Rosetta Energy Analysis of AlphaFold2 models: Point Mutations and Conformational Ensembles

Richard A. Stein and Hassane S. Mchaourab

Department of Molecular Physiology and Biophysics and Center for Applied AI in Protein Dynamics
Vanderbilt University

Richard Stein
741B Light Hall
2215 Garland Ave.
Dept. of Mol. Physiol. and Biophys.
Vanderbilt University
Nashville, TN 37232

Hassane Mchaourab
747 Light Hall
2215 Garland Ave.
Dept. of Mol. Physiol. and Biophys.
Vanderbilt University
Nashville, TN 37232

Running Title: Protein Energetic Landscapes with AlphaFold2 and Rosetta

Keywords: AlphaFold, Ensemble, Conformation, Energy Landscape

Statement: AlphaFold2's ability to accurately predict protein structures from a multiple sequence alignment has raised many questions about the utility of the models generated in downstream structural analysis. Two outstanding questions are the prediction of the consequences of point mutations and the completeness of the landscape of protein conformational ensembles. The results presented support that AlphaFold2 can predict structural changes via point mutation and has learned information about protein structural energetics encoded in the multiple sequence alignment.

Author's Contributions: RAS: Conceptualization, Analysis, and Writing. HSM: Conceptualization, Funding Acquistion and Writing.

ABSTRACT

AlphaFold2's ability to accurately predict protein structures from a multiple sequence alignment (MSA) has raised many questions about the utility of the models generated in downstream structural analysis. Two outstanding questions are the prediction of the consequences of point mutations and the completeness of the landscape of protein conformational ensembles. We previously developed a method, SPEACH_AF, to obtain alternate conformations by introducing residue substitutions across the MSA and not just within the input sequence. Here, we compared the structural and energetic consequences of having the mutation(s) in the input sequence versus in the whole MSA (SPEACH_AF). Both methods yielded models different from the wild-type sequence, with more robust changes when the mutation(s) were in the whole MSA. To evaluate models of conformational diversity, we used SPEACH_AF and a new MSA subsampling method, AF_cluster, combined with model relaxation in Rosetta. We find that the energetics of the conformations generated by AlphaFold2 correspond to those seen in experimental crystal structures and explored by standard molecular dynamic methods. Combined, the results support the fact that AlphaFold2 can predict structural changes due to point mutations and has learned information about protein structural energetics that are encoded in the MSA.

INTRODUCTION

The results from CASP 14, the competition for de novo structure prediction, revolutionized the field of structural biology with models produced by AlphaFold2 at experimental accuracy.[1] The subsequent publication of the methodology[2] and release of pipelines utilizing the methodology intensified the interest in AlphaFold2 and its use in structural biology. At the same time, the DeepMind team in conjunction with EMBL-EBI released putative structures of the human proteome[3] along with over 45 proteomes of other species. Initial discussions centered around the generation of models adopting only a single conformation, epitomized by the presentation of a single model per sequence in the AlphaFold2 database. The subsequent ability to run the AlphaFold2 pipeline using a convenient interface, ColabFold,[4] expanded the ability to understand this tool and to gauge its ability to address a variety of protein structure/function questions.

One of the major questions was whether AlphaFold2 could predict the effects of single point mutations. An early short letter concluded that AlphaFold2 could not predict the effect of point mutations. The models generated by AlphaFold2 were folded and similar to the wild-type models whereas experimentally the mutations induced unfolding.[5] In contrast, several studies have generated AlphaFold2 models of wild-type proteins to extract sequence and physical properties and create predictors of mutational effects that do appear to have predictive value.[6–8] A recently published report found that the internal AlphaFold2 quality metric, pLDDT, was not an adequate measure for the energetic changes due to point mutations.[9] Conversely, another study found a correlation between pLDDT and pathogenicity in a set of cancer proteins.[10] In addition, a neural network was able to learn the pathogenicity of BRCA1 mutations from an AlphaFold2 model.[11] Unfortunately this network was BRCA1 specific, limiting the methodology to targets with already characterized data. While these studies are promising, the dogma remains that AlphaFold2 is not able to predict the effect of point mutations.[12–15]

Another application of interest is the generation and understanding of protein conformational ensembles. The initial perception was that AlphaFold2 only generates a single conformation even

though there are five independently trained modelers. We have shown that for some proteins it is possible to generate multiple conformations via multiple runs with different random seeds.[16] We further tailored the pipeline by modifying the MSA, either making changes at specific residues or scanning across the protein sequence. These *in silico* mutations were made across the whole MSA as modifications of only the input sequence did not yield robust results. The methodology has been applied to the Major Facilitator Superfamily (MFS) membrane transporters to generate both inward- and outward-facing models.[17] Other methodologies have also been successful in generating alternate conformations with AlphaFold2, for example using structural templates from homologous sequences[18,19] or subsampling the MSA.[20] In particular, MSA subsampling modifies two input parameters, max_msa_clusters and max_extra_msa, to randomly select small subsets of the larger MSA.[20] A recent preprint has taken subsampling a step further by clustering the sequences in the MSA leading to multiple smaller input MSAs.[21] Unlike the unmodified AlphaFold2 algorithm,[22] application of this subsampling approach promotes conformational sampling of fold-switch proteins. Despite these methodological developments, there is still a perception that AlphaFold2 can only generate a single conformation.[23,24]

Here we combine methodologies for modeling point mutations and conformational ensembles with energetic analysis using Rosetta. We extend our observation that a more robust change in the modeling occurs if the mutations are placed in the whole MSA compared to just the input sequence. In contrast to previous reports, we find that AlphaFold2 is capable of predicting structural changes due to point mutations. However, we demonstrate that an integrated structural/energetic parameterization is necessary. We then expand our previous work on the exploration of protein conformational space for the models generated from AlphaFold2, by relaxing and scoring the energetics of the models with Rosetta. These models map out energy landscapes that track with previously reported conformations and energetics. Combined, the results support that AlphaFold2 can predict structural changes via point mutation and has learned information about protein structural energetics encoded in the MSA.

METHODS

Point Mutations. For all samples an MSA was generated with MMSeqs2,[25] via ColabFold.[4] The MSA was modified in two ways: the mutation was placed in the input sequence only or across the MSA replacing the amino acid in the same position in the sequence ignoring gaps.[16] The parent MSA and the two modified MSAs were then used as input for ColabFold[4] ten times to generate 50 models for each MSA. These models were then compared to either the same domain of the protein from the AlphaFold2 structure database or the appropriate experimental structures using TM-Align.[26] These models were also minimized and scored with Rosetta using FastRelax with the backbone constrained.[27–29] The outlier scores were removed if the values exceeded 4 times the standard deviation of the mean. A two-tailed t-test was carried out to determine whether the TM-scores or Rosetta Energy Scores for the sets of mutant models is different from the set of wild-type models.

Conformational Sampling. The MSAs used were generated with MMSeqs2,[25] via ColabFold.[4] The models generated with SPEACH_AF were those used previously.[16] For MSA subsampling, AF_cluster[21] was used with the minimum number of sequences per cluster (min_samples) set to 3, 7, or 11. The subsampled MSAs were then used to generate protein models using ColabFold.[4] The resultant models from both methods were processed to exclude models with a pLDDT less than 70, where a pLDDT of 70 or greater corresponds to a generally correct backbone prediction.[3] This parsed set of models were further

evaluated with principle component analysis (PCA) using ProDy[30] to remove models that have a high pLDDT but are misfolded or misthreaded. The remaining models were then subjected to minimization in Rosetta utilizing FastRelax with the backbone constrainted.[27–29] For the SPEACH_AF models, the residues mutated to alanine were mutated back to the native residues prior to relaxation. The default Rosetta score function was used for soluble proteins and the membrane_highres_Menv_smooth weights were used for membrane proteins. The membrane spanning regions were determined with TOPCONS.[31] The Rosetta relaxed structures were subsequently analyzed with the eigenvalues and eigenvectors from the PCA run on the models prior to relaxation. The Rosetta Energy Scores were adjusted by setting the lowest score to zero yielding a ΔG.

The collective variables for adenylate kinase and ribose binding protein were measured with MDAnalysis.[32,33] For adenylate kinase the collective variables are the AMP binding domain and ATP binding domain angles relative to the core domain. The core domain is comprised of residues 1-29, 60-121, and 160-214, the AMP binding domain (Hinge) is residues 30-59, and the ATP binding domain (Lid) is residues 122-159. The collective variables for ribose binding protein define the tilt and twist of the two domains relative to each other.[34,35] The N-terminal domain is comprised of residues 1-100 and 236-259 and the C-terminal domain is comprised of residues 108-231 and 269-271. The tilt angle is defined as the angle between the center of mass of the N- and C-terminal domains and the center of mass of the hinge point comprising residues 101-107, 232-235, and 160-268. The twist angle is the dihedral angle of the center of mass of the N- and C-terminal domains and two regions near the ribose binding site on the top of the N-terminal domain, 124-125, 262-262, and 283-284 and the bottom of the C-terminal domain, 133-134, 255-256, 294-295.

RESULTS AND DISCUSSION

AlaphFold2 prediction of point mutations. Given the incongruent results in modeling point mutations with AlpahFold2, we selected a set of fives proteins including three used in the Buel and Walters report[5] to support the contention that AlphaFold2 cannot predict the effect of point mutations. To model amino acid substitutions, we implemented two approaches. In the first approach, the mutation was included in the input sequence only. In the second approach, the MSA was edited to incorporate the substitution across all the sequences. The protein set consisted of: the BRCT domain of BRCA1, the ubiquitin-associated domain (UBA) of Rad23 (HR23A), the MyUb domain of Myosin VI, isocyanide hydratase, and XylE. The mutation A1708E in BRCA1 has been associated with breast cancer[36] and leads to an increase in proteolytic degradation of BRCA1.[37] In Rad23, L198A causes disorder of the UBA domain.[38] The Myosin VI mutation R1117A disrupts the structure of the MyUb domain.[39] Isocyanide hydratase and its point mutant D183A, were targets in CASP 15 (T1109/T1110). The mutation induces a small change in the helix containing the mutation and a switch in the dimeric structure. XylE, an MFS transporter, was selected because two point mutations lead to a shift in its structure.[40,41] Hydrogen/deuterium exchange experiments have shown that only one of the mutations is needed to induce the conformational change.[42]

Modeling of isocyanide hydratase by AlphaFold2 yielded TM scores consistent with accurate prediction of its wild-type structure (Fig. 1A). Both methods of modeling point mutations, i.e. in the input sequence and across the MSA, yielded two populations of models. One set of models is similar to the wild type structure (Fig. 1A). Close inspection of these models shows a perturbation around the mutation site, which is reflected by the slightly reduced TM score to the wild-type target (T1110) and a

slight increase in the TM-score relative to the mutant target (T1109). The second set of models appears to be a good match for the mutant protein with high TM scores to the mutant target (T1109) (Fig. 1A and S1B-C). Interestingly, only one of the modelers predicted the complete conformational change, but it is not the same modeler in both cases.

XylE has been crystallized in inward- and outward-facing conformations (Fig. S2A).[40,41] The outward-facing conformation was stabilized by two point mutations G58A/L315W. Examination of the models for XylE indicate that there is very little change in TM scores when the mutation is only in the input sequence (Fig. 1B). On the other hand, placement of the mutation(s) across the whole MSA leads to larger changes in TM scores. Both the single, G58W, and double mutant, G58W/L315W, lead to a complete shift in conformation evidenced by the high TM score when comparing to the alternate conformation (Fig. S2B-C). The G58W mutation also promotes the conformational change as reported by hydrogen/deuterium exchange experiments.[42]

In light of these results, we considered the possibility that the effects of the mutations in the previous report[5] were not observed because the mutation was not placed across the whole MSA. Therefore, we tested out the two methodologies on the protein domains used in that report. For this purpose, and in keeping with the approach of Buel and Walters,[5] AlphaFold2 models of mutants were compared to the corresponding domain from the model of the full-length protein from the AlphaFold2 database (Fig. S3). We found that the wild-type BRCT domain of BRAC1 is highly comparable to the domain from the full-length model (Fig. 2A). Moreover, there is a decrease in the TM-score for the mutant protein regardless of how the mutation is introduced to AlphaFold2. Yet there is a larger change in structure corresponding to a greater average shift in TM-score when the mutation is introduced across the whole MSA. The extent of change of TM-scores between each population of mutant models relative to the wild-type is statistically significant in both cases.

The TM-scores for the wild-type UBA and MyUb domains modeled in isolation are lower and more diverse than for the wild-type BRCT domain, suggesting that the context of these domains in the full-length protein is important for AlphaFold2 modeling of their structures (Fig. 2B-C). For the UBA domain, there is a smaller change in the TM-score for the models with the mutation in the input sequence compared to a more significant shift with the mutation across the whole MSA (Fig. 2B). The TM-scores for the mutant MyUb domain are slightly shifted up compared to the scores for wild-type models, though the effects are different depending on the placement of the mutation in the input sequence only or across the whole MSA (Fig. 2C). This shift is significant only when the mutation is in the whole MSA.

Energetics of points mutations modeled by AlphaFold2. These disparate results bring into the forefront the more general question of how mutations affect protein structure. The wealth of experimental studies highlights that while these tend to be protein specific, they can be grouped into two major classes. In one class, the mutations do not lead to major distortion in the native structure but mostly are associated with changes in the free energy of unfolding. If the change is large enough, the protein does not fold, misfolds, or is unstable and unfolds. In the second class, the mutations lead to notable distortions in the native structure. Thus, at the origin of the different conclusions between our results and those reported earlier is the expectation that AlphaFold2 should not generate a folded structure since experimentally the proteins were reported to be disordered or unfolded.

A previous study, examining the role of mutations in protein stability, reported that the Rosetta Energy Score (RES) of the mutant protein compared to wild type protein ($\Delta\Delta G$) correlated with the

experimental changes in the melting temperature of the protein.[43] Therefore, we examined whether the RES could provide insight into the mutations for the three proteins in the Buel and Walters report.[5] Shown in Figure S4 are the RESs for the BRCT, UBA, and MyUb domains. The change in average RES for the BRCT domain of BRCA1 is shifted 10 units higher for both methods of inputting the mutation. The scale of change in energy could suggest that this mutation is destabilizing, potentially driving unfolding. In contrast, the mutants of the UBA domain of HR23A lead to a small increase in the average value, while the change for the MyUb domain of MyoVI is marginal. In both cases the change in RES is slightly larger for the models predicted from the mutation being in only the input sequence. This appears to be contrary to the results for the TM-score, where there was a slightly larger effect when the mutation was in the whole MSA. It is possible that these results reflect that the structures generated by AlphaFold2 with the mutation in the MSA are more structurally perturbed, which would suggest that AlphaFold2 attempted to compensate for the mutation, in effect minimizing the energetic consequences of the mutation. Without this minimization, the change in free energy, as calculated by Rosetta, is increased in the presence of the mutation in the more wild-type structure. These results would suggest then that AlphaFold2 is capable of predicting changes due to mutations, though the structural or energetic consequences of that mutation might not be accurately predicted, especially in the case of mutations that lead to folding defects. Overall, our analysis supports the conclusion that AlphaFold2 has learned energetics of the folded protein that is encoded in the MSA.[44] Therefore, the parameterization of the change energetically and/or structurally is critical when testing the ability of AlphaFold2 to predict the effect of mutations.

Energetics of conformational sampling by AlphaFold2. Multiple methods have been advanced to generate conformational ensembles from AlphaFold2 including MSA subsampling methods and the *in silico* mutagenesis method SPEACH_AF.[16,20,21] Here we compare the structural diversity and energetics of models generated by the MSA subsampling technique AF_cluster[21] to the same set of proteins previously examined as being amenable to conformational sampling with SPEACH_AF.[16] Following generation of models by both methods, a two-step filtering method to parse out misfolded and misthreaded structures was carried out. The first step removed models with a pLDDT less than 70, where a pLDDT of 70 or greater corresponds to a generally correct backbone prediction.[3] The second step entails principal component analysis (PCA) carried out with models generated from both AF_cluster and SPEACH_AF to identify outliers that are misfolded or misthreaded. Initial tests with AF_cluster, based on pLDDT, yielded very few well-folded models when the minimum number of sequences per cluster was set to 3. Therefore, to examine the effects of the number of sequences in the subsampled MSA, additional subsampling was done with the minimum set to 7 and 11 sequences. For almost all of the cases, increasing the minimum number of sequences improved the overall quality of the models, but generally reduced the total number of models (Table 1). For RBP and 3 out of 4 GPCRs, CGRPR, FZD7, and PTH1R, AF_cluster yielded few or no usable models. Table 1 also reports the percentage of models generated by SPEACH_AF that passed the two-step parsing. Subsequent to the two-step parsing, the remaining models underwent relaxation with Rosetta and were subsequently compared to the unrelaxed models by using the principal components of the unrelaxed models.

A canonical example of conformational flexibility, crystal structures of *E. coli* adenylate kinase (AK) in various catalytic states highlight two flexible domains that independently bind ATP and AMP.[45] AK models generated by SPEACH_AF and AF_cluster demonstrated little overlap (Fig. 3A), which was not affected by subsequent Rosetta relaxation, although changes in the layout of the models relative to each

other was observed (Fig. 3C). Evaluation of the models with the relative Rosetta Energy, ΔG, indicates several regions with low energy models (Fig. 3E). A plot of the energy vs the first principal component indicates that the lowest energy model is near the closed state, 1AKE, with the next 4 lowest models lying between the fully open and closed state (Fig. 3G and 3I). If the movements of the ATP lid domain (Lid) and the AMP hinge domain (Hinge) relative to the core domain are independent there would be four main conformations of AK: Lid closed/Hinge closed, Lid open/Hinge closed, Lid closed/Hinge open, and Lid closed/Hinge closed.[46–48] As per most analysis of AK domain movements, the angle of the Lid and Hinge domains relative to the core domain were calculated (Fig. S5). This plot is similar to the plot of the two principal components (Fig. 3C) supporting that the PCA is reflective of these domain movements. It is apparent that three of the four main conformations are well populated by low energy models with only the Lid closed/Hinge open conformation not represented by a low energy model. Regardless, the correspondence of the PCA and the known movements of AK highlights the robustness of this approach. Identifying the energetics of the conformations, both low points and saddle points along conformational space, would aid in picking seeds for downstream molecular dynamic simulations and should allow for a decrease in computational time as was shown for identifying cryptic drug binding sites in AlphaFold2 models.[49]

Unlike AK, modeling of ribose binding protein (RBP) with AF_cluster generated few models that passed the two-step filtering (Table 1). There is a change in the conformational space covered by the Rosetta-minimized models (Fig. 3B and 3D). The decrease in conformational space explored by the Rosetta minimized models is due to a decrease in the twist of the two domains relative to each other. The majority of the low energy models cluster near the closed conformation (Fig. 3F), though there are some lower energy models midway between the closed and open conformations of RBP (Fig. 3H and 3J). We parameterized the AlphaFold2 models with known collective variables.[34,35] The plot of the tilt and twist angle of the two domains relative to each other indicate that the two collective variables are interdependent (Fig. S6). The linear spread of AlphaFold2 models between the closed (holo) and open (apo) structures are consistent with the trajectories from molecular dynamic simulations.[34,35]

The rest of the benchmark proteins are transporters or receptors used in previous examinations of protein conformational diversity generated by AlphaFold2.[16,20] The conformational space explored by the AF_cluster and SPEACH_AF models differ for the two MFS transporters, MCT1 and STP10. For MCT1, the models span the range between the two crystal structures (Fig. S7A-B). The SPEACH_AF models tend to segregate near the two crystal structures whereas the AF_cluster models fill the space between the two crystal structures. There are models of similar low energy across the entire conformational space as defined by PC1. It is unclear if the low energy models in the center are distinct energy wells or distinct points due to under-sampling. In contrast, there is a clearer segregation of the models generated by AF_cluster and SPEACH_AF for STP10 (Fig. S7A-B). The SPEACH_AF models span PC1 between the two crystal structures whereas the AF_cluster models span a more diffuse conformational space. The low energy models reside between the two crystal structures (Fig. S7C-D).

For the MOPS transporter MurJ, there is overlap between the conformations generated by AF_cluster and SPEACH_AF (Fig. S7A-B). The low energy models are all in a single cluster slightly biased towards the outward-open conformation, 6NC9 (Fig. S7C-D). Conversely, for PfMATE, which is a member of the same MOPS superfamily, there is almost no overlap between the models generated by AF_cluster and SPEACH_AF (Fig. S7A-B). The low energy models appear to span most of the conformational space from

the outward-open crystal structure, 3VVN, to the models corresponding to an occluded conformation (Fig. S7C-D).

Similar to most examples in the benchmark set, the AF_cluster and SPEACH_AF models have very little overlap for the two LeuT-fold transporters, LAT1 and SERT (Fig. S8A-B). For LAT1, the SPEACH_AF models are in a tighter cluster spanning the range between the two crystal structures, while the AF_cluster models span a more diffuse conformational space between the two crystal structures. Examining the energetics of the models for LAT1, we find one main cluster near the outward-open structure, 7DSQ, with a few slightly higher energy models closer to the inward-open structure, 6IRS (Fig. S8C-D). For SERT the configuration of the AF_cluster and SPEACH_AF models are different compared to LAT1, as they span the conformational space between the two crystal structures, but in a continuous nature with little overlap (Fig. S8A-B). Similar to LAT1, the low energy models all cluster near the outward-open structure, 5I6X (Fig. S8C-D).

The next two proteins of the benchmark set are not functionally related. ASCT2 is a sodium-dependent exchanger of neutral amino acids that is thought to work via an elevator-type mechanism. There are very few conformations generated by AF_cluster that pass the two-step filtering method, but in this case the models appear to overlap with the SPEACH_AF models (Fig. S8A-B). The AlphaFold2 models are strongly bifurcated and align with the two crystal structures with the low energy models clustered near the 6RVX structure, where the elevator is in the up position. ZnT8 is a zinc transporter that is a member of the cation diffusion facilitator family. Here again, the AF_cluster and SPEACH_AF models have very little overlap. Interestingly, the SPEACH_AF models appear to form a structural progression across the conformational space, but it is the AF_cluster models that collect near one of the crystal structures, 6XPF_A (Fig. S8A-B). The low energy models span the conformational space from one crystal structure to the other.

There are four G-protein coupled receptor (GPCR) proteins in the benchmark set. For three of the proteins there are few or no viable models generated with AF_cluster (Table 1). One of these, CGRPR, also had a lower percentage of viable SPEACH_AF models compared to the rest of the benchmark set. We speculate that the information content of the MSA may contribute to the sparsity of models that survive the two-step filtering. Methods to optimize the information content of the MSA for AlphaFold2 predictions is still an open question.[50] Similarly, the AF_cluster models of CCR5 have little overlap with the SPEACH_AF models in PCA space (Fig. S9A) and relaxation of these models with Rosetta leads to only minor structural changes (Fig. S9B). For the other three GPCRs there are not enough AF_cluster models to compare with the SPEACH_AF models. The pattern of low energy models relative to the experimentally determined structures varies across the four GPCRS (Fig. S9C). This could be due to the conditions that the structures were obtained, such as in the presence of G-protein or inhibitor, and that these conformations are only low energy states in their presence.[51] The inability to sample the full conformational space with the GPCRs is similar to two recent studies examining the ability of AlphaFold2 to generate ensembles for GPCRs.[18,19] In both studies they were only able to obtain more diverse models by using templates to bias AlphaFold2. This would suggest that while AlphaFold2 has learned elements of the protein energy landscape as encoded in the MSA, it is unable to predict how the landscape is altered in the presence of ligands or accessory proteins.

For the majority of proteins in the benchmark set, the models generated by SPEACH_AF and AF_cluster have very little overlap in the conformational space depicted by PCA. By mapping the energy

scores of the models on the conformational space, it is clear that the low energy models generally reside in one subset or the other. To further examine this, histograms of the energetics of the models for the two methods are plotted (Fig. S10-11). In some cases, there is a clear correlation of one of the methods with the low energy models and this correlation might track with the fold/function of the protein. For example, SPEACH_AF yielded more low energy models for the LeuT-fold proteins, while for the cation diffusion facilitator ZnT8 AF_cluster models clearly have lower energy. The differential position in conformational space and energetics seen here clearly supports a previous study showing that the MSA content has a role in the models generated by AlphaFold2.[50]

How then does the information content of the MSA reflect on methods that generate models using a single sequence?[52,53] Could the condensed MSAs highlight how to boost the information content or aid in the learning of the linear sequence for these methods? As we approach the expanded use of AlphaFold2 in structural biology pipelines, further work on how to maximize the information content of the MSA is warranted.[50] In fact, in CASP 15 two of the methods achieved increased performance over the baseline AlphaFold2 by optimizing the MSA.[54,55] The question remains how to best create an MSA to yield the breadth of conformations or increase sampling in a computationally tractable way. One possible method is to combine AF_cluster with SPEACH_AF as was done with KaiB, where point mutations placed across the MSA could cause the fold switch in KaiB to occur with a single sub-sampled MSA.[21]

CONCLUSIONS

While AF_cluster was shown to be able to generate both conformations for fold-switch proteins, we find that it is also able to generate conformational ensembles that are complementary to those generated with SPEACH_AF. Rosetta scoring of these models effectively map out energy landscapes that track with previously reported conformations and energetics. We have also shown that AlphaFold2 is capable of generating models that are sensitive to point mutations. The structural changes in the models are more robust when the mutation(s) are in the whole MSA versus just in the input sequence. The combined results support that AlphaFold2 has learned information about protein energy landscapes from information encoded in the MSAs and the crystal structures that it was trained on.

ACKNOWLEDGEMENTS

We thank Derek Claxton for many helpful discussions and critically reading the manuscript. This work was supported by the National Institute of Health grants GM128087 and GM077659 awarded to HSM.

REFERENCES

1. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J (2021) Critical assessment of methods of protein structure prediction (CASP)—Round XIV. Proteins: Structure, Function, and Bioinformatics 89:1607–1617.

2. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589.

3. Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Žídek A, Bridgland A, Cowie A, Meyer C, Laydon A, et al. (2021) Highly accurate protein structure prediction for the human proteome. Nature 596:590–596.

4. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. Nat Methods:1–4.

5. Buel GR, Walters KJ (2022) Can AlphaFold2 predict the impact of missense mutations on structure? Nat Struct Mol Biol 29:1–2.

6. Chuah A, Li S, Do A, Field M, Andrews D (2022) StabilitySort: assessment of protein stability changes on a genome-wide scale to prioritize potentially pathogenic genetic variation. Bioinformatics 38:4220–4222.

7. Iqbal S, Ge F, Li F, Akutsu T, Zheng Y, Gasser RB, Yu D-J, Webb GI, Song J (2022) PROST: AlphaFold2-aware Sequence-Based Predictor to Estimate Protein Stability Changes upon Missense Mutations. J. Chem. Inf. Model. 62:4270–4282.

8. Schmidt A, Röner S, Mai K, Klinkhammer H, Kircher M, Ludwig KU (2023) Predicting the pathogenicity of missense variants using features derived from AlphaFold2. Bioinformatics:btad280.

9. Pak MA, Markhieva KA, Novikova MS, Petrov DS, Vorobyev IS, Maksimova ES, Kondrashov FA, Ivankov DN (2023) Using AlphaFold to predict the impact of single mutations on protein stability and function. PLOS ONE 18:e0282689.

10. Keskin Karakoyun H, Yüksel ŞK, Amanoglu I, Naserikhojasteh L, Yeşilyurt A, Yakıcıer C, Timuçin E, Akyerli CB (2023) Evaluation of AlphaFold structure-based protein stability prediction on missense variations in cancer. Frontiers in Genetics [Internet] 14. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2023.1052383

11. Li C, Zhang L, Zhuo Z, Su F, Li H, Xu S, Liu Y, Zhang Z, Xie Y, Yu X, et al. (2023) Artificial intelligence-based recognition for variant pathogenicity of BRCA1 using AlphaFold2-predicted structures. Theranostics 13:391–402.

12. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK (2023) Before and after AlphaFold2: An overview of protein structure prediction. Frontiers in Bioinformatics [Internet] 3. Available from: https://www.frontiersin.org/articles/10.3389/fbinf.2023.1120370

13. Wodak SJ, Vajda S, Lensink MF, Kozakov D, Bates PA (2023) Critical Assessment of Methods for Predicting the 3D Structure of Proteins and Protein Complexes. Annual Review of Biophysics 52:183–206.

14. Varadi M, Bordin N, Orengo C, Velankar S (2023) The opportunities and challenges posed by the new generation of deep learning-based protein structure predictors. Current Opinion in Structural Biology 79:102543.

15. Yang Z, Zeng X, Zhao Y, Chen R (2023) AlphaFold2 and its applications in the fields of biology and medicine. Sig Transduct Target Ther 8:1–14.

16. Stein RA, Mchaourab HS (2022) SPEACH_AF: Sampling protein ensembles and conformational heterogeneity with Alphafold2. PLOS Computational Biology 18:e1010483.

17. Xiao Q, Xu M, Wang W, Wu T, Zhang W, Qin W, Sun B (2022) Utilization of AlphaFold2 to Predict MFS Protein Conformations after Selective Mutation. International Journal of Molecular Sciences 23:7235.

18. Heo L, Feig M (2022) Multi-state modeling of G-protein coupled receptors at experimental accuracy. Proteins: Structure, Function, and Bioinformatics 90:1873–1885.

19. Sala D, Hildebrand PW, Meiler J (2023) Biasing AlphaFold2 to predict GPCRs and kinases with user-defined functional or structural properties. Frontiers in Molecular Biosciences [Internet] 10. Available from: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1121962

20. del Alamo D, Sala D, Mchaourab HS, Meiler J (2022) Sampling alternative conformational states of transporters and receptors with AlphaFold2. eLife 11:e75751.

21. Wayment-Steele HK, Ovchinnikov S, Colwell L, Kern D (2022) Prediction of multiple conformational states by combining sequence clustering with AlphaFold2. :2022.10.17.512570. Available from: https://www.biorxiv.org/content/10.1101/2022.10.17.512570v1

22. Chakravarty D, Porter LL (2022) AlphaFold2 fails to predict protein fold switching. Protein Science 31:e4353.

23. Lane TJ (2023) Protein structure prediction has reached the single-structure frontier. Nat Methods 20:170–173.

24. Akdel M, Pires DEV, Pardo EP, Jänes J, Zalevsky AO, Mészáros B, Bryant P, Good LL, Laskowski RA, Pozzati G, et al. (2022) A structural biology community assessment of AlphaFold2 applications. Nat Struct Mol Biol 29:1056–1067.

25. Mirdita M, Steinegger M, Söding J (2019) MMseqs2 desktop and local web server app for fast, interactive sequence searches. Bioinformatics 35:2856–2858.

26. Zhang Y, Skolnick J (2005) TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Research 33:2302–2309.

27. Fleishman SJ, Leaver-Fay A, Corn JE, Strauch E-M, Khare SD, Koga N, Ashworth J, Murphy P, Richter F, Lemmon G, et al. (2011) RosettaScripts: A Scripting Language Interface to the Rosetta Macromolecular Modeling Suite. PLOS ONE 6:e20161.

28. Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popović Z, Baker D, Players F (2011) Algorithm discovery by protein folding game players. Proceedings of the National Academy of Sciences 108:18949–18953.

29. Maguire JB, Haddox HK, Strickland D, Halabiya SF, Coventry B, Griffin JR, Pulavarti SVSRK, Cummins M, Thieker DF, Klavins E, et al. (2021) Perturbing the energy landscape for improved packing during computational protein design. Proteins: Structure, Function, and Bioinformatics 89:436–449.

30. Bakan A, Meireles LM, Bahar I (2011) ProDy: Protein Dynamics Inferred from Theory and Experiments. Bioinformatics 27:1575–1577.

31. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A (2015) The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Research 43:W401–W407.

32. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O (2011) MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. Journal of Computational Chemistry 32:2319–2327.

33. Gowers RJ, Linke M, Barnoud J, Reddy TJE, Melo MN, Seyler SL, Domański J, Dotson DL, Buchoux S, Kenney IM, et al. (2016) MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the 15th Python in Science Conference:98–105.

34. Dokainish HM, Sugita Y (2021) Exploring Large Domain Motions in Proteins Using Atomistic Molecular Dynamics with Enhanced Conformational Sampling. International Journal of Molecular Sciences 22:270.

35. Ravindranathan KP, Gallicchio E, Levy RM (2005) Conformational Equilibria and Free Energy Profiles for the Allosteric Transition of the Ribose-binding Protein. Journal of Molecular Biology 353:196–210.

36. Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro ANA, Iversen ES, Couch FJ, et al. (2007) A Systematic Genetic Assessment of 1,433 Sequence Variants of Unknown Clinical Significance in the BRCA1 and BRCA2 Breast Cancer–Predisposition Genes. The American Journal of Human Genetics 81:873–883.

37. Williams RS, Green R, Glover JNM (2001) Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nat Struct Mol Biol 8:838–842.

38. Wang Q, Goh AM, Howley PM, Walters KJ (2003) Ubiquitin Recognition by the DNA Repair Protein hHR23a. Biochemistry 42:13529–13535.

39. He F, Wollscheid H-P, Nowicka U, Biancospino M, Valentini E, Ehlinger A, Acconcia F, Magistrati E, Polo S, Walters KJ (2016) Myosin VI Contains a Compact Structural Motif that Binds to Ubiquitin Chains. Cell Reports 14:2683–2694.

40. Jiang X, Wu J, Ke M, Zhang S, Yuan Y, Lin JY, Yan N (2019) Engineered XylE as a tool for mechanistic investigation and ligand discovery of the glucose transporters GLUTs. Cell Discov 5:1–4.

41. Quistgaard EM, Löw C, Moberg P, Trésaugues L, Nordlund P (2013) Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. Nat Struct Mol Biol 20:766–768.

42. Martens C, Shekhar M, Borysik AJ, Lau AM, Reading E, Tajkhorshid E, Booth PJ, Politis A (2018) Direct protein-lipid interactions shape the conformational landscape of secondary transporters. Nat Commun 9:4151.

43. Peccati F, Alunno-Rufini S, Jiménez-Osés G (2023) Accurate Prediction of Enzyme Thermostabilization with Rosetta Using AlphaFold Ensembles. J. Chem. Inf. Model. 63:898–909.

44. Roney JP, Ovchinnikov S (2022) State-of-the-Art Estimation of Protein Model Accuracy Using AlphaFold. Phys. Rev. Lett. 129:238101.

45. Miller MD, Phillips GN (2021) Moving beyond static snapshots: Protein dynamics and the Protein Data Bank. Journal of Biological Chemistry 296:100749.

46. Ping J, Hao P, Li Y-X, Wang J-F (2013) Molecular Dynamics Studies on the Conformational Transitions of Adenylate Kinase: A Computational Evidence for the Conformational Selection Mechanism. BioMed Research International 2013:e628536.

47. Li D, Liu MS, Ji B (2015) Mapping the Dynamics Landscape of Conformational Transitions in Enzyme: The Adenylate Kinase Case. Biophysical Journal 109:647–660.

48. Beckstein O, Denning EJ, Perilla JR, Woolf TB (2009) Zipping and Unzipping of Adenylate Kinase: Atomistic Insights into the Ensemble of Open↔Closed Transitions. Journal of Molecular Biology 394:160–176.

49. Meller A, Bhakat S, Solieva S, Bowman GR (2023) Accelerating Cryptic Pocket Discovery Using AlphaFold. J. Chem. Theory Comput. [Internet]. Available from: https://doi.org/10.1021/acs.jctc.2c01189

50. Petti S, Bhattacharya N, Rao R, Dauparas J, Thomas N, Zhou J, Rush AM, Koo P, Ovchinnikov S (2023) End-to-end learning of multiple sequence alignments with differentiable Smith–Waterman. Bioinformatics 39:btac724.

51. Deupi X, Kobilka BK (2010) Energy Landscapes as a Tool to Integrate GPCR Structure, Dynamics, and Function. Physiology 25:293–303.

52. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, et al. (2022) Evolutionary-scale prediction of atomic level protein structure with a language model. :2022.07.20.500902. Available from: https://www.biorxiv.org/content/10.1101/2022.07.20.500902v3

53. Chowdhury R, Bouatta N, Biswas S, Floristean C, Kharkar A, Roy K, Rochereau C, Ahdritz G, Zhang J, Church GM, et al. (2022) Single-sequence protein structure prediction using a language model and deep learning. Nat Biotechnol 40:1617–1623.

54. Peng Z, Wang W, Wei H, Li X, Yang J Improved protein structure prediction with trRosettaX2, AlphaFold2, and optimized MSAs in CASP15. Proteins: Structure, Function, and Bioinformatics [Internet] n/a. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.26570

55. Oda T Improving protein structure prediction with extended sequence similarity searches and deep-learning-based refinement in CASP15. Proteins: Structure, Function, and Bioinformatics [Internet] n/a. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.26551

Table 1: Initial Screen of AlphaFold2-generated Models[a]				
protein	AFC_3[b]	AFC_7[b]	AFC_11[b]	SPEACH_AF
AK	289/2395 (12.1)	45/305 (14.8)	11/120 (9.2)	300/345 (87.0)
RBP	5/1715 (0.3)	5/330 (1.5)	5/100 (5)	420/422 (99.5)
MCT1	129/1320 (9.8)	198/360 (55.0)	169/235 (71.9)	690/692 (99.7)
STP10	950/1695 (56.0)	518/545 (95.0)	301/305 (98.7)	735/735 (100)
MurJ	644/1380 (46.7)	361/445 (81.1)	199/245 (81.2)	507/677 (74.9)
PfMATE	588/665 (88.4)	120/120 (100)	125/125 (100)	690/692 (99.7)
Lat1	486/1780 (27.3)	429/670 (64.0)	269/325 (82.8)	705/705 (100)
SERT	176/1500 (11.7)	72/570 (12.6)	154/420 (36.7)	825/825 (100)
ASCT2	60/1245 (4.8)	52/310 (16.8)	66/130 (50.8)	690/693 (99.6)
ZnT8	604/1905 (31.7)	376/515 (91.7)	220/240 (91.7)	450/465 (96.8)
CCR5	430/1095 (39.3)	303/435 (69.7)	221/270 (81.8)	447/452 (98.9)
CGRPR	0/520 (0)	0/205 (0)	3/140 (2.1)	217/390 (55.6)
FZD7	10/560 (1.8)	7/170 (4.1)	5/105 (4.8)	508/510 (99.6)
PTH1R	0/105 (0)	0/50 (0)	0/35 (0)	389/452 (86.1)

a: Shown are the number of models that passed the initial two-step screen relative to the total number of models generated with the percentage given in parentheses.
b: AFC_3, AFC_7, and AFC_11 utilizing AF_cluster with the minimum number of sequences set to 3, 7, and 11 respectively.

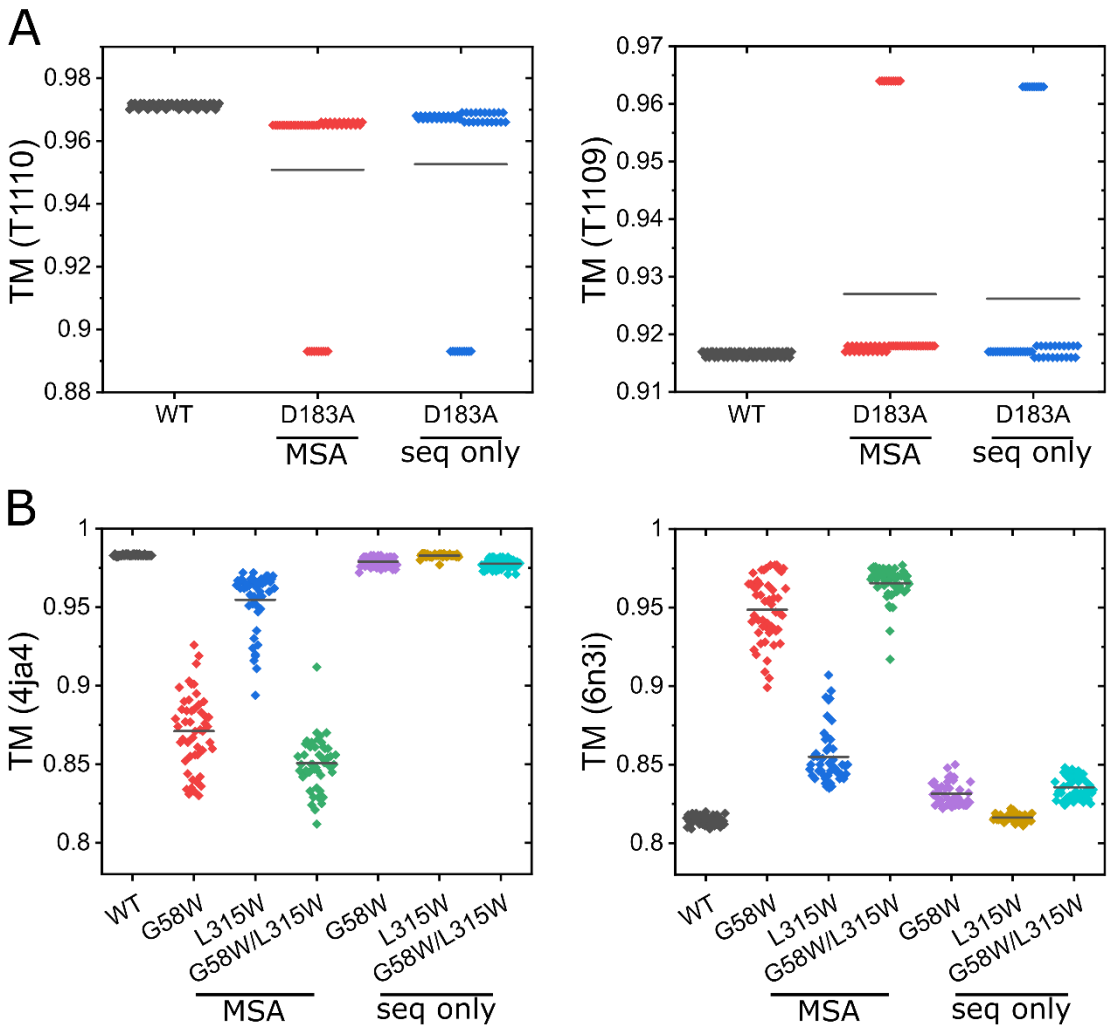

Figure 1. Comparison of WT and Mutant models to crystal structures. A) TM scores for the AlphaFold2 predicted models for wild-type and D183A isocyanide hydratase compared to the target monomers from CASP 15. On the left is the wild-type target (T1110) and on the right is the mutant target (T1109). B) TM scores for the AlphaFold2 predicted models for wild-type and mutant models compared to wild-type (4ja4) and G58W/L315W (6n3i) crystal structures of XylE.

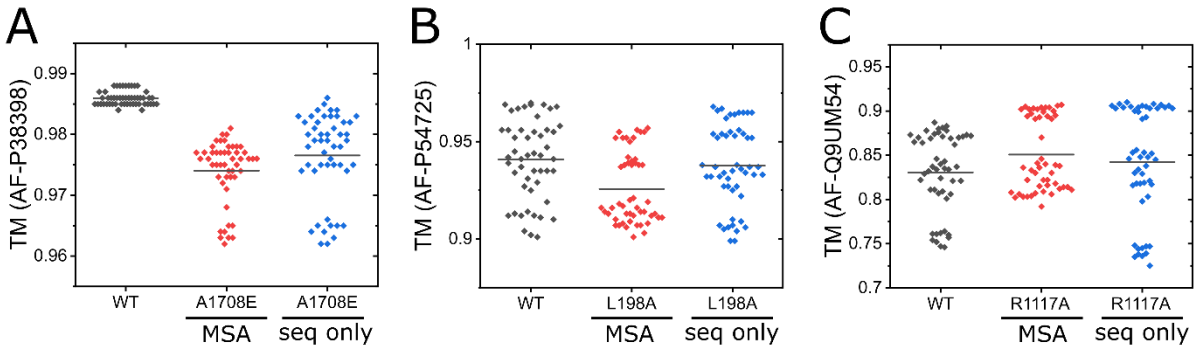

Figure 2. Comparison of WT and Mutant models to reference models. For each of the sets of models a t-test was done to ascertain the confidence that each mutant model set is different from the wild-type model set. A) BRCT domain of BRCA1. The p-values are 7.46E-28 for in the whole MSA and 2.02E-14 for in the input sequence only. B) UBA domain of HR23A. The p-values are 1.6E-4 for in the whole MSA and 0.462 for in the input sequence only. C) MyUB domain of Myosin VI. The p-values are 0.044 for in the whole MSA and 0.456 for in the input sequence only.

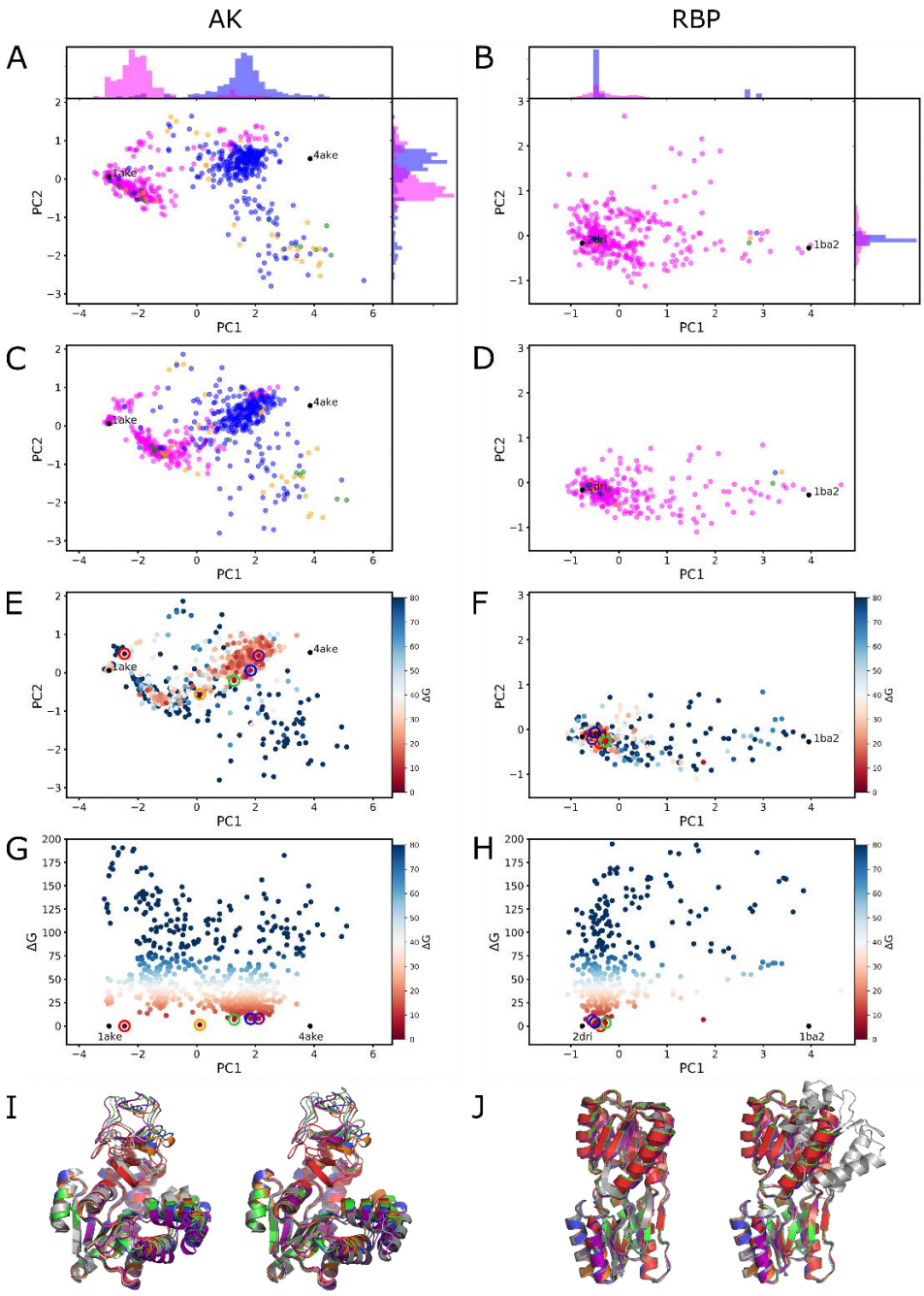

Figure 3. Principal Component Analysis of Adenylate Kinase (AK) and Ribose Binding Protein (RBP). A&B) The first two principal components for the two-step parsed models. The SPEACH_AF models are in magenta, and the AF_cluster models are blue (3), gold (7), and green (11). C&D) The first two principal components for the models after Rosetta relaxation. Colors as in A&B. E&F) Same plots as in C&D except with the coloring based on relative Rosetta Energy Score, ΔG. The five lowest energy models in order, red, orange, green, blue, and purple, are circled. G&H) Plot of ΔG vs PC1. The five lowest energy models, same as in E&F, in order, red, orange, green, blue, and purple, are circled. The crystal structures were not energy minimized and are placed at zero ΔG and their PC1 position for comparison. I) Five lowest energy models shown with the structure from 1ake (left) and 4ake (right). J) Five lowest energy models shown with the structure from 2dri (left) and 1ba2 (right).

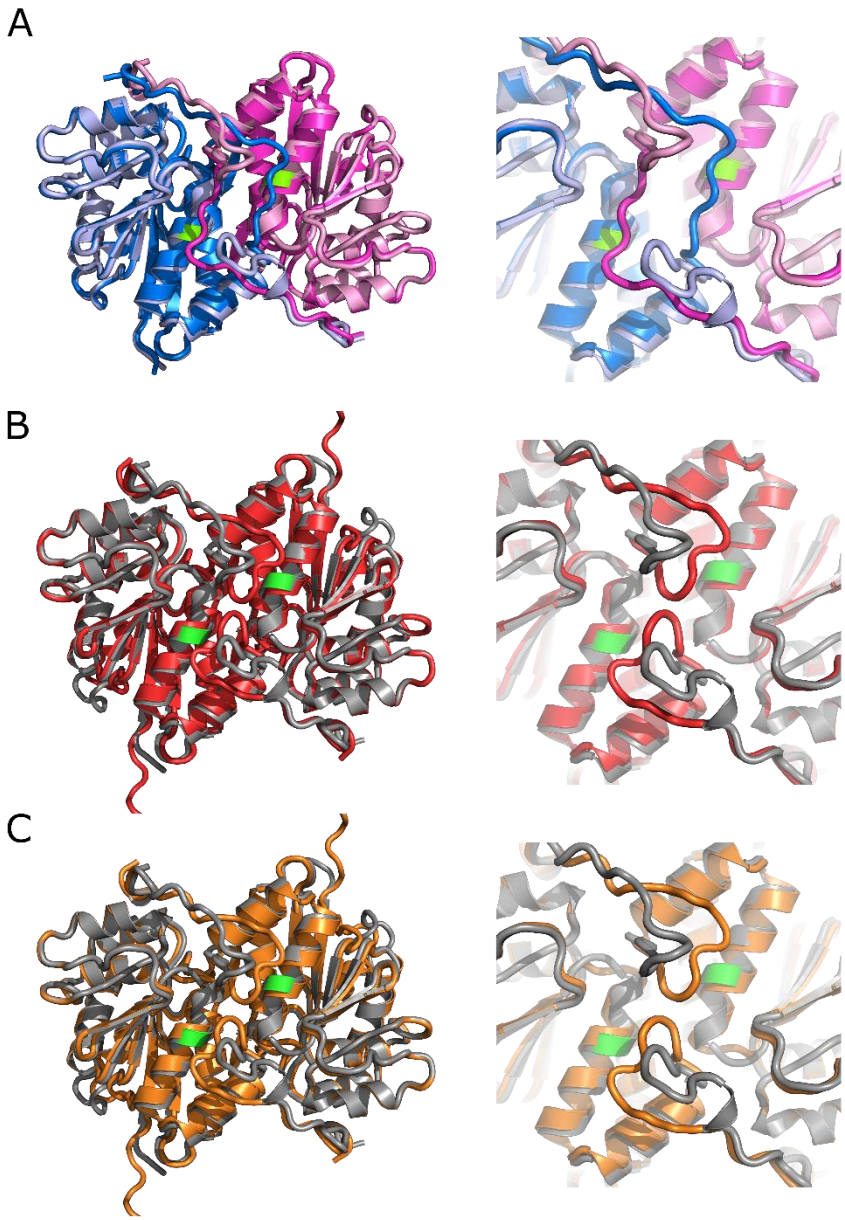

Figure S1. Effect of mutation on isocyanide hydratase. A) Dimer models from CASP 15. The wild-type protein (T1110) is colored slate and magenta. The mutant protein (T1109) is colored lightblue and lightpink. The D183A mutation is colored green. On the right is a close up of the N-terminal tail that switches structure upon mutation. B) Mutant protein (T1109, grey) and the AlphaFold2 model (red) generated with the mutation across the whole MSA. While different from the crystal structure the AlphaFold2 model has a similar configuration of the N-terminal region. C) Mutant protein (T1109, grey) and the AlphaFold2 model (orange) generated with the mutation in the input sequence only. While different from the crystal structure the AlphaFold2 model has a similar configuration of the N-terminal region.

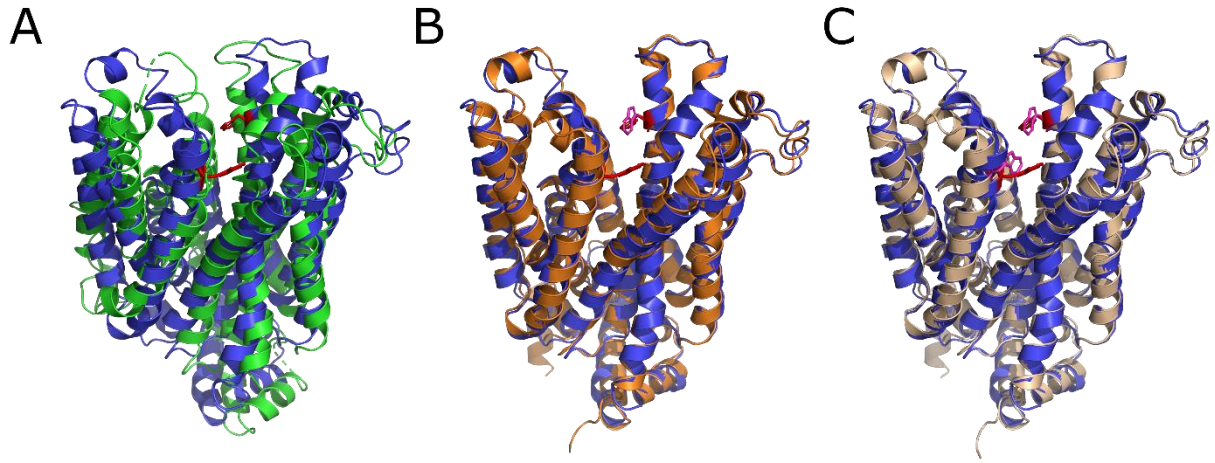

Figure S2. XylE conformational switch. A) Crystal structures of outward-facing (blue, 6n3i) and inward-facing (green, 4ja4) XylE. In red are the mutant residues, G58W/L315W, in 6n3i. B) Outward-facing crystal structure 6n3i (blue) and outward-facing AlphaFold2 model (orange) with the single G58W mutation (magenta). C) Outward-facing crystal structure 6n3i (blue) and outward-facing AlphaFold2 model (wheat) with the G58W/L215W mutation (magenta).

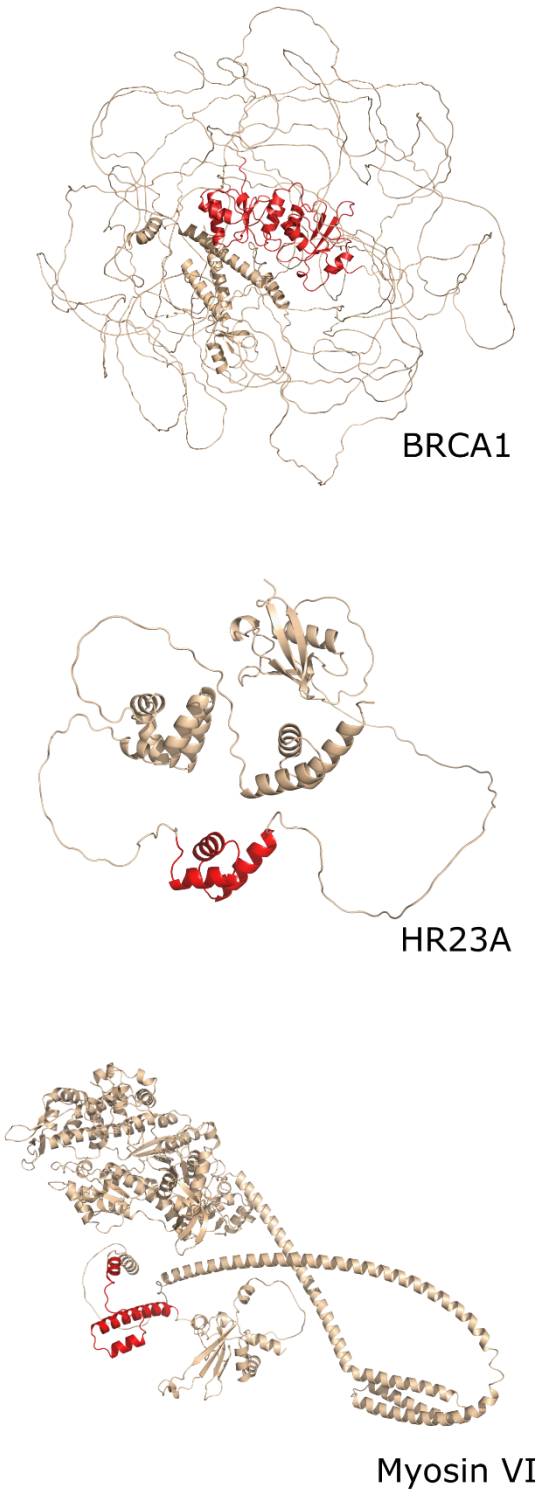

Figure S3. Full length models of BRCA1, HR23A, and Myosin VI. Highlighted in red are the BRCT of BRCA1, UBA of HR23A, and MyUb of Myosin VI domains. The regions in red were used in the comparison to the models of wild-type and mutant domains.

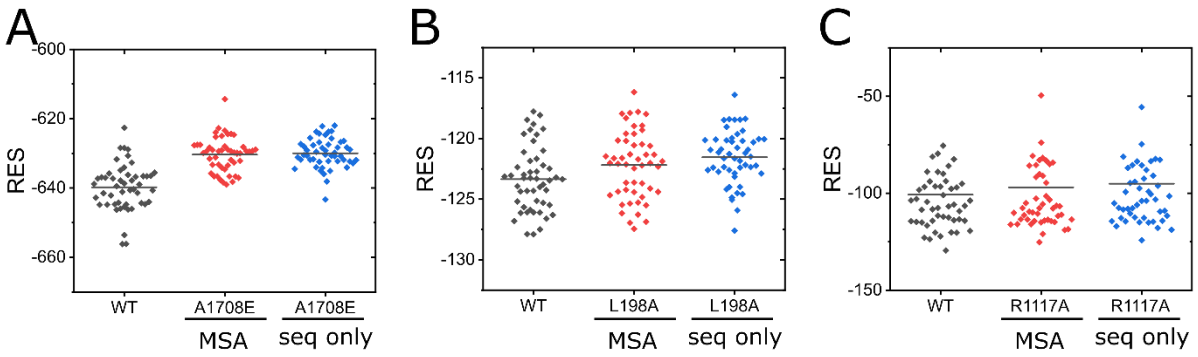

Figure S4. Rosetta Energy Scores for the AlphaFold2 models. For each of the sets of models a t-test was done to ascertain the confidence that the mutant model set is different from the wild-type model set. A) BRCT domain of BRCA1. The p-values are 1.82E-12 for in the whole MSA and 4.45E-14 for in the input sequence only. B) UBA domain of HR23A. The p-values are 0.035 for in the whole MSA and 3.6E-4 for in the input sequence only. C) MyUB domain of Myosin VI. The p-values are 0.531 for in the whole MSA and 0.172 for in the input sequence only.

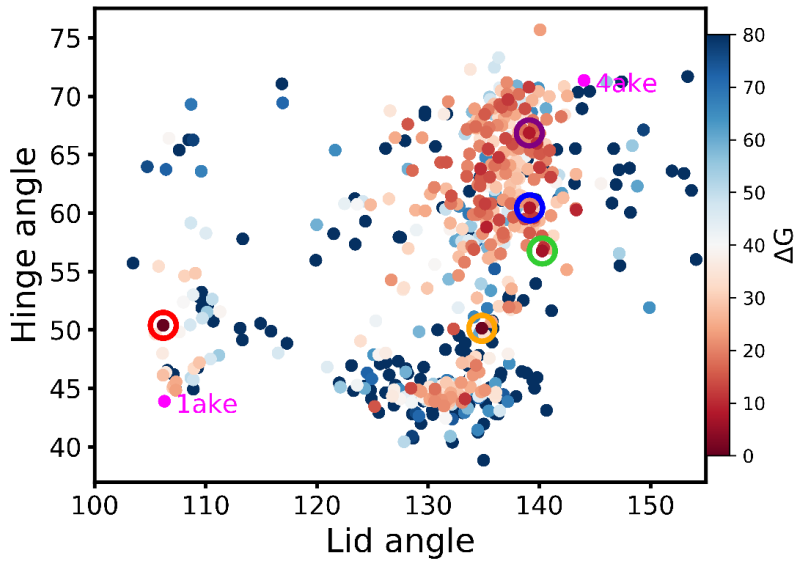

Figure S5. Domain movements in Adenylate Kinase. Plot of the angle between the Lid (ATP binding) domain and Hinge (AMP binding) domain to the core domain. The core domain is comprised of residues 1-29, 60-121, and 160-214; the hinge domain is residues 30-59; and the lid domain is residues 122-159. The coloring of this plot corresponds to Fig. 3C.

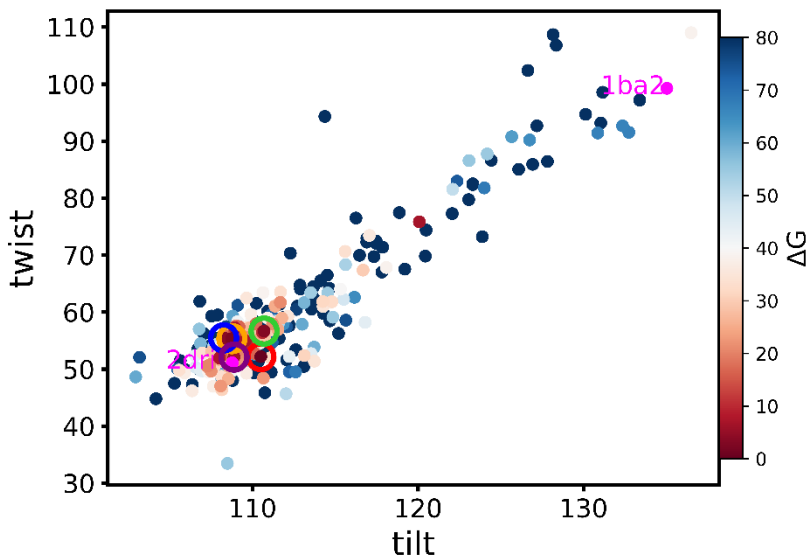

Figure S6. Domain movement in Ribose Binding Protein. Plot of the tilt and twist angle between the N- and C-terminal domains. The N-terminal domain is comprised of residues 1-100 and 236-259 and the C-terminal domain is comprised of residues 108-231 and 269-271. The tilt angle is defined as the angle between the center of mass of the N- and C-terminal domains and the center of mass of the hinge point comprising residues 101-107, 232-235, and 160-268. The twist angle is the dihedral angle of the center of mass of the N- and C-terminal domains and two regions near the ribose binding site on the top of the N-terminal domain, 124-125, 262-262, and 283-284 and the bottom of the C-terminal domain, 133-134, 255-256, 294-295.

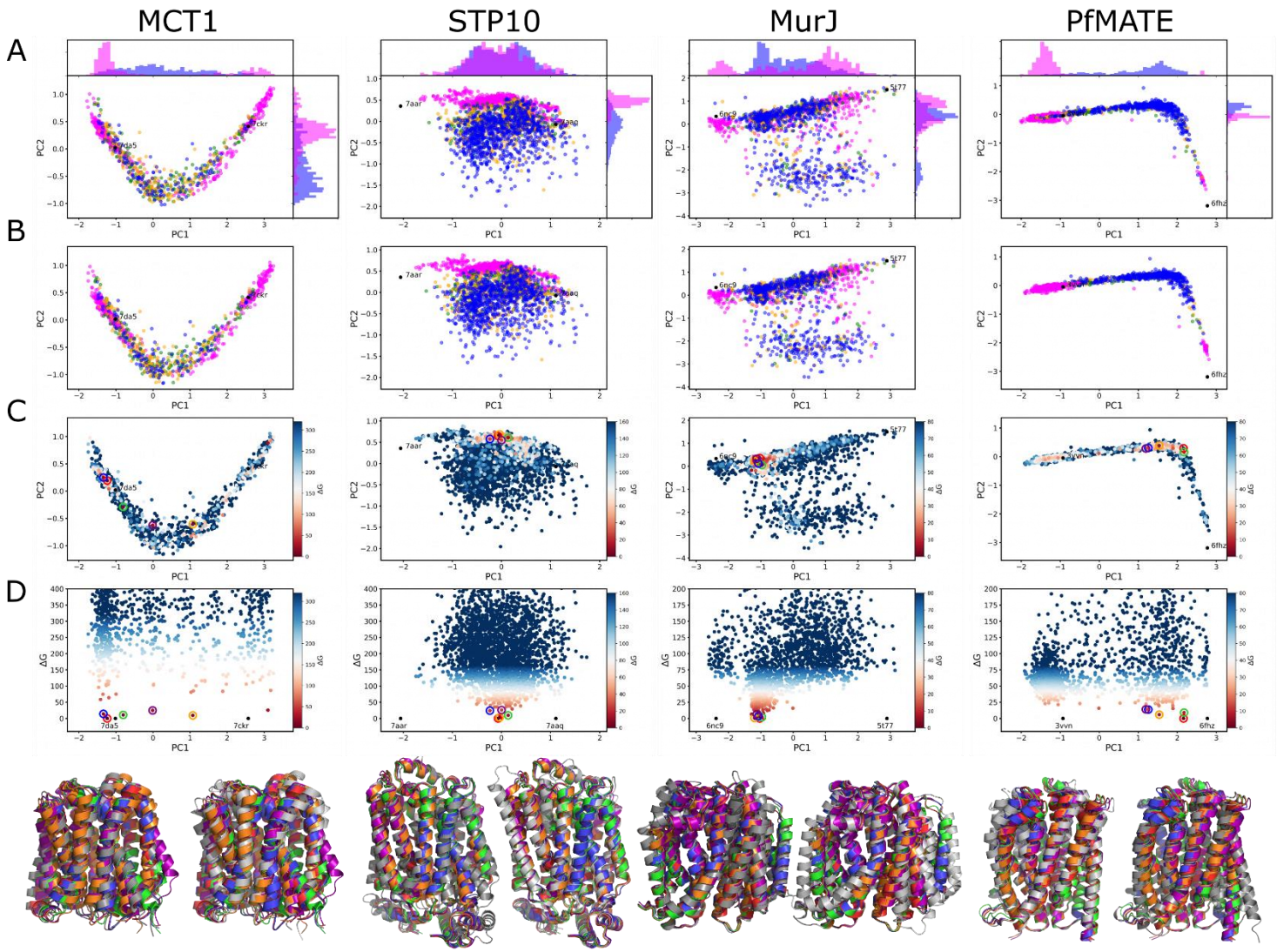

Fig. S7. Principal Component Analysis. Plots are the same as in Fig. 3 of the main paper.

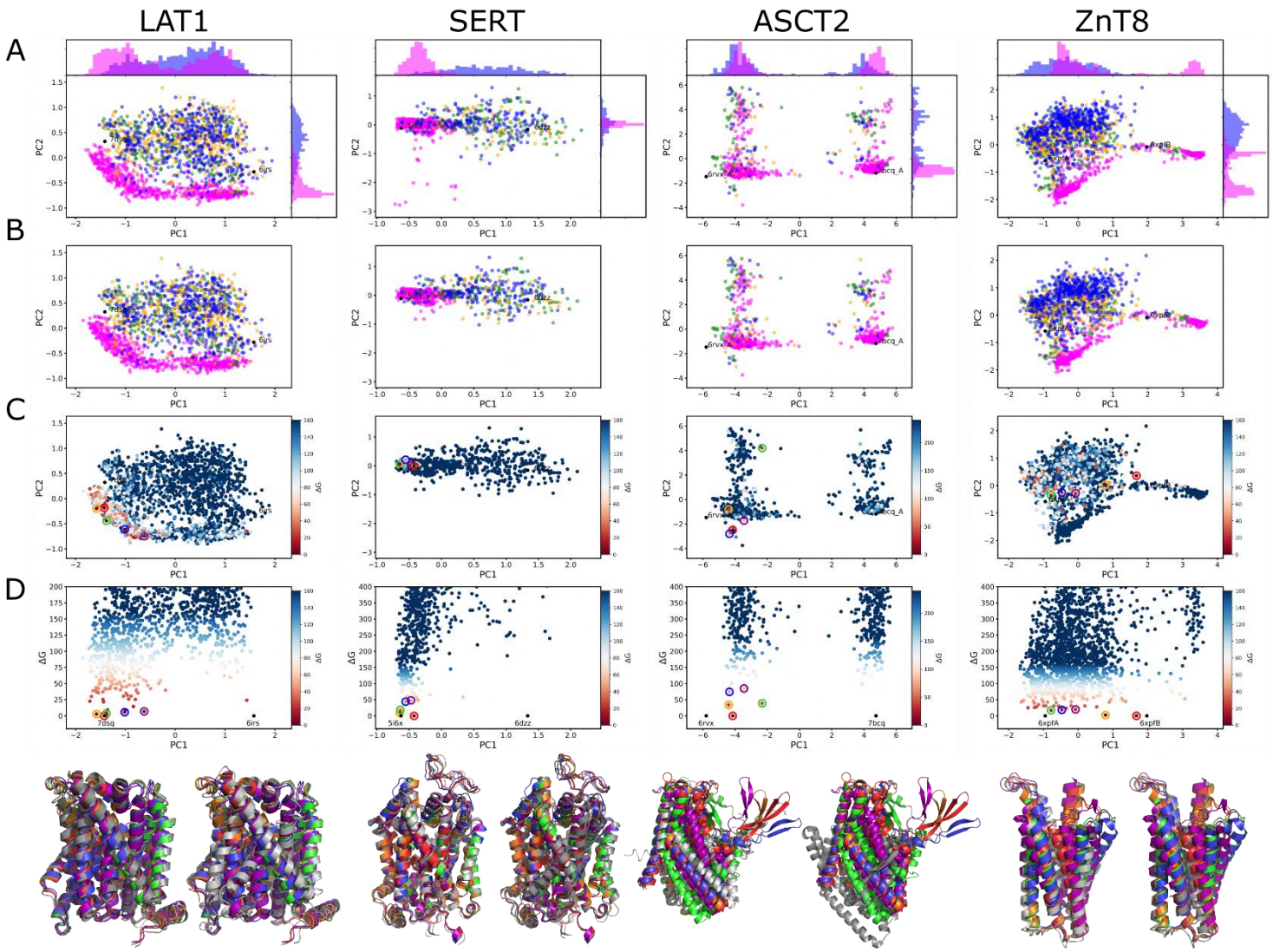

Fig. S8. Principal Component Analysis. Plots are the same as in Fig. 3 of the main paper.

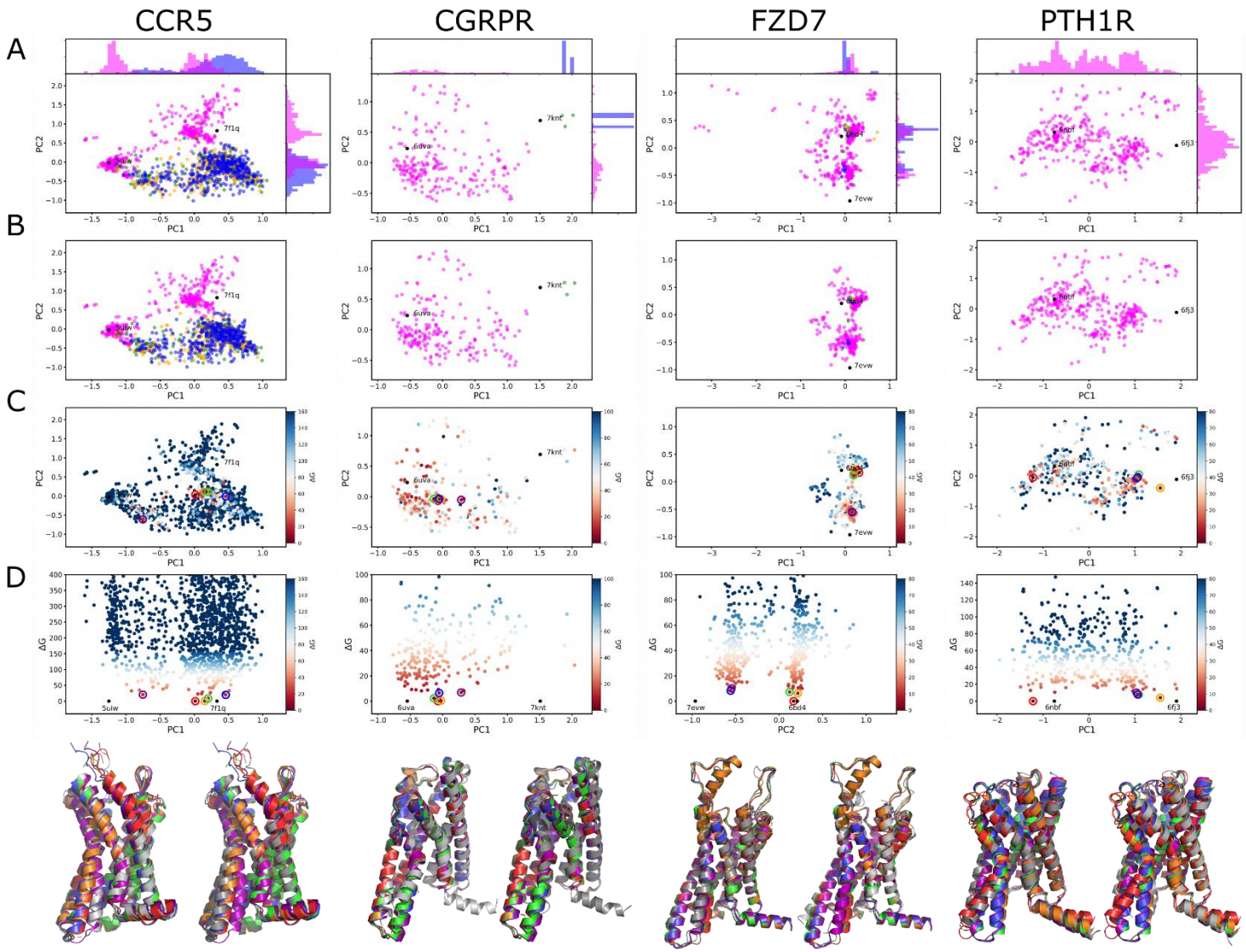

Fig. S9. Principal Component Analysis. Plots are the same as in Fig. 3 of the main paper.

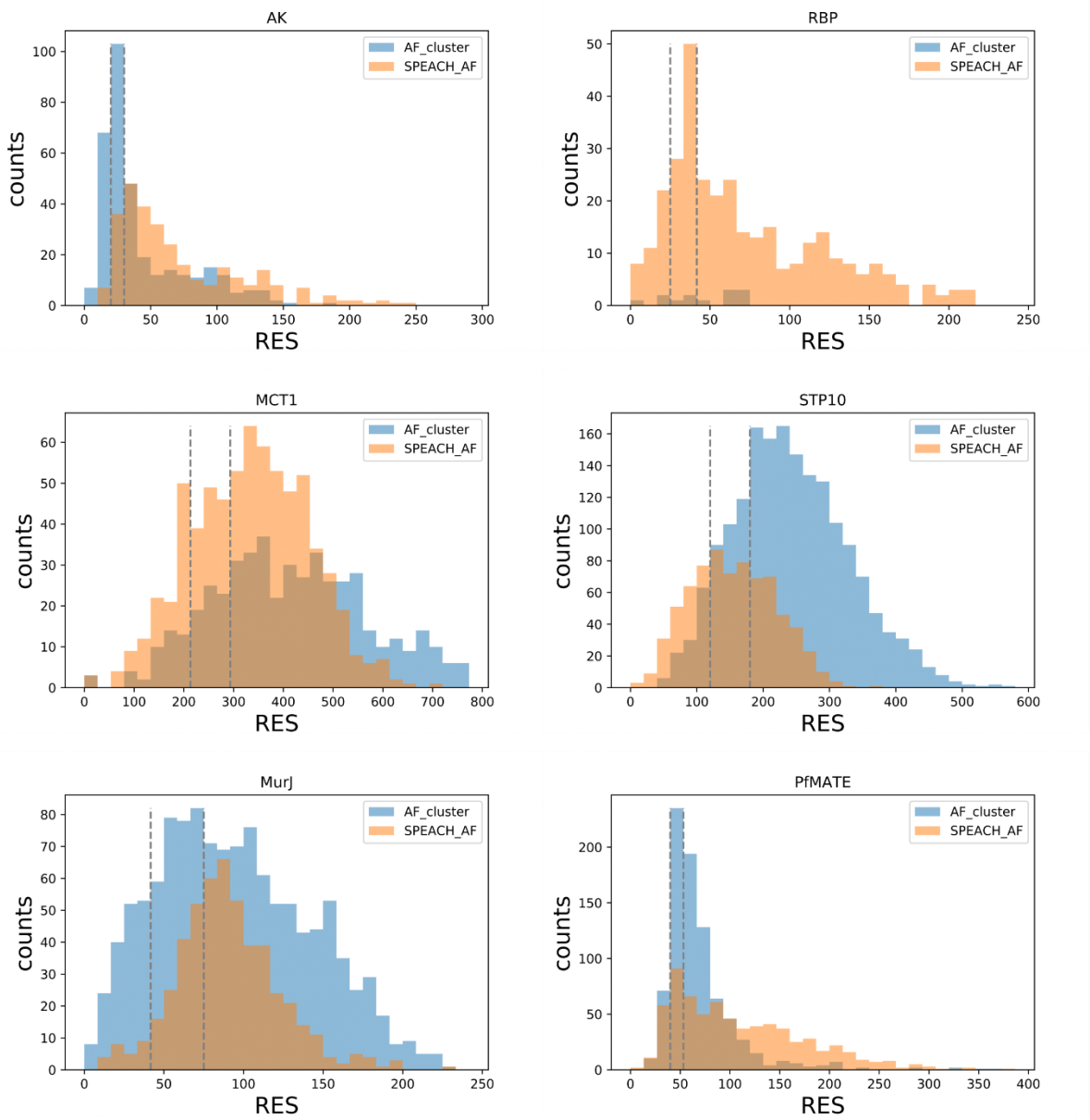

Fig. S10. Histograms of RES for AF_cluster and SPEACH_AF. The gray dashed lines demarcate the ~10% and ~30% lowest scoring models from both methods combined.

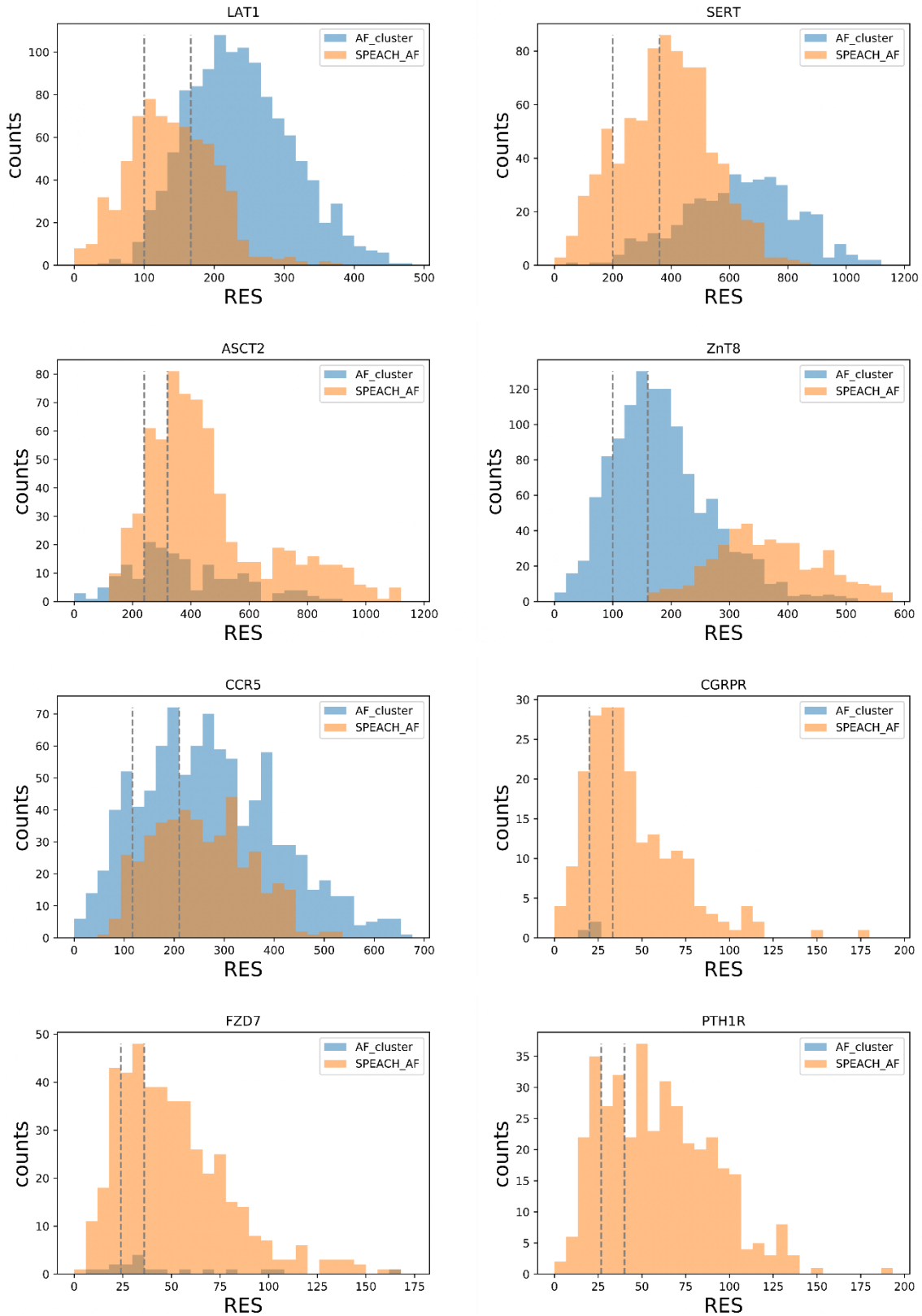

Fig. S11. Histograms of RES for AF_cluster and SPEACH_AF. The gray dashed lines demarcate the ~10% and ~30% lowest scoring models from both methods combined.