.15	-15488.66
	102-430	32.6	Original	0.35	7.11	2.14	91.13	1.91	-14512.62
			FFX	1.39	0.00	0.92	94.19	1.00	-15256.71
	99-430	27.0	Original	1.72	3.33	1.21	92.73	1.76	-14609.35
			FFX	1.03	0.00	0.61	94.24	0.90	-15357.72
	99-431	27.0	Original	0.69	5.91	0.91	94.26	1.71	-14988.56
			FFX	1.03	0.00	1.21	95.17	0.85	-15782.58
ESRRB-Dimer	209-430	79.6	Original	1.02	0.27	0.00	98.64	0.61	-20082.23
			FFX	0.51	0.27	0.00	96.82	0.80	-20615.47
EYA1	322-592	85.0	Original	0.87	58.03	0.74	95.17	2.57	-11383.47
			FFX	0.87	0.00	0.37	95.54	0.81	-12281.43
EYA4	369-639	82.0	Original	2.16	51.02	0.37	95.54	2.75	-11545.13
			FFX	0.43	0.00	0.37	96.28	0.75	-12327.92
FGF3	38-86	53.0	Original	0.00	87.58	0.00	93.62	2.83	-1965.90
			FFX	2.50	0.00	4.26	93.62	1.22	-2085.72
FGF3-Dimer	34-196	44.4	Original	2.56	1.14	1.57	91.54	1.64	-14359.07

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.10	0.00	0.31	93.10	0.97	-15034.82
FGFR1	288-771	30.0	Original	2.59	120.34	2.49	92.32	3.33	-19802.51
			FFX	3.54	0.00	1.66	91.49	1.42	-21460.86
FGFR2	479-801	30.0	Original	3.85	152.48	3.12	87.85	3.69	-13668.07
			FFX	3.85	0.39	1.56	88.79	1.67	-15020.36
FOXI1	121-212	67.4	Original	3.70	1.37	1.11	90.00	1.85	-4245.86
			FFX	1.23	0.00	1.11	94.44	0.95	-4398.66
	122-220	56.6	Original	2.27	0.62	2.06	86.60	1.61	-5186.87
			FFX	1.14	0.00	3.09	91.75	1.04	-5375.18
GIPC3	105-198	57.0	Original	3.75	47.75	0.00	96.74	2.80	-4473.00
			FFX	0.00	0.00	0.00	95.65	0.80	-4679.23
	70-193	18.1	Original	1.90	7.74	2.46	90.98	2.16	-5419.67
			FFX	0.95	0.00	0.00	93.44	0.93	-5653.88
GIPC3-Dimer	37-306	68.2	Original	1.34	1.32	0.56	95.90	1.24	-24916.75
			FFX	0.67	0.00	0.37	96.08	0.77	-25856.15
GJB3	2-213	58.0	Original	2.12	83.69	2.38	92.38	3.11	-7705.75
			FFX	0.53	0.00	0.95	94.29	0.89	-8523.55
GJB6	2-217	79.0	Original	3.06	87.49	1.40	91.59	3.28	-7833.48
			FFX	2.04	0.00	1.87	93.93	1.14	-8699.71
GPR98(ADGRV1)	3254-3579	18.5	Original	0.35	13.47	1.23	87.04	2.26	-12294.56
			FFX	2.43	0.00	0.31	90.12	1.34	-13159.68
	4270-4339	39.0	Original	1.75	69.98	2.94	92.65	2.96	-3093.62
			FFX	3.51	0.00	1.47	86.76	1.54	-3302.16
	5593-5900	13.5	Original	2.58	21.10	3.27	90.20	2.68	-12773.41
			FFX	4.43	0.21	1.31	93.14	1.51	-13419.72
GPSM2	117-648	97.0	Original	5.66	52.45	1.51	95.47	3.08	-23631.57
			FFX	3.49	0.00	0.57	94.53	1.29	-25193.32
	169-509	34.4	Original	2.78	2.42	2.06	90.86	1.89	-15950.79
			FFX	0.69	0.00	0.59	92.92	0.95	-16716.64
GRHL2	248-489	82.0	Original	5.05	47.45	0.83	96.25	2.94	-10926.10
			FFX	2.29	0.00	1.25	95.42	1.09	-11718.01
GRXCR1	231-288	38.0	Original	1.96	68.60	1.79	83.93	3.21	-2450.97
			FFX	3.92	0.00	3.57	83.93	1.63	-2605.75
GRXCR1-Heteromer	91-218	25.2	Original	3.11	1.45	2.39	91.43	1.76	-22461.96
			FFX	0.44	0.00	1.20	93.63	0.92	-23762.37

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
GRXCR2	187-247	38.0	Original	3.85	56.37	1.69	93.22	3.11	-2673.31
			FFX	1.92	0.00	0.00	93.22	1.15	-2789.45
	195-240	39.4	Original	0.00	13.43	2.27	75.00	2.43	-2350.18
			FFX	0.00	0.00	2.27	77.27	1.27	-2442.84
GSDME(DFNA5)	17-125	32.0	Original	2.04	80.16	0.93	85.98	3.25	-4572.65
			FFX	1.02	0.00	1.87	87.85	1.11	-4910.86
HARS2	55-502	80.0	Original	5.68	77.41	1.35	95.29	3.26	-19109.03
			FFX	3.10	0.14	0.90	93.05	1.37	-20522.59
HARS2-Dimer	54-502	77.9	Original	0.39	0.90	0.56	96.20	1.03	-41206.82
			FFX	0.64	0.00	0.22	97.20	0.65	-42575.75
HGF	104-723	28.3	Original	1.82	8.42	3.88	85.28	2.31	-26789.58
			FFX	2.37	0.00	1.62	90.13	1.33	-28790.75
	128-720	42.5	Original	2.48	3.38	1.86	92.05	1.91	-26310.95
			FFX	2.10	0.22	0.51	92.39	1.30	-27701.16
	200-725	26.8	Original	1.72	8.72	4.39	88.17	2.24	-22816.06
			FFX	1.72	0.00	1.91	89.50	1.24	-24431.93
	210-288	53.0	Original	1.37	48.45	0.00	94.81	2.62	-4183.65
			FFX	4.11	0.00	0.00	93.51	1.39	-4359.49
	210-469	47.8	Original	2.95	4.99	2.33	91.47	2.12	-12744.38
			FFX	0.84	0.00	0.78	89.92	1.05	-13296.22
	305-720	36.1	Original	1.38	6.38	3.38	87.92	2.06	-17517.62
			FFX	0.55	0.16	0.97	90.82	1.09	-18807.03
	38-722	40.8	Original	1.15	5.90	1.02	92.68	1.83	-30590.94
			FFX	0.82	0.00	0.73	92.24	0.98	-32159.77
HOMER2	1-128	82.0	Original	2.65	32.96	0.79	97.62	2.40	-5558.27
			FFX	1.77	0.00	0.00	97.62	0.77	-5877.25
	1-142	77.5	Original	0.00	0.45	0.00	97.14	0.81	-6322.14
			FFX	0.00	0.00	0.71	96.43	0.74	-6492.85
	173-301	15.1	Original	0.83	5.14	0.00	97.64	1.35	-7330.05
			FFX	0.00	0.00	0.00	98.43	0.50	-7588.77
	178-341	15.8	Original	0.65	1.45	1.85	91.98	1.37	-9254.21
			FFX	0.65	0.00	0.62	90.74	1.03	-9684.81
	179-307	21.7	Original	1.64	0.93	1.57	93.70	1.36	-7472.92
			FFX	1.64	0.00	0.00	97.64	0.74	-7738.62
	179-337	18.4	Original	1.32	4.47	2.55	92.99	1.77	-9071.76
			FFX	0.66	0.00	0.00	91.72	1.00	-9402.83
	182-345	17.6	Original	0.00	5.05	0.62	95.68	1.57	-9688.88

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.00	0.00	0.00	96.30	0.75	-9973.96
	184-320	12.4	Original	1.54	1.74	0.74	93.33	1.51	-8141.61
			FFX	0.77	0.00	0.00	97.04	0.67	-8405.78
	189-316	10.2	Original	0.83	1.41	0.00	96.83	1.07	-7565.58
			FFX	0.83	0.00	0.00	99.21	0.50	-7760.58
	189-333	16.3	Original	2.17	0.82	0.70	94.41	1.39	-8360.12
			FFX	0.72	0.00	0.70	97.20	0.65	-8658.59
	218-350	18.9	Original	0.00	6.22	0.76	96.95	1.52	-8226.06
			FFX	1.56	0.00	0.00	99.24	0.65	-8474.48
HSD17B4	3-304	92.0	Original	3.36	50.82	0.00	98.00	2.58	-12922.94
			FFX	0.84	0.00	0.33	96.67	0.71	-13626.82
HSD17B4-Dimer	5-606	54.5	Original	1.35	3.97	1.92	91.25	1.79	-49290.79
			FFX	1.35	0.16	0.75	94.17	1.06	-52671.92
ILDR1	24-155	18.0	Original	2.54	83.29	0.00	92.31	3.17	-5393.05
			FFX	1.69	0.00	1.54	90.00	1.22	-5888.80
	24-166	22.4	Original	2.33	8.75	7.80	80.85	2.47	-6223.82
			FFX	2.33	0.00	2.13	89.36	1.34	-6551.44
KARS	124-595	20.0	Original	5.30	153.33	2.98	91.28	3.71	-20476.23
			FFX	6.27	0.00	1.06	88.51	1.69	-22024.15
KCNE1	1-129	98.0	Original	3.48	35.44	0.79	96.85	2.63	-5157.13
			FFX	1.74	0.00	0.00	96.06	0.95	-5480.14
KCNJ10	27-348	44.0	Original	2.12	88.16	1.25	95.00	3.01	-11841.26
			FFX	1.77	0.00	0.62	94.38	1.07	-12971.09
KCNJ10-Tetramer	26-343	43.1	Original	0.36	8.48	1.27	94.70	1.82	-49492.06
			FFX	1.08	0.20	0.32	94.62	0.97	-52624.89
	28-338	45.0	Original	2.47	3.42	2.02	93.45	1.86	-47479.16
			FFX	1.56	0.00	0.40	93.69	1.06	-50941.79
KCNQ1	352-535	97.0	Original	7.14	45.66	0.55	98.35	2.79	-7901.49
			FFX	2.98	0.00	0.00	95.05	1.20	-8417.19
KCNQ4	317-369	46.0	Original	2.27	47.03	0.00	98.04	2.42	-2222.05
			FFX	0.00	0.00	0.00	98.04	0.50	-2340.36
	80-328	24.6	Original	0.97	5.46	0.40	96.76	1.50	-8543.28
			FFX	0.00	0.00	0.00	98.79	0.50	-9002.39
KITLG	35-166	80.0	Original	3.15	50.00	0.77	96.15	2.82	-5854.88
			FFX	3.15	0.00	0.00	94.62	1.25	-6186.32
LARS2	50-902	37.0	Original	0.95	0.67	1.53	92.83	1.17	-35950.27

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.09	0.00	0.59	93.89	0.94	-37519.62
	55-835	23.3	Original	2.24	0.89	3.08	89.09	1.61	-32531.02
			FFX	1.79	0.08	1.67	91.14	1.24	-33988.55
	60-901	39.0	Original	3.04	130.50	2.50	90.00	3.49	-32663.97
			FFX	4.28	0.00	1.31	89.29	1.55	-36443.15
LHFPL5	131-190	34.0	Original	2.17	57.21	1.72	96.55	2.71	-2199.09
			FFX	2.17	0.00	0.00	96.55	0.98	-2313.81
LOXHD1	1253-1371	32.4	Original	2.94	10.42	1.71	93.16	2.33	-5715.04
			FFX	0.00	0.00	0.85	93.16	0.94	-5961.18
	1591-1784	12.8	Original	1.16	14.44	3.12	90.62	2.25	-8879.86
			FFX	1.16	0.00	2.60	91.15	1.06	-9306.84
	173-281	35.3	Original	2.25	19.52	1.87	73.83	2.86	-5567.95
			FFX	1.12	0.00	0.93	84.11	1.21	-5837.29
	1861-2065	26.0	Original	2.22	3.63	3.45	90.15	1.96	-9032.35
			FFX	0.00	0.00	0.99	94.09	0.90	-9621.85
	1888-2057	32.7	Original	2.68	12.79	1.79	87.50	2.55	-7601.60
			FFX	0.67	0.00	0.60	94.64	0.87	-8028.69
	275-384	38.0	Original	2.13	82.75	0.93	87.96	3.24	-5388.92
			FFX	0.00	0.00	0.00	85.19	1.16	-5736.96
	297-430	32.1	Original	0.00	5.55	2.27	93.18	1.74	-5506.01
			FFX	0.00	0.00	0.76	95.45	0.82	-5731.04
	554-685	32.0	Original	0.00	6.80	2.31	91.54	1.88	-6692.70
			FFX	0.00	0.00	0.77	93.08	0.94	-6941.61
	781-935	28.9	Original	0.74	2.81	3.27	92.16	1.55	-7237.93
			FFX	2.22	0.00	0.65	93.46	1.19	-7516.26
	816-947	29.2	Original	0.00	9.56	1.54	90.00	2.05	-6291.13
			FFX	0.00	0.00	0.77	93.85	0.91	-6493.64
LOXL3	167-527	32.9	Original	1.41	9.89	2.51	89.97	2.18	-14710.24
			FFX	0.70	0.00	1.67	93.31	0.93	-15690.38
	184-388	37.8	Original	4.40	5.94	3.94	88.67	2.40	-8099.56
			FFX	1.89	0.00	1.48	92.61	1.17	-8609.37
	307-744	63.8	Original	8.31	3.76	1.15	93.12	2.31	-19944.74
			FFX	2.22	0.00	0.69	94.50	1.14	-21472.74
	42-408	30.5	Original	1.71	4.38	4.93	83.56	2.08	-15337.46
			FFX	1.37	0.00	3.29	89.04	1.18	-16396.54
	43-144	48.0	Original	2.47	80.13	4.00	93.00	3.12	-4605.90
			FFX	2.47	0.00	1.00	93.00	1.25	-4830.31

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
LRTOMT	118-287	26.4	Original	0.72	2.23	1.79	92.86	1.45	-6833.86
			FFX	3.60	0.00	1.19	92.26	1.40	-7112.79
	75-290	39.3	Original	0.00	5.61	0.00	94.86	1.66	-8886.91
			FFX	0.56	0.00	0.93	95.79	0.79	-9250.56
	76-290	37.7	Original	0.00	2.66	0.47	94.84	1.41	-8725.89
			FFX	1.13	0.00	0.47	96.71	0.75	-9040.89
	77-290	38.6	Original	0.00	4.17	0.47	92.45	1.67	-8760.98
			FFX	1.14	0.00	0.47	95.28	0.87	-9099.92
	79-291	39.0	Original	1.15	86.07	0.95	95.73	2.74	-8470.37
			FFX	4.60	0.00	0.95	94.31	1.39	-9039.09
MARVELD2	442-551	30.0	Original	1.90	62.63	0.93	98.15	2.48	-6246.62
			FFX	3.81	0.00	1.85	96.30	1.19	-6593.96
MCM2	189-821	52.3	Original	0.73	3.02	1.90	91.28	1.60	-27240.82
			FFX	2.20	0.00	1.27	92.23	1.24	-28883.93
	192-805	37.1	Original	3.02	1.04	1.14	90.85	1.69	-27107.95
			FFX	1.89	0.00	0.98	92.97	1.16	-28427.32
	199-798	39.0	Original	4.26	92.25	1.51	95.32	3.24	-25159.52
			FFX	1.55	0.00	1.00	92.64	1.11	-27492.89
MET	1082-1380	39.0	Original	1.13	100.40	2.02	94.28	2.90	-12426.36
			FFX	1.89	0.00	1.68	91.25	1.22	-13390.96
	44-919	25.6	Original	1.77	4.10	4.81	85.47	2.03	-34865.68
			FFX	1.52	0.00	1.26	89.70	1.20	-37452.36
MITF	323-402	96.0	Original	4.11	35.23	0.00	100.00	2.50	-4301.54
			FFX	0.00	0.00	0.00	98.72	0.50	-4477.61
MITF-Dimer	324-402	97.5	Original	0.00	0.37	0.00	99.35	0.63	-8691.33
			FFX	0.00	0.00	0.00	98.70	0.50	-8958.68
MSRB3	36-171	53.7	Original	6.72	7.20	5.97	78.36	2.78	-6116.73
			FFX	3.36	0.00	2.24	85.07	1.56	-6440.73
	41-167	58.0	Original	2.70	47.37	0.00	96.00	2.76	-5823.79
			FFX	2.70	0.00	0.00	95.20	1.16	-6074.23
MYH14	15-863	46.8	Original	6.79	1.54	1.77	92.80	1.99	-36828.72
			FFX	1.39	0.00	1.18	94.69	0.97	-39251.64
	17-863	46.8	Original	6.94	1.84	2.25	92.54	2.05	-36762.29
			FFX	1.94	0.00	1.30	94.20	1.11	-39080.37
	18-863	47.2	Original	5.01	0.96	2.49	91.47	1.82	-37086.67
			FFX	1.53	0.00	0.95	93.84	1.05	-39116.26
	18-983	38.4	Original	1.47	1.16	3.63	86.72	1.58	-42927.11

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.98	0.00	1.97	92.32	0.97	-45619.24
	18-985	38.4	Original	2.32	1.68	3.52	86.44	1.83	-43098.83
			FFX	1.46	0.13	1.86	90.99	1.20	-45878.51
	21-860	50.0	Original	6.44	5.28	2.74	89.62	2.46	-33826.45
			FFX	4.20	0.00	1.43	92.48	1.44	-38511.85
	21-862	49.9	Original	2.37	2.82	1.90	92.98	1.80	-37322.04
			FFX	1.26	0.00	1.07	94.29	0.96	-39078.35
	22-947	69.3	Original	1.28	1.48	1.52	92.86	1.42	-41912.86
			FFX	0.77	0.07	1.08	94.59	0.90	-43756.53
	22-954	68.8	Original	6.20	0.33	1.83	92.91	1.68	-42074.66
			FFX	2.03	0.00	0.75	96.03	1.01	-44297.87
	24-983	68.3	Original	6.39	1.04	2.30	89.67	1.97	-43333.54
			FFX	2.70	0.00	1.25	93.32	1.26	-45836.83
	46-832	76.0	Original	4.91	110.75	2.55	91.46	3.54	-33827.51
			FFX	3.42	0.16	1.27	92.74	1.43	-36806.16
	804-955	17.5	Original	0.00	0.80	4.00	90.67	1.28	-7308.23
			FFX	0.00	0.00	2.00	96.00	0.77	-7614.00
	804-957	17.5	Original	1.56	3.15	1.97	91.45	1.76	-7430.33
			FFX	0.00	0.00	0.66	97.37	0.62	-7775.44
MYH9	1-895	83.5	Original	6.91	0.41	1.34	93.06	1.73	-40874.20
			FFX	2.56	0.00	0.34	95.18	1.15	-43224.10
	3-836	52.3	Original	7.16	2.73	1.92	91.35	2.22	-37113.74
			FFX	3.99	0.00	1.20	93.27	1.39	-39398.78
	3-838	52.4	Original	2.20	1.77	1.68	94.12	1.59	-37995.78
			FFX	0.96	0.00	0.60	95.08	0.84	-39856.15
	3-839	49.1	Original	4.39	1.40	1.44	92.93	1.81	-37891.89
			FFX	2.47	0.00	0.72	95.57	1.11	-39763.09
	3-930	80.0	Original	13.70	1.79	1.30	92.76	2.26	-42488.88
			FFX	1.23	0.00	0.54	95.36	0.89	-45482.92
	3-959	78.7	Original	8.60	1.03	2.30	89.53	2.07	-44553.71
			FFX	3.70	0.00	0.73	93.72	1.35	-46912.41
	3-961	41.4	Original	3.34	0.71	2.40	88.30	1.72	-44363.09
			FFX	1.67	0.00	1.67	92.27	1.15	-47112.65
	4-822	81.8	Original	0.98	0.45	0.73	94.86	1.01	-37633.44
			FFX	0.70	0.00	0.24	95.10	0.84	-39341.73
	6-808	84.0	Original	3.72	81.41	1.50	92.38	3.29	-35717.79
			FFX	2.87	0.00	0.62	95.26	1.18	-38513.52

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
MYO15A	1215-1973	39.7	Original	1.20	2.51	1.45	94.06	1.49	-32732.14
			FFX	0.30	0.08	0.40	94.98	0.88	-34414.53
	1217-1946	43.3	Original	0.47	3.30	1.24	94.51	1.50	-31766.64
			FFX	1.25	0.00	0.69	95.60	0.88	-32980.55
	1217-1949	37.7	Original	1.86	2.19	2.05	92.61	1.66	-31584.53
			FFX	0.78	0.00	0.82	94.80	0.86	-33085.18
	1217-1960	35.3	Original	1.98	2.73	1.75	94.47	1.66	-31898.35
			FFX	1.07	0.00	0.81	94.61	0.89	-33583.03
	1217-1971	38.5	Original	3.16	4.39	2.39	92.16	2.08	-32178.53
			FFX	1.05	0.08	1.06	93.76	0.96	-33927.40
	1217-1976	38.3	Original	5.83	1.69	4.49	87.07	2.12	-31936.39
			FFX	2.39	0.08	0.92	91.42	1.33	-34070.24
	1224-1898	44.0	Original	2.69	110.88	3.57	91.68	3.33	-27856.11
			FFX	3.87	0.00	2.08	91.38	1.45	-30070.82
	1224-1972	41.5	Original	3.33	1.80	2.14	91.03	1.85	-31233.69
			FFX	1.82	0.00	0.94	93.44	1.13	-33390.91
MYO3A	293-1084	36.4	Original	1.56	7.09	1.90	92.15	2.02	-34859.62
			FFX	0.71	0.00	1.01	94.43	0.88	-36639.40
	337-1105	35.2	Original	4.54	3.86	2.74	91.00	2.19	-32598.46
			FFX	3.66	0.00	1.04	93.48	1.36	-35047.45
	337-1121	31.7	Original	2.44	1.81	2.81	91.19	1.75	-34283.28
			FFX	1.00	0.00	0.89	94.51	0.88	-36029.26
	337-1134	33.5	Original	1.55	0.93	1.38	93.22	1.36	-35604.20
			FFX	0.84	0.08	0.50	95.23	0.86	-37039.84
	337-1141	34.2	Original	2.23	2.15	2.37	91.41	1.76	-35112.19
			FFX	1.81	0.00	0.62	93.52	1.12	-37088.28
	337-1144	34.3	Original	4.87	4.13	2.73	88.59	2.30	-34309.50
			FFX	1.25	0.00	0.87	92.43	1.04	-37189.39
	341-1141	36.4	Original	1.68	3.01	1.88	91.99	1.75	-34880.48
			FFX	0.98	0.00	0.38	94.99	0.85	-36959.65
	9-292	53.0	Original	4.40	75.40	0.71	93.62	3.26	-12719.88
			FFX	4.80	0.22	0.35	93.26	1.54	-13445.29
MYO6	1-768	41.0	Original	3.55	122.72	1.96	93.34	3.40	-33402.75
			FFX	2.36	0.00	0.91	92.95	1.23	-36924.83
	1166-1294	94.6	Original	26.79	3.72	4.72	84.25	2.92	-5982.26
			FFX	5.36	0.00	0.79	92.91	1.51	-6479.49
	2-825	98.3	Original	7.26	0.45	0.49	97.08	1.48	-38294.58

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.23	0.00	0.49	96.72	0.78	-40215.77
	4-819	97.9	Original	4.16	1.37	0.61	96.07	1.61	-37813.07
			FFX	1.11	0.00	0.25	96.68	0.74	-39722.30
	61-833	42.0	Original	6.55	139.62	3.63	90.79	3.75	-14162.08
			FFX	3.49	0.16	1.69	91.05	1.49	-36786.67
MYO7A	2-792	36.4	Original	5.29	2.74	2.28	92.27	2.09	-34567.67
			FFX	3.15	0.00	0.89	93.79	1.29	-36594.65
	2-800	36.2	Original	1.13	4.10	1.00	93.98	1.64	-35215.32
			FFX	1.42	0.00	0.50	96.24	0.87	-37243.87
	3-800	36.9	Original	6.68	5.35	2.51	91.21	2.43	-34400.84
			FFX	1.14	0.00	1.13	92.96	0.99	-37168.60
	3-801	36.6	Original	6.38	3.64	2.01	90.84	2.29	-34816.36
			FFX	1.42	0.00	0.88	94.10	1.01	-37198.43
	3-859	40.1	Original	1.72	2.15	0.70	94.04	1.57	-38082.79
			FFX	0.80	0.07	0.70	96.49	0.76	-39912.59
	35-861	42.8	Original	1.79	2.15	0.97	93.45	1.61	-36245.95
			FFX	0.69	0.00	0.61	96.73	0.71	-38554.77
	56-874	40.1	Original	3.47	0.74	2.33	90.82	1.67	-36058.68
			FFX	0.69	0.07	0.49	94.61	0.90	-37787.06
	993-1686	98.0	Original	4.49	66.25	0.58	96.10	3.05	-28861.43
			FFX	2.66	0.00	0.14	95.52	1.13	-31460.80
NARS2	28-476	35.0	Original	4.31	109.51	2.46	90.38	3.53	-18930.09
			FFX	2.54	0.00	0.89	91.05	1.33	-20500.29
NARS2-Dimer	24-476	30.5	Original	0.63	8.26	2.00	91.01	1.97	-39334.62
			FFX	1.26	0.07	1.44	92.90	1.06	-41771.93
	3-474	18.9	Original	1.33	3.66	1.49	89.15	1.82	-40516.60
			FFX	1.69	0.00	1.17	90.64	1.20	-43131.52
NLRP3	699-1031	39.2	Original	1.67	8.70	0.00	95.17	1.97	-13151.50
			FFX	2.00	0.00	0.30	95.17	1.06	-13800.50
	725-897	48.0	Original	0.65	65.32	1.75	97.08	2.45	-6555.98
			FFX	1.31	0.00	0.58	97.66	0.66	-7059.46
OPA1	262-707	22.4	Original	0.75	5.48	1.35	91.89	1.79	-20127.19
			FFX	1.00	0.14	1.13	93.69	0.97	-21168.48
	262-709	26.5	Original	0.50	6.44	1.57	94.84	1.71	-20217.92
			FFX	1.73	0.00	0.45	94.39	1.06	-21382.78
	272-577	27.0	Original	2.59	112.97	2.30	93.09	3.28	-12921.66
			FFX	2.59	0.00	0.66	93.75	1.23	-13886.11

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
OPA1-Dimer	261-706	24.7	Original	1.00	5.27	2.25	92.22	1.76	-39892.72
			FFX	0.87	0.14	0.79	93.91	0.96	-42168.82
OPA1-Tetramer	262-707	23.7	Original	1.43	5.45	1.91	92.06	1.90	-80139.24
			FFX	0.81	1.76	0.90	93.02	1.38	-83370.60
OSBPL2	59-480	35.5	Original	1.86	2.35	2.14	92.38	1.69	-19582.91
			FFX	2.13	0.00	1.19	91.90	1.24	-20470.55
	60-480	38.0	Original	2.67	105.26	3.58	90.21	3.35	-18445.66
			FFX	1.60	0.29	1.67	88.78	1.35	-20083.10
	84-116	36.4	Original	0.00	7.29	3.23	77.42	2.17	-1363.63
			FFX	0.00	0.00	3.23	93.55	0.92	-1428.61
OTOA	1-533	14.0	Original	3.40	137.94	1.32	93.97	3.41	-20815.35
			FFX	4.26	0.12	0.94	93.79	1.44	-22702.08
OTOF	1-124	91.0	Original	2.65	46.12	0.82	95.08	2.80	-5804.22
			FFX	2.65	0.00	0.00	95.90	1.10	-6024.71
	250-542	18.4	Original	3.05	11.43	2.75	89.35	2.51	-12824.48
			FFX	2.67	0.00	2.06	89.69	1.38	-13566.43
	251-543	24.2	Original	0.00	3.67	4.81	83.85	1.84	-12738.80
			FFX	1.91	0.00	2.41	87.97	1.31	-13388.82
	950-1095	19.4	Original	1.60	16.47	3.47	86.81	2.50	-6301.44
			FFX	0.00	0.00	4.17	89.58	1.06	-6578.52
	951-1226	22.9	Original	0.42	5.25	5.47	86.13	1.92	-11426.39
			FFX	0.00	0.00	1.46	92.70	0.96	-12051.09
OTOG	1244-1398	17.5	Original	1.63	5.42	0.65	94.12	1.85	-6253.09
			FFX	0.81	0.00	0.00	96.73	0.70	-6486.41
	1285-1398	19.6	Original	0.00	8.92	0.89	91.96	1.97	-4760.31
			FFX	1.09	0.00	0.00	95.54	0.84	-4926.79
	1880-2032	21.8	Original	2.61	9.37	3.97	86.09	2.45	-4972.10
			FFX	1.74	0.00	1.99	86.09	1.32	-5309.56
	408-511	33.7	Original	5.68	2.05	2.94	83.33	2.24	-4088.90
			FFX	0.00	0.00	0.98	91.18	1.01	-4371.26
	768-877	30.5	Original	5.43	4.65	7.41	85.19	2.45	-4836.07
			FFX	3.26	0.00	2.78	91.67	1.39	-5066.96
	906-982	47.0	Original	4.48	78.36	0.00	96.00	3.13	-2999.11
			FFX	5.97	0.00	0.00	97.33	1.22	-3184.02
OTOGL	1242-1352	27.3	Original	1.00	5.03	0.92	93.58	1.69	-5058.92
			FFX	1.00	0.00	0.92	96.33	0.75	-5255.68
	394-470	38.7	Original	10.45	7.51	5.33	86.67	2.81	-3138.40

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.49	0.00	1.33	93.33	1.07	-3315.85
	480-602	13.0	Original	0.00	10.20	3.31	90.08	2.08	-4681.90
			FFX	0.93	0.00	1.65	85.95	1.14	-4924.05
	724-827	27.8	Original	5.32	4.61	3.92	88.24	2.38	-4142.00
			FFX	2.13	0.00	1.96	86.27	1.39	-4399.41
	846-935	42.5	Original	10.39	3.11	10.23	76.14	2.66	-3161.11
			FFX	2.60	0.00	1.14	85.23	1.47	-3344.74
P2RX2	120-178	56.0	Original	4.08	82.64	0.00	98.25	2.85	-2370.78
			FFX	0.00	0.00	0.00	94.74	0.86	-2504.03
P2RX2-Trimer	25-381	51.7	Original	0.75	4.84	0.56	94.84	1.61	-43180.54
			FFX	0.54	0.00	0.28	93.43	0.93	-45458.83
PAX3	111-274	21.6	Original	1.36	0.37	3.09	87.04	1.36	-8938.18
			FFX	0.68	0.00	0.00	94.44	0.88	-9344.49
	112-272	24.5	Original	1.39	6.87	2.52	85.53	2.14	-8746.00
			FFX	2.78	0.00	1.89	93.08	1.28	-9164.32
	112-276	24.5	Original	4.73	6.26	4.91	86.50	2.49	-8950.34
			FFX	1.35	0.00	1.23	92.02	1.08	-9401.07
	124-277	21.9	Original	4.38	4.77	4.61	88.82	2.31	-8297.61
			FFX	2.19	0.00	0.00	95.39	1.08	-8708.82
	295-476	10.0	Original	0.61	92.78	3.33	89.44	3.00	-5896.49
			FFX	2.44	0.00	3.33	86.67	1.42	-6548.27
	34-162	71.1	Original	4.39	2.90	3.15	83.46	2.25	-5572.81
			FFX	1.75	0.00	1.57	88.19	1.28	-5797.47
	35-158	78.7	Original	6.36	1.01	1.64	92.62	1.87	-5378.74
			FFX	1.82	0.00	0.82	96.72	0.90	-5637.72
	37-161	69.3	Original	1.79	0.00	0.00	95.93	0.97	-5513.05
			FFX	1.79	0.00	0.00	99.19	0.69	-5738.66
PCDH15	1140-1405	95.0	Original	0.00	0.23	0.00	95.08	0.93	-11351.62
			FFX	1.30	0.00	0.00	95.83	0.87	-11966.15
	1804-1836	56.0	Original	0.00	29.41	0.00	100.00	1.95	-824.47
			FFX	12.90	0.00	0.00	96.77	1.54	-863.96
	380-814	30.2	Original	0.26	4.81	1.62	92.61	1.71	-19590.68
			FFX	0.79	0.15	0.92	93.30	0.99	-20467.96
	398-812	93.5	Original	3.61	0.79	0.48	94.67	1.54	-19582.55
			FFX	1.39	0.00	0.00	93.70	1.03	-20371.44
PCDH15-Dimer	28-387	97.5	Original	0.62	0.62	0.00	98.32	0.70	-36577.40
			FFX	0.77	0.00	0.28	97.49	0.60	-37684.46

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	337-756	34.9	Original	0.96	5.37	3.11	89.00	1.86	-37557.11
			FFX	1.64	0.08	1.32	91.87	1.19	-39943.43
	400-918	27.1	Original	1.77	1.95	2.61	90.04	1.70	-44213.41
			FFX	1.66	0.00	1.45	91.97	1.16	-46932.94
	509-1034	22.0	Original	1.55	5.46	3.24	88.36	2.03	-43488.41
			FFX	1.99	0.00	1.62	92.94	1.18	-46428.23
	545-1026	30.8	Original	0.84	8.18	2.40	89.90	2.00	-39401.63
			FFX	1.57	0.00	1.25	90.62	1.18	-42360.67
	621-1136	26.3	Original	0.34	7.19	1.46	92.70	1.86	-42060.04
			FFX	0.89	0.00	1.36	92.12	0.98	-45169.74
	816-1357	95.0	Original	0.42	0.47	0.56	95.09	1.00	-45197.19
			FFX	1.16	0.12	0.28	94.07	0.99	-48093.70
PCDH15-Heteromer	623-1139	24.8	Original	2.42	9.20	4.52	86.03	2.42	-39595.14
			FFX	2.42	0.00	2.31	89.39	1.36	-42772.42
PDZD7	4-294	27.0	Original	3.32	10.52	3.46	89.27	2.51	-11934.29
			FFX	2.07	0.00	1.38	91.70	1.24	-12704.20
	76-170	94.0	Original	0.00	74.89	1.08	92.47	2.81	-3967.22
			FFX	0.00	0.00	1.08	95.70	0.80	-4269.21
	83-293	33.0	Original	1.14	7.81	3.83	86.12	2.11	-9072.93
			FFX	2.27	0.00	0.00	91.87	1.26	-9474.97
	84-292	28.3	Original	0.00	6.01	4.35	87.92	1.93	-8950.79
			FFX	0.00	0.00	0.97	89.86	1.05	-9429.77
	85-293	22.7	Original	2.87	10.11	5.31	83.57	2.56	-8877.05
			FFX	2.30	0.00	1.93	90.82	1.30	-9456.83
PEX1	13-179	89.0	Original	2.08	33.67	0.00	97.58	2.34	-6813.34
			FFX	1.39	0.00	0.61	93.94	1.01	-7102.25
	405-1063	27.5	Original	2.41	3.25	3.04	91.02	1.93	-27332.24
			FFX	2.59	0.10	0.46	94.37	1.24	-29076.06
PEX1-Hexamer(*used monomer)	405-1064	28.0	Original	2.75	0.76	4.26	86.02	1.72	-26998.71
			FFX	1.55	0.00	1.06	88.45	1.23	-28770.29
PEX6	202-975	29.3	Original	2.20	5.11	3.76	87.82	2.13	-30173.01
			FFX	2.51	0.00	2.07	90.80	1.33	-32131.23
	203-976	27.3	Original	1.72	15.06	3.76	89.90	2.41	-29687.16
			FFX	2.51	0.09	1.94	91.19	1.35	-32408.08

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	757-935	47.0	Original	1.99	111.88	5.65	86.44	3.37	-7146.98
			FFX	5.30	0.00	2.82	87.01	1.67	-7708.58
PEX6-Hexamer(*used monomer)	202-979	29.0	Original	2.50	0.85	3.87	86.86	1.69	-30409.87
			FFX	2.66	0.00	2.58	90.98	1.34	-32276.23
	435-930	19.0	Original	2.49	3.48	4.05	86.44	2.07	-18814.25
			FFX	1.25	0.00	2.63	89.68	1.13	-20024.72
PJVK	1-236	22.7	Original	0.98	12.08	2.14	89.32	2.16	-9790.25
			FFX	0.49	0.27	1.28	94.02	1.00	-10355.57
	27-95	30.0	Original	6.35	61.84	1.49	92.54	3.34	-3121.04
			FFX	3.17	0.00	1.49	85.07	1.54	-3311.64
PNPT1	45-754	37.0	Original	3.80	125.40	3.39	86.86	3.63	-28238.18
			FFX	3.47	0.00	2.82	88.98	1.49	-30994.57
	606-752	17.9	Original	1.57	3.44	2.07	92.41	1.76	-6502.94
			FFX	0.79	0.00	0.69	95.86	0.79	-6744.39
PNPT1-Heteromer	45-749	38.2	Original	1.91	4.05	1.43	93.20	1.84	-85144.13
			FFX	1.85	0.03	0.64	94.28	1.10	-89951.07
PNPT1-Trimer	43-744	34.3	Original	1.51	3.49	2.52	90.10	1.82	-88386.97
			FFX	1.17	0.09	1.57	92.29	1.07	-93356.95
	45-593	38.7	Original	1.36	2.27	2.74	90.49	1.64	-70064.50
			FFX	1.36	0.27	0.91	91.41	1.21	-73893.79
POLR1C	40-340	47.0	Original	1.88	84.54	0.33	97.32	2.73	-13874.72
			FFX	2.63	0.00	0.33	95.99	1.10	-14514.71
	51-332	34.7	Original	0.81	8.46	2.14	90.71	1.99	-12663.51
			FFX	1.21	0.00	2.14	92.86	1.02	-13196.05
POLR1D	20-120	50.0	Original	2.22	2.52	3.03	90.91	1.82	-4829.22
			FFX	1.11	0.00	1.01	90.91	1.06	-5027.74
POU3F4	187-338	63.2	Original	2.96	0.81	0.67	92.67	1.57	-6883.42
			FFX	0.00	0.00	0.00	98.00	0.50	-7156.86
	188-338	87.0	Original	3.73	177.48	2.01	94.63	3.51	-6419.51
			FFX	2.99	0.00	2.01	94.63	1.23	-7009.59
	190-331	41.9	Original	3.20	3.49	0.71	95.00	1.87	-6367.05
			FFX	1.60	0.00	0.00	97.86	0.69	-6623.25
	194-336	62.4	Original	5.56	4.27	2.13	92.91	2.23	-6359.23
			FFX	0.00	0.00	0.71	94.33	0.89	-6638.38
POU4F3	180-333	49.0	Original	0.00	3.58	3.29	90.79	1.67	-7176.66

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.76	0.00	0.66	95.39	0.82	-7467.15
	183-327	38.5	Original	3.28	2.56	2.80	89.51	1.99	-6600.18
			FFX	1.64	0.00	0.70	93.71	1.08	-6925.45
	183-331	38.5	Original	3.17	5.35	0.68	94.56	2.04	-6852.23
			FFX	0.79	0.00	0.68	96.60	0.72	-7176.18
	184-333	49.7	Original	3.15	2.85	1.35	95.95	1.73	-6889.21
			FFX	0.79	0.00	0.00	95.27	0.83	-7128.31
PRPS1	3-318	48.0	Original	4.48	127.69	3.50	90.76	3.59	-12861.62
			FFX	4.85	0.20	2.23	91.72	1.60	-13881.79
PTPRQ	2013-2328	39.4	Original	0.70	6.16	1.59	92.04	1.82	-13981.36
			FFX	1.41	0.00	0.64	93.63	1.03	-14648.04
	614-894	17.1	Original	0.77	8.24	2.15	87.10	2.07	-11639.20
			FFX	1.15	0.00	1.43	93.19	0.99	-12400.27
	711-988	19.1	Original	2.72	1.36	3.26	89.13	1.77	-11167.89
			FFX	2.33	0.00	1.81	91.30	1.29	-11814.13
	714-1091	18.1	Original	0.57	8.04	3.46	90.43	1.98	-14577.67
			FFX	0.86	0.00	1.33	91.76	1.00	-15614.36
	797-993	13.0	Original	3.33	107.52	3.08	86.67	3.52	-7577.96
			FFX	3.33	0.00	1.54	82.05	1.60	-8267.21
	948-1378	18.7	Original	2.53	11.98	2.10	86.71	2.53	-15894.05
			FFX	1.27	0.00	2.10	90.44	1.11	-17136.20
PTPRQ-Dimer	571-898	17.3	Original	3.62	14.17	5.06	88.65	2.67	-25571.82
			FFX	2.14	0.00	1.69	89.57	1.31	-28097.56
	664-992	18.3	Original	0.99	3.19	3.98	91.13	1.63	-26754.80
			FFX	1.32	0.00	1.68	92.66	1.05	-28679.84
RDX	3-583	59.1	Original	0.78	3.01	1.04	97.75	1.15	-31483.40
			FFX	0.58	0.00	0.17	97.58	0.59	-32700.66
	337-445	16.8	Original	0.00	1.10	3.74	91.59	1.32	-6936.13
			FFX	0.00	0.00	0.00	94.39	0.88	-7093.26
	497-583	71.3	Original	1.30	0.00	0.00	100.00	0.59	-4967.98
			FFX	0.00	0.00	0.00	100.00	0.50	-5082.64
RIPOR2 (FAM65B)	1011-1065	19.0	Original	4.26	113.61	1.89	94.34	3.38	-2734.91
			FFX	4.26	0.00	0.00	98.11	0.98	-2850.65
	832-1036	13.6	Original	1.09	7.69	2.46	92.61	1.91	-8316.03
			FFX	1.09	0.00	0.49	93.10	0.97	-8758.21
	883-1032	14.4	Original	0.75	8.36	2.70	91.22	1.97	-6129.71

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.49	0.42	0.68	93.92	1.19	-6490.87
ROR1	442-742	38.4	Original	0.75	8.53	5.69	86.29	2.10	-12934.43
			FFX	1.49	0.00	1.67	92.31	1.11	-13705.75
	464-749	70.0	Original	0.78	52.34	0.70	95.42	2.51	-12451.63
			FFX	0.39	0.00	0.00	94.01	0.90	-13090.59
	467-749	39.2	Original	1.19	3.75	1.42	93.95	1.63	-12334.67
			FFX	0.40	0.00	1.07	96.09	0.77	-12938.67
S1PR2	6-301	52.0	Original	1.97	90.68	0.68	94.56	3.02	-9461.44
			FFX	2.36	0.00	1.36	94.22	1.17	-10159.44
	7-302	51.9	Original	1.18	4.24	0.68	94.90	1.61	-9722.30
			FFX	0.79	0.00	0.34	97.62	0.58	-10288.58
SERPINB6	3-376	48.3	Original	2.45	3.39	0.00	98.39	1.43	-16873.97
			FFX	0.31	0.17	0.27	96.77	0.77	-17578.23
	347-376	55.0	Original	0.00	37.58	3.57	92.86	2.51	-1152.52
			FFX	3.70	0.00	0.00	96.43	1.17	-1229.01
SIX5	129-259	21.2	Original	0.93	2.28	5.43	86.05	1.65	-6175.02
			FFX	0.93	0.00	1.55	89.15	1.07	-6432.00
	222-255	56.0	Original	0.00	46.71	0.00	100.00	2.15	-1675.27
			FFX	0.00	0.00	0.00	100.00	0.50	-1752.58
	86-257	61.6	Original	2.14	1.06	1.18	95.88	1.34	-8061.04
			FFX	1.43	0.00	1.18	98.24	0.62	-8269.43
SLC17A8	78-520	10.0	Original	3.59	163.18	2.27	94.56	3.47	-13946.29
			FFX	3.87	0.00	1.59	90.70	1.48	-15416.53
SLC22A4	150-217	35.0	Original	1.72	101.40	4.55	89.39	3.22	-2027.87
			FFX	1.72	0.00	0.00	90.91	1.20	-2174.70
SLC26A4	516-562	33.9	Original	0.00	0.00	0.00	95.56	0.81	-1945.74
			FFX	2.38	0.00	0.00	97.78	0.84	-2021.96
	516-727	39.0	Original	5.18	110.72	2.38	90.95	3.57	-8510.95
			FFX	4.15	0.00	1.43	90.48	1.50	-9491.82
	655-731	34.6	Original	0.00	3.97	0.00	97.33	1.31	-3452.13
			FFX	0.00	0.00	0.00	98.67	0.50	-3562.96
SLC26A5	505-718	93.0	Original	2.78	40.18	0.00	97.64	2.50	-9148.25
			FFX	2.22	0.00	0.00	96.70	0.97	-9774.35
	638-718	92.7	Original	0.00	0.00	0.00	100.00	0.50	-3457.98
			FFX	0.00	0.00	0.00	94.94	0.85	-3571.74
SLITRK6	31-267	53.0	Original	4.19	70.55	0.85	92.34	3.27	-9795.96
			FFX	3.72	0.00	0.43	93.19	1.37	-10395.82

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	330-567	46.9	Original	1.34	2.09	0.85	91.95	1.56	-10003.81
			FFX	2.23	0.00	2.12	93.22	1.20	-10453.29
SMPX	6-75	31.0	Original	3.45	66.48	1.47	97.06	2.87	-2975.18
			FFX	3.45	0.00	0.00	92.65	1.37	-3157.80
	13-42	33.3	Original	4.35	4.42	21.43	64.29	2.59	-1232.88
			FFX	0.00	0.00	3.57	78.57	1.25	-1313.04
SNAI2	124-263	36.5	Original	3.17	1.79	0.00	92.75	1.78	-7367.52
			FFX	0.00	0.00	0.72	91.30	1.01	-7643.96
	125-263	29.6	Original	4.00	3.16	2.92	88.32	2.16	-6887.70
			FFX	0.80	0.00	2.92	92.70	0.96	-7244.36
	126-263	34.1	Original	7.20	4.09	6.62	81.62	2.56	-6107.16
			FFX	1.60	0.00	2.21	92.65	1.12	-6492.47
	129-262	42.8	Original	0.82	4.66	2.27	92.42	1.71	-7381.13
			FFX	0.00	0.00	3.79	88.64	1.08	-7646.15
	158-264	87.0	Original	8.25	47.51	0.95	95.24	3.18	-4708.72
			FFX	3.09	0.00	0.95	93.33	1.31	-5007.63
SOX10	102-173	96.0	Original	3.12	34.04	0.00	98.57	2.39	-3771.96
			FFX	3.12	0.00	0.00	100.00	0.88	-3907.17
STRC	859-1074	31.0	Original	4.35	176.85	5.14	82.71	3.89	-7392.04
			FFX	4.35	0.30	3.74	85.05	1.75	-8165.92
SYNE4	100-201	8.0	Original	1.28	0.62	3.00	94.00	1.19	-4392.20
			FFX	1.28	0.00	1.00	96.00	0.86	-4539.33
	122-249	10.4	Original	1.00	3.57	1.59	92.06	1.63	-5385.38
			FFX	1.00	0.00	2.38	92.86	0.95	-5588.68
	24-202	11.4	Original	1.44	7.08	1.13	92.66	1.97	-7712.27
			FFX	0.72	0.00	0.56	96.05	0.77	-8075.28
	25-195	13.7	Original	0.76	3.54	4.14	86.98	1.77	-7379.36
			FFX	1.52	0.00	1.18	93.49	1.06	-7692.95
	41-253	21.0	Original	4.76	124.45	1.42	94.31	3.46	-8804.74
			FFX	4.17	0.00	0.47	92.89	1.42	-9519.97
TBC1D24	124-292	15.0	Original	0.68	6.24	1.80	91.62	1.84	-6910.31
			FFX	2.05	0.00	0.60	96.41	0.98	-7245.65
	27-310	26.9	Original	1.22	8.77	0.35	93.97	1.94	-11757.11
			FFX	0.81	0.00	0.00	95.74	0.80	-12335.34
	339-554	33.7	Original	0.55	6.28	5.14	84.58	2.01	-8822.86
			FFX	1.09	0.00	0.00	91.59	1.03	-9293.55
	442-522	38.0	Original	1.56	71.07	1.27	94.94	2.82	-3175.68

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.56	0.00	0.00	93.67	1.07	-3347.27
TBX1	37-79	53.0	Original	0.00	39.86	0.00	95.12	2.42	-1393.20
			FFX	7.14	0.00	2.44	87.80	1.75	-1457.50
TCOF1	302-465	20.0	Original	4.59	49.77	0.00	98.77	2.68	-7143.75
			FFX	2.75	0.00	1.23	96.30	1.08	-7581.41
TECTA	1801-2066	17.8	Original	0.84	9.40	1.52	96.21	1.75	-11250.03
			FFX	0.42	0.00	0.38	94.70	0.86	-11843.56
	584-691	27.6	Original	6.32	0.66	1.89	91.51	1.83	-4183.96
			FFX	0.00	0.00	2.83	95.28	0.83	-4373.30
	973-1078	28.9	Original	3.09	2.60	3.85	87.50	2.03	-4550.77
			FFX	1.03	0.00	0.96	93.27	0.95	-4786.39
	993-1100	19.0	Original	5.26	2.61	1.89	94.34	1.98	-4792.71
			FFX	1.05	0.65	0.94	95.28	1.06	-4991.64
TECTA-Heteromer	1772-2085	36.8	Original	0.37	4.22	1.33	94.52	1.58	-26318.22
			FFX	0.55	0.00	0.17	94.68	0.87	-27657.23
TECTB	188-280	37.0	Original	4.71	85.75	5.49	87.91	3.51	-3739.66
			FFX	1.18	0.00	2.20	84.62	1.22	-4026.46
	27-286	17.5	Original	0.86	7.05	3.10	88.37	1.98	-10322.39
			FFX	3.00	0.00	0.78	90.70	1.39	-10944.05
TECTB-Dimer	27-287	22.5	Original	0.64	6.42	1.54	91.12	1.87	-20632.12
			FFX	2.35	0.00	0.97	92.66	1.24	-22097.80
TIMM8A	27-84	41.9	Original	1.82	0.00	0.00	94.64	1.07	-2827.67
			FFX	0.00	0.00	0.00	100.00	0.50	-2916.38
TJP2	23-132	93.0	Original	2.20	62.02	0.93	93.52	2.95	-4665.02
			FFX	1.10	0.00	0.93	93.52	0.96	-4961.45
	24-127	97.1	Original	2.30	1.25	3.92	84.31	1.79	-4549.09
			FFX	0.00	0.00	0.98	93.14	0.94	-4803.10
	511-887	66.6	Original	1.88	1.34	2.13	93.87	1.48	-17177.87
			FFX	0.62	0.00	0.53	94.93	0.85	-17936.64
TMC1	68-137	36.0	Original	4.92	67.17	2.94	94.12	3.22	-4278.30
			FFX	3.28	0.00	0.00	92.65	1.35	-4505.95
TMEM132E	337-460	17.9	Original	1.82	3.63	3.28	86.07	1.99	-4880.51
			FFX	0.00	0.00	1.64	91.80	0.99	-5098.02
	532-706	29.0	Original	2.68	102.30	1.73	93.64	3.22	-6238.41
			FFX	0.67	0.00	0.00	93.64	0.92	-6809.73
	542-675	21.4	Original	0.87	3.97	2.27	93.18	1.62	-4704.34

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.00	0.00	2.27	91.67	1.00	-4920.92
TMIE	1-48	48.0	Original	0.00	34.74	0.00	95.65	2.33	-1477.67
			FFX	3.03	0.00	0.00	97.83	0.91	-1560.68
TMPRSS3	105-451	39.9	Original	1.03	4.58	1.16	91.01	1.76	-13588.80
			FFX	1.03	0.19	0.87	93.62	1.01	-14260.37
	242-446	50.0	Original	3.53	108.62	5.91	85.22	3.57	-7273.74
			FFX	3.53	0.00	2.96	85.22	1.57	-7949.33
TNC	1043-1340	24.5	Original	2.44	6.18	1.35	92.91	2.09	-12396.62
			FFX	3.66	0.00	1.01	94.26	1.32	-12942.89
	1132-1420	22.9	Original	2.10	11.72	1.74	90.59	2.36	-11402.42
			FFX	2.94	0.00	0.35	91.99	1.34	-12101.86
	1223-1511	23.4	Original	0.81	4.55	0.70	90.59	1.76	-11786.17
			FFX	1.63	0.00	0.70	89.55	1.22	-12298.44
	1254-1603	27.1	Original	3.02	1.50	3.45	89.08	1.83	-13972.21
			FFX	2.01	0.00	1.72	92.53	1.20	-14704.18
	1545-1885	28.8	Original	1.03	4.97	2.06	93.81	1.68	-13941.24
			FFX	1.03	0.00	0.88	93.22	0.95	-14575.78
	1580-1878	15.5	Original	1.18	5.23	3.03	88.89	1.91	-12200.97
			FFX	0.00	0.00	1.01	92.26	0.98	-12768.28
	1585-1886	16.7	Original	0.78	11.04	2.33	92.67	2.02	-12162.46
			FFX	1.17	0.22	1.33	93.00	1.08	-12897.68
	1620-1968	24.6	Original	0.66	5.63	2.02	92.80	1.76	-14168.54
			FFX	1.00	0.00	0.29	93.95	0.90	-14873.07
	1620-1975	24.2	Original	0.98	6.99	3.67	90.96	1.91	-14342.97
			FFX	1.95	0.00	1.69	92.09	1.20	-15032.24
	1625-1797	49.0	Original	3.33	81.93	0.58	97.08	2.94	-6989.31
			FFX	2.67	0.00	0.00	95.91	1.11	-7495.39
	175-374	39.2	Original	1.20	1.48	1.01	89.90	1.50	-9304.26
			FFX	0.00	0.00	0.00	89.90	1.05	-9766.58
	1980-2191	39.6	Original	0.57	5.15	0.95	94.29	1.66	-10212.48
			FFX	0.57	0.00	0.00	95.24	0.83	-10599.75
	206-425	38.1	Original	1.10	1.68	0.92	91.28	1.46	-10407.75
			FFX	1.66	0.00	0.00	91.28	1.18	-10818.11
	237-498	38.1	Original	0.46	5.59	3.85	86.15	1.94	-12678.03
			FFX	0.46	0.00	0.38	87.31	1.11	-13290.20
	297-560	33.2	Original	0.00	3.63	2.67	85.11	1.81	-12970.87
			FFX	0.45	0.00	0.38	91.60	1.00	-13576.63

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	300-529	40.3	Original	1.05	7.35	0.44	90.79	1.95	-11412.66
			FFX	0.00	0.00	0.88	92.11	0.98	-11821.36
	623-974	29.3	Original	0.32	7.06	1.14	92.00	1.88	-16127.06
			FFX	0.32	0.00	0.29	94.29	0.89	-16856.75
	623-983	29.5	Original	0.63	22.15	0.84	92.20	2.32	-16255.07
			FFX	1.58	0.00	0.56	94.43	1.03	-17087.80
	625-893	13.6	Original	0.41	19.47	3.37	86.52	2.42	-11983.81
			FFX	1.65	0.00	1.87	91.76	1.16	-12615.37
	712-1063	25.7	Original	0.65	7.42	1.43	92.29	1.88	-16105.37
			FFX	1.95	0.00	0.86	93.43	1.15	-16762.76
	712-1071	23.5	Original	1.27	10.27	1.12	91.62	2.11	-16260.65
			FFX	0.96	0.00	0.84	93.30	0.94	-17086.75
	766-984	28.5	Original	3.21	11.92	4.15	85.71	2.62	-9507.50
			FFX	2.14	0.00	4.15	90.78	1.28	-10175.25
	854-1147	29.3	Original	1.21	4.40	2.05	93.15	1.72	-12802.49
			FFX	1.21	0.00	1.03	95.89	0.85	-13386.34
	854-1162	25.3	Original	0.77	5.87	2.28	90.55	1.85	-13259.80
			FFX	1.54	0.00	0.98	92.83	1.10	-13919.22
	893-1235	26.5	Original	0.00	1.52	0.88	91.50	1.40	-14380.96
			FFX	1.42	0.00	0.88	92.38	1.09	-15036.88
TPRN	159-189	68.0	Original	0.00	35.26	0.00	96.55	2.25	-873.77
			FFX	9.52	0.00	0.00	93.10	1.69	-913.37
TRIOBP	1807-1887	42.0	Original	1.39	83.96	1.27	89.87	3.05	-3867.36
			FFX	2.78	0.00	1.27	93.67	1.26	-4083.53
	2065-2251	17.0	Original	0.00	2.99	1.62	94.05	1.49	-9877.05
			FFX	0.62	0.00	1.08	96.76	0.70	-10281.07
	2072-2355	13.1	Original	0.00	4.13	1.42	91.84	1.69	-14628.96
			FFX	1.20	0.00	1.42	91.84	1.05	-15173.87
	2113-2332	12.1	Original	2.03	0.56	1.83	93.12	1.37	-11556.28
			FFX	1.02	0.00	0.00	93.58	0.93	-12051.77
	2115-2255	13.6	Original	0.80	2.17	0.00	97.12	1.15	-7853.07
			FFX	0.00	0.00	0.72	97.84	0.54	-8058.03
	2115-2262	15.0	Original	0.00	4.96	0.68	95.21	1.59	-8036.34
			FFX	0.00	0.00	0.00	98.63	0.50	-8299.32
	2115-2263	18.5	Original	0.76	1.23	0.68	95.24	1.17	-8075.60
			FFX	0.76	0.00	0.00	97.28	0.64	-8356.36
	2115-2321	13.4	Original	0.00	3.86	0.49	96.59	1.39	-10939.76

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	0.00	0.00	0.00	96.59	0.72	-11368.16
	2117-2335	22.6	Original	0.00	3.36	2.76	91.24	1.64	-11507.78
			FFX	0.00	0.00	0.92	91.71	1.00	-11940.14
	2127-2345	12.1	Original	2.54	6.11	2.30	92.17	2.12	-11779.01
			FFX	0.00	0.00	0.46	95.85	0.79	-12283.39
	2129-2349	16.1	Original	1.01	1.37	3.20	87.21	1.49	-11590.46
			FFX	0.51	0.00	1.37	92.69	0.96	-12131.63
	2130-2287	8.2	Original	0.00	7.07	1.28	94.23	1.78	-8328.24
			FFX	0.00	0.00	0.00	96.79	0.70	-8722.64
	2130-2324	18.9	Original	0.00	1.57	2.59	92.75	1.36	-10347.82
			FFX	0.58	0.00	0.00	93.78	0.91	-10727.58
	2137-2335	11.1	Original	2.26	4.32	3.55	90.36	2.02	-10214.41
			FFX	1.13	0.00	1.52	93.91	0.95	-10701.98
	2154-2317	11.9	Original	0.00	7.53	1.85	93.83	1.82	-8474.53
			FFX	0.69	0.00	0.62	96.91	0.68	-8854.40
	2166-2311	17.3	Original	0.00	0.84	0.00	94.44	1.14	-7682.42
			FFX	0.00	0.00	0.69	97.22	0.64	-7998.45
	2188-2350	13.6	Original	2.01	11.86	3.73	90.06	2.37	-8608.91
			FFX	0.67	0.00	0.62	94.41	0.88	-9076.94
	2190-2360	18.2	Original	0.65	7.47	1.78	89.94	1.96	-8924.89
			FFX	1.29	0.00	0.59	95.27	0.91	-9368.68
	2200-2355	10.0	Original	0.71	4.28	2.60	92.21	1.69	-8213.13
			FFX	0.00	0.00	0.00	94.16	0.89	-8548.80
TSPEAR	106-160	42.0	Original	4.26	92.45	1.89	92.45	3.38	-2148.61
			FFX	2.13	0.00	0.00	90.57	1.28	-2283.72
TWNK	174-631	19.0	Original	0.50	14.05	2.85	88.60	2.24	-19079.66
			FFX	1.00	0.00	0.44	90.57	1.03	-20395.36
	404-453	36.0	Original	0.00	31.83	0.00	97.92	2.01	-2036.69
			FFX	0.00	0.00	0.00	97.92	0.52	-2141.70
	57-359	17.9	Original	2.33	15.47	3.32	89.04	2.55	-12491.61
			FFX	2.71	0.00	0.00	93.36	1.26	-13352.53
USH1C	108-301	22.1	Original	2.42	9.86	3.12	90.10	2.35	-7836.33
			FFX	0.61	0.00	1.04	92.71	0.96	-8230.81
	300-383	18.0	Original	1.30	32.06	0.00	100.00	2.08	-5397.15
			FFX	2.60	0.00	0.00	100.00	0.81	-5648.61
	82-291	35.3	Original	1.68	11.24	6.25	85.58	2.39	-8556.78
			FFX	1.12	0.00	1.92	90.87	1.06	-8985.90

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
	83-281	33.3	Original	0.59	8.88	3.55	85.79	2.12	-8280.33
			FFX	1.18	0.00	3.05	90.86	1.08	-8686.03
	86-281	34.0	Original	6.59	7.06	4.64	80.41	2.74	-7865.19
			FFX	1.80	0.00	2.06	90.21	1.24	-8398.80
USH1G	1-157	31.4	Original	0.00	9.29	1.94	94.19	1.89	-7572.73
			FFX	0.76	0.00	0.00	96.13	0.76	-7813.57
	1-166	24.2	Original	2.14	2.63	0.61	95.12	1.64	-8160.69
			FFX	0.71	0.00	0.00	96.95	0.68	-8479.82
	2-161	36.5	Original	0.00	3.95	2.53	93.04	1.63	-7839.43
			FFX	0.00	0.00	0.63	96.20	0.76	-8128.72
	388-461	99.0	Original	0.00	31.75	0.00	95.83	2.27	-3532.39
			FFX	1.54	0.00	0.00	98.61	0.64	-3630.99
USH2A	1509-1893	24.9	Original	1.85	13.01	3.13	89.03	2.40	-15049.74
			FFX	1.85	0.00	1.04	90.60	1.24	-16074.76
	2325-2921	17.1	Original	3.35	9.92	4.71	86.89	2.54	-20997.61
			FFX	3.35	0.22	3.36	88.24	1.58	-22643.89
	2326-2921	19.1	Original	2.24	6.81	3.87	87.21	2.26	-20939.33
			FFX	3.36	0.00	3.03	88.72	1.48	-22746.31
	2327-2816	16.7	Original	2.06	10.47	5.33	86.48	2.41	-16756.46
			FFX	1.38	0.13	2.66	88.11	1.26	-18417.38
	2409-2823	16.9	Original	2.16	9.44	4.36	88.86	2.33	-14460.91
			FFX	1.35	0.16	1.94	90.07	1.21	-15647.65
	287-715	31.7	Original	2.36	8.65	2.58	89.23	2.32	-18464.57
			FFX	2.36	0.00	1.87	91.80	1.28	-19742.57
	287-728	35.0	Original	0.76	6.29	1.36	87.27	1.96	-19133.75
			FFX	1.02	0.00	0.23	91.59	1.01	-20375.77
	326-728	38.5	Original	1.39	9.38	2.74	86.28	2.24	-17655.70
			FFX	1.39	0.16	1.50	92.77	1.13	-18658.26
	950-991	43.0	Original	8.57	108.62	2.50	87.50	3.82	-1696.90
			FFX	0.00	0.00	0.00	87.50	1.11	-1814.62
WFS1	759-835	43.0	Original	0.00	36.10	1.33	94.67	2.41	-3422.21
			FFX	0.00	0.00	1.33	89.33	1.07	-3610.99
	82-188	14.7	Original	0.00	5.27	1.90	91.43	1.79	-5440.82
			FFX	2.25	0.00	0.95	93.33	1.20	-5633.96
	83-186	17.2	Original	1.15	1.80	2.94	88.24	1.58	-5226.81
			FFX	1.15	0.00	0.98	95.10	0.89	-5479.67
	86-188	15.6	Original	1.16	3.62	0.00	87.13	1.82	-5160.64

Gene	Residue Range	Sequence Identity	Model	% Poor Rotamers	Clash Score	Ramachandran Outliers	Ramachandran Favored	MolProbity Score	Amoeba Energy (Kcal/Mol)
			FFX	1.16	0.00	0.99	95.05	0.89	-5340.43
WHRN	260-378	97.0	Original	2.06	55.43	2.56	92.31	2.93	-5267.44
			FFX	1.03	0.00	0.85	94.02	0.91	-5615.43
Averages		41.74	Original	2.33	25.00	2.03	91.94	2.16	-15342.08
			FFX	1.60	0.03	0.93	93.48	1.04	-16287.37