State-specific Peptide Design Targeting G Protein-coupled Receptors

Yang Xue*[1], Jun Li*[1], Hong Wang*[2], Jianguo Hu[2], Zhi Zheng[2], Jingzhou He[1], Huanzhang Gong[2,✉], Xiangqun Li[2,✉], Xiaonan Zhang[1,✉], and Xiaomin Fang[1,✉]

[1]Baidu Inc.
[2]Zonsen PepLib Biotech Inc.

G protein-coupled receptors (GPCRs) play critical roles in various physiological processes and serve as pivotal targets for drug discovery. Peptides are particularly compelling therapeutic agents for targeting GPCRs, as they frequently exhibit superior affinity, selectivity, and potency compared to small molecule drugs. However, the dynamic conformational states of GPCRs profoundly influence their interactions with ligands, highlighting the need for state-specific strategies in peptide design. To address this, we developed an efficient state-specific peptide design framework targeting GPCRs. This approach includes an analysis of GPCR state-transition mechanisms and a specially optimized peptide structure prediction model for GPCRs, HelixFold-Multistate, allowing us to evaluate state-specific scores for the designed peptides. On recent GPCR-peptide PDBs, HelixFold-Multistate not only maintains the GPCR active state, similar to AlphaFold-Multistate, but also demonstrates high interaction prediction capabilities comparable to AlphaFold-Multimer in terms of DockQ and iRMS. Experiments on CXCR4-peptide affinity data indicate that its confidence scores exhibit a higher correlation and screening capability than those of AlphaFold-Multistate. By employing this design method, we successfully identified both the agonist and antagonist peptides for the Growth Hormone Secretagogue Receptor (GHSR) and Apelin Receptor (APJ), exhibiting EC50 values below 100 nM and 10 nM respectively. Despite the challenges with antagonist peptides, our method also identified inhibitors for GHSR and APJ, with respective IC50 values of 3.3 µM and 20.3 µM.

Introduction

G-protein-coupled receptors (GPCRs) are a large and diverse group of membrane receptors that play crucial roles in cellular signal transduction. Their involvement in various physiological processes makes them attractive targets for therapeutic intervention (Santos et al., 2017). Despite the success of small molecules and biologics in modulating GPCR activity, these approaches often face limitations such as off-target effects (Weiss et al., 2018). Peptide-based therapeutics offer a promising alternative due to their remarkable pharmacodynamic properties: high affinity, selectivity and potency (Wang et al., 2022; Davenport et al., 2020).

Conventional methods for developing peptide drugs include leveraging natural peptides and hormones found in the human body and creating peptides that mimic these hormones (Alavi et al., 2019; Wilson et al., 2007; Pokuri et al., 2016). Peptides can also be sourced from natural products, including those derived from bacteria, fungi, and plants (Brown et al., 2007; Yamazaki et al., 2009; Toivanen et al., 2017; Mygind et al., 2005; Matsubayashi and Sakagami, 2006). Another approach involves the use of phage display techniques to discover and optimize peptide candidates (Omidfar and Daneshpour, 2015).

Recently, AI-based peptide rational design has drawn more and more attention (Hill et al., 2014; Nevola and Giralt, 2015; Chang et al., 2022). EvoBind (Bryant and Elofsson, 2022; Li et al., 2024) utilizes in-silico directed evolution to design peptides, then evaluate the designs using confidence scores from the folding model

*These authors contributed equally to this work
✉ Corresponding authors: gong_huanzhang@peplib.com (H. Gong), li_xiangqun@peplib.com (X. Li), zhangxiaonan@baidu.com (X. Zhang), fangxiaomin01@baidu.com (X. Fang)

AlphaFold-Multimer (Evans et al., 2021). EvoPlay (Wang et al., 2023) improves it by replacing directed evolution with Monte Carlo tree search (MCTS) and reinforcement learning policy (Silver et al., 2016), leading to more data-efficient peptide design. Additionally, PepMLM (Chen et al., 2023) uses protein language model (PLM) to design peptide binders via modeling the binder as a protein sequence mask filling task. Approaches for designing mini-protein might also be applicable for peptide designs, such as AlphaFold-2 hallucination (Vázquez Torres et al., 2024), RFDiffusion-MPNN (Vázquez Torres et al., 2024), AlphaProteo (Zambaldi et al., 2024) and BindCraft Pacesa et al., 2024.

While the above AI-driven peptide design approaches have demonstrated effectiveness in creating binders, targeting GPCRs presents a distinct and complex challenge. Unlike standard targets, GPCR-directed peptides must not only achieve binding but also induce conformational shifts, altering the GPCR energy landscape across local minima. The dynamic conformational states of GPCRs play a critical role in these interactions, emphasizing the necessity of incorporating GPCR dynamics into rational peptide design. A straightforward approach would be to employ the state-sensitive folding model AlphaFold-Multistate (Heo and Feig, 2022) instead of AlphaFold-Multimer (Evans et al., 2021) to better estimate the GPCR structures in various states, but we found that AlphaFold-Multistate underperforms in predicting the GPCR-peptide complex structures, making it difficult to understand the interactions between GPCRs and peptide candidates. A bypass way is to directly generate peptide sequences from fixed state-specifc GPCRs structures, then use the generator score alone to filter plausible peptide candidates avoiding the refolding process. Following this approach, HelixDiff (Xie et al., 2024) recently introduced an all-atom diffusion model for designing helix peptides targeting GPCRs. However, it remains untested for generating both agonists and antagonists and focuses only on structurally stable helix peptides, rather than unconstrained designs like peptides with coil regions. Developing unconstrained, state-specific functional peptides that target GPCRs thus remains a significant challenge.

In this study, we present an efficient computational method for state-specific peptide design targeting GPCRs (Figure 1). This pipeline leverages diversified seed peptides backbones and protein sequence generation models (Hsu et al., 2022; Dauparas et al., 2022) to generate peptide sequences, and the state-specific folding model to evaluate the target-peptide complex. After sequence design, we use an optimized GPCRs-peptide structure prediction model to evaluate state-specific scores, then peptides with higher pLDDT and lower interaction pAE are selected as top peptides and their backbones are utilized as new backbone seeds to refine the design if necessary. After that, we run molecule simulation to analyze the key interactions related to GPCRs state-transition mechanism to filter plausible peptides to synthesize and test affinity and functionality. In addition, the optimized GPCRs-peptide structure prediction model (Figure 1b) is based on AlphaFold-Multistate (Heo and Feig, 2022), HelixFold-Multimer (Fang et al., 2024), and fine-tuned on curated protein-peptide, GPCR-peptide structure data. Particularly, we standardize AlphaFold-Multistate method for all GPCRs classes, and re-implement it in HelixFold-Multimer. Then we fine-tune the model on filtered protein-peptide data from Propedia database (Martins et al., 2021), and GPCR-peptide data from GPCRdb (Pándy-Szekeres et al., 2023). This model can keep the GPCR in a correct activation state, while predict GPCR-peptide interaction pose much better than AlphaFold-Multistate in terms of iRMS and DockQ. When evaluated using the publicly available CXCR4-peptides affinity dataset, the structural confidence scores derived from HelixFold-Multistate demonstrated superior correlation compared to those from AlphaFold-Multistate. Notably, the disparity in confidence scores between active and inactive states proved to be indicative of the peptide's potential functional properties. Over two experiments conducted on GHSR and APJ, our pipeline found <100nM affinity peptides using only <3000 designed sequences. Notably, on GHSR which the state-transition mechanism is relatively clear (Qin et al., 2022), our pipeline only requires 200 sequences to meet the affinity standard. On APJ, using a relatively large design pool (3k), our pipeline can find peptides with strong affinity (< 10 nM). For challenging antagonist peptides, we still found inhibitor for GHSR with modest 3.3 µM affinity and for APJ with 20.3 µM affinity.

Results

In our pipeline (Figure 1), the state-specific refolding and evaluating process plays an essential role, because it not only does coarse filtering, but also effects significantly on the results of refined selection based on molecule simulation. Thus, at first, we will report the results of our proposed state-specific GPCR-peptide complex structure prediction model, HelixFold-Multistate, which is built on top of HelixFold-Multimer (Fang et al., 2024). Then, we take GHSR and APJ as two experimental targets, and show the results of each pipeline step.

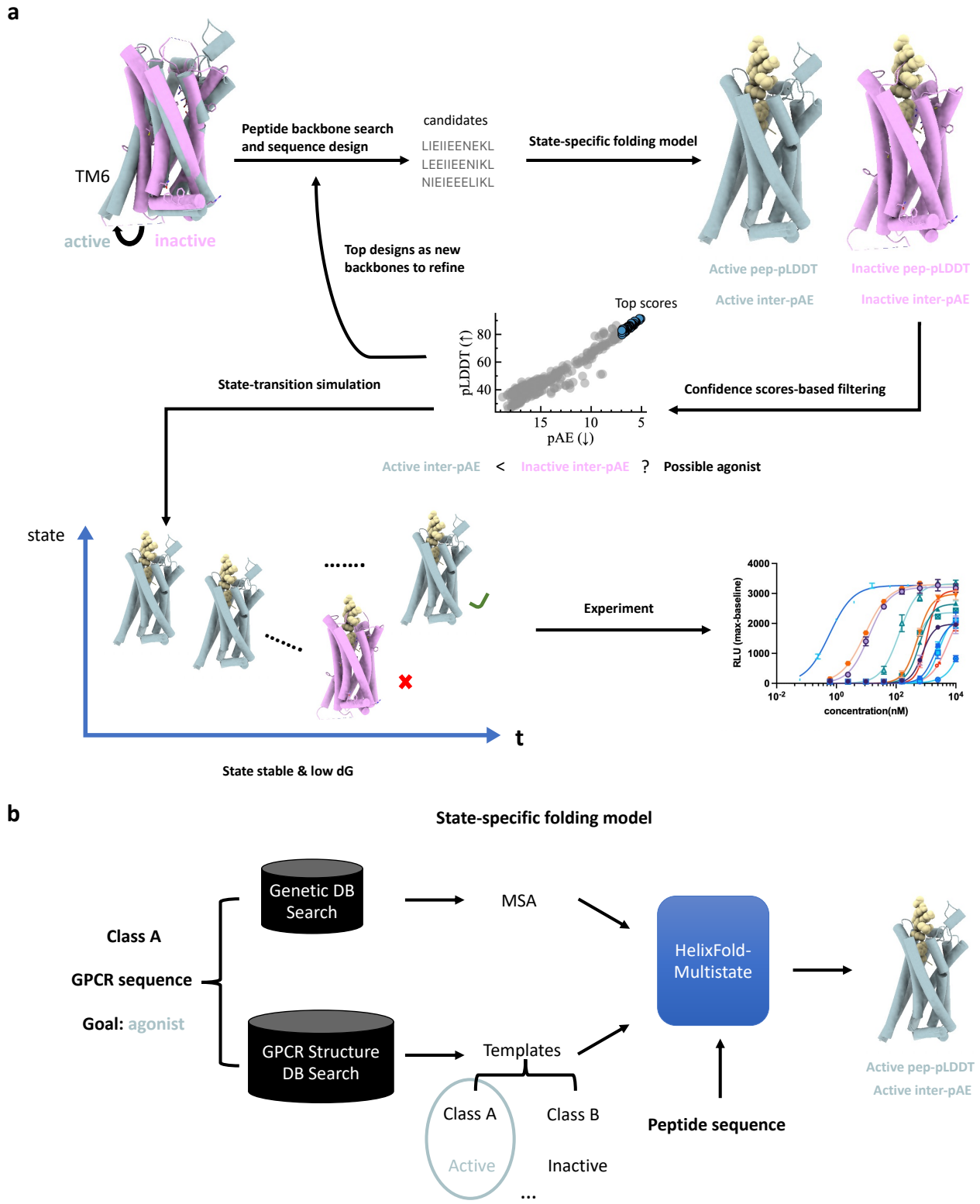

Figure 1. State-specific peptide design pipeline.
(a) Overview of pipeline. **(b)** HelixFold-Multistate, an optimized HelixFold-Multimer model for predicting state-specific GHCR-peptide complex and more correlated confidence scores interaction pAE and peptide pLDDT.

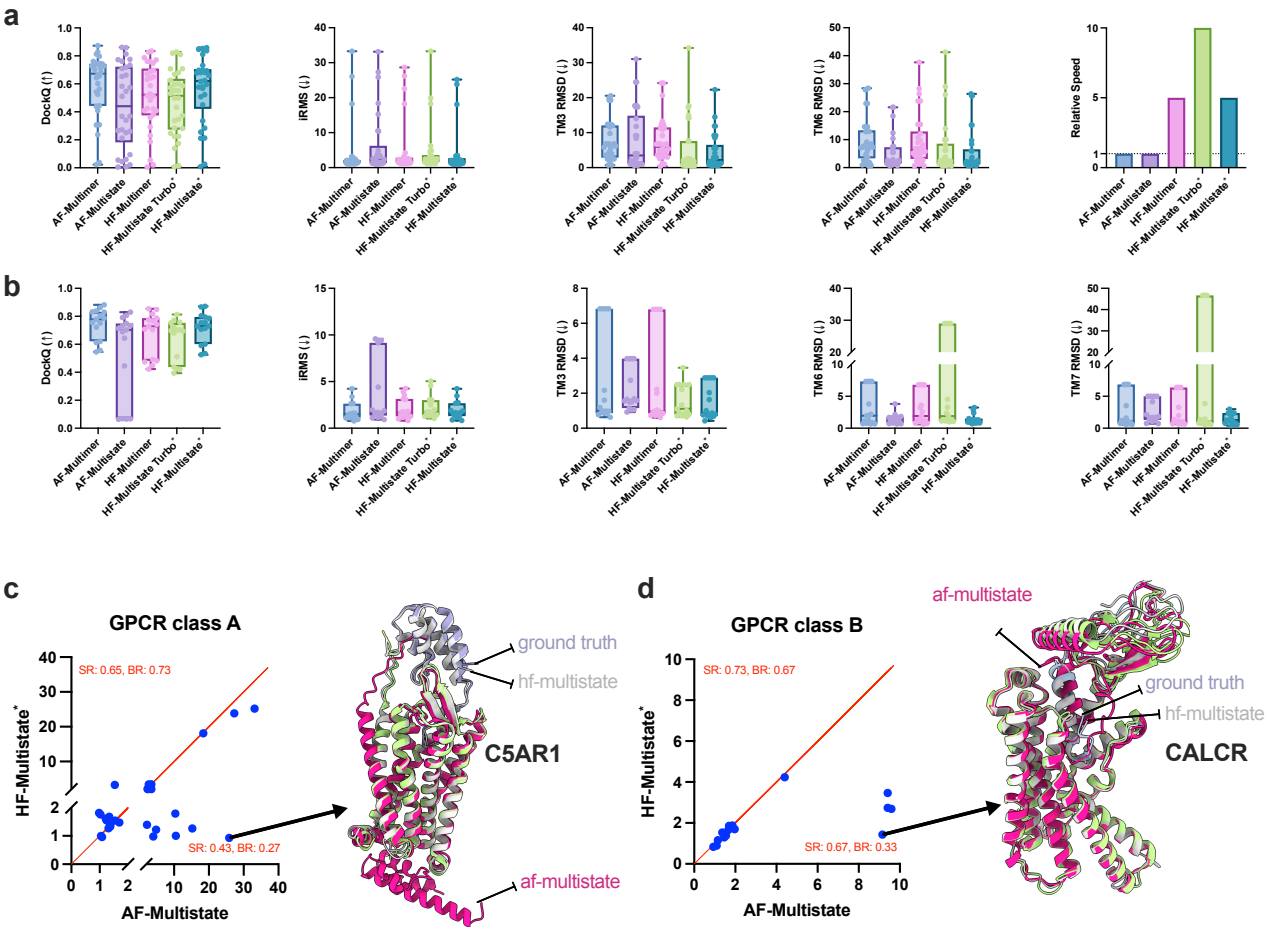

Figure 2. Multimer models performance on recent GPCR class A-peptide and class B-peptide PDB data. **(a)** The box and whiskers plots show DockQ (↑), iRMS (↓), TM3 RMSD (↓), TM6 RMSD (↓) scores distributions for multimer models on recent GPCR class A-peptide PDBs: AF-Multimer, AF-Multistate, HF-Multimer, HF-Multistate Turbo (our), and HF-Multistate (our). The last bar plot shows relative speed (number of AlphaFold network forwards) of five methods. **(b)** The plots show the related metrics on recent GPCR class B-peptide PDBs. **(c)** and **(d)** The scatter plot compares iRMS between AF-Multistate and HF-Multistate (our) on GPCR class A and class B respectively. In the plot, SR (success rate) is the ratio of predictions where iRMS is less than 2Å compared to the experimental PDB. BR (beat rate) is the ratio of one model beats another model in terms of lower iRMS. GPCRs C5AR1 (PDB: **8HK5**) and CALCR (PDB: **8F0K**) are taken as examples to illustrate the different predictions. The N-terminal of the peptide is indicated and labeled. The GPCR protein is shown in light lime green. The peptide in the experimental PDB is shown in pale sky blue. The predicted structure of HF-Multistate is in light gray. The predicted structure of AF-Multistate is in deep pink.

In-Silico Evaluation

We collected recent GPCR-peptide PDBs from GPCRdb (Pándy-Szekeres et al., 2023) as evaluation data, with a release date cutoff of June 30, 2022. We excluded all PDBs from the training data where the GPCR target exists in this set. After this curation process, we obtained 45 PDBs, with 30 belonging to class A and the remaining to class B. In training data, there are 153 GPCR-peptide complex PDBs left. Notably, Data for antagonist peptides is extremely scarce, with only one entry available in GPCRdb. To alleviate the limitation of scarce GPCR-peptide data, we incoporated 7.9k protein-peptide data from Propedia database (Martins et al., 2021), and weight the sampling of GPCR-peptide to balance the data.

To evaluate the accuracy of GPCR-peptide complex structure prediction model, we first utilize the conventional metrics like DockQ and iRMS (Evans et al., 2021; Fang et al., 2024). However, the three components of DockQ (Basu and Wallner, 2016) score mainly focus on the interaction interface, and the activation state of the GPCR protein cannot be assessed. Thus, we use the RMSD of TM domain (Heo and Feig, 2022) as a supplement to assess the GPCR state. Furthermore, for DockQ, 0.49 to 0.80 signifies medium quality; scores of 0.80 or higher reflect high-quality predictions, closely matching the actual complex structure. For iRMS, we take 2Å as a threshold as it is

close to the PDB resolution, indicating an experimental level structure prediction.

The baselines we tested include canonical methods AlphaFold-Multimer (Evans et al., 2021) (AF-Multimer), AlphaFold-Multistate (Heo and Feig, 2022) (AF-Multistate). Since our proposed model, HelixFold-Multistate (HF-Multistate), is finetuned from HelixFold-Multimer (Fang et al., 2024) (HF-Multimer), we compare with it as well. Besides, inspired by the model speed-up idea from EvoPlay (Wang et al., 2023), we also finetune a lighter version of HF-Multistate, named HF-Multistate Trubo, which only executes single recycle forward. Another major difference is that both HF-Multimer and HF-Multistate do not use model ensembles. In contrast, AF-Multimer and AF-Multistate use 5 models, each running 5 times with different seeds, while HF-Multimer and HF-Multistate use a single model running 5 times. Therefore, the relative speed up compared to AF-Multimer is $\times 5$ or $\times 10$ as illustrated in Figure 2a.

Class A. As shown in Figure 2a, the distribution of DockQ scores for our proposed model, HF-Multistate, closely matches that of AF-Multimer and outperforms its initial model, HF-Multimer. Furthermore, while the AF-Multistate model can accurately predict the activation state of monomer GPCR proteins as reported in its paper (Heo and Feig, 2022), we demonstrate that it struggles with poor interaction predictions on GPCR-peptide. In statistics (Figure S1a), the percentage of high-quality predictions (DockQ > 0.8) of HF-Multistate beats all baselines, and the percentage on medium-quality (DockQ > 0.49) is on par with AF-Multimer. On the iRMS metric, HF-Multistate is comparable with AF-Multimer as well. Notably, the turbo model performs slightly better than AF-Multistate while achieving over 10 times the speed. This makes it highly suitable for large peptide library screening. In the aspect of GPCR activation state, we use TM3 RMSD and TM6 RMSD for class A according to existing analysis works (Hauser et al., 2021). We propose that there may be better metrics than RMSD to evaluate the movement of TM3 and TM6 domains, as the movements of these two domains are inherently different. Specifically, the TM6 domain can trigger a large degree of swing from inactive to active state. While the TM3 domain only does tilt and rotation during the transition. This direction should be investigated in our future work. As a result in Figure 2a, the multistate models designed specially for different GPCR states, does lead to lower TM RMSD, especially in the term of TM6 RMSD. Moreover, HF-Multistate surpasses AF-Multistate with a lower TM RMSD medium. The scatter plots in Figure S2 the comparison details for each test data, in which the **success rate** (SR) is ratio of predictions where TM RMSD is less than 2Å, and **beat rate** (BR) is ratio of predictions of one model beats another in terms of lower TM RMSD.

Class B. HF-Multistate demonstrates similar performance on GPCR class B as it does on GPCR class A in terms of DockQ and iRMS, as shown in Figure 2b. Furthermore, to assess the accuracy of predicting GPCR activation states, we analyzed the RMSDs of TM3, TM6, and TM7 following the guidelines of a prior study (Hauser et al., 2021). The results in Figure 2b indicate that HF-Multistate achieves the lowest TM RMSD. However, HF-Multistate Turbo model failed to fold GPCRs in some test examples, highlighting the importance of multiple recycle forwards in accurately modeling the swing movements of TM6 and TM7. Figure S3 compares HF-Multistate with other baselines on each individual test data in terms of TM RMSD. HF-Multistate consistently achieves a higher beat rate; however, at the success rate, the margin is not as significant as it is for GPCR class A. We argue that this may be due to the smaller amount of training data available for GPCR class B.

Since the multimer models cannot accurately predict the activation states of GPCRs, we focus on comparing HF-Multistate with AF-Multistate. In Figure 2(c,d), the iRMS scores of the two models are presented as scatter points. As shown, the majority of the scatter points fall below the $y = x$ line, indicating that HF-Multistate achieves overall lower iRMS scores than AF-Multistate. Specifically, the beat rates are 0.73 and 0.67 for GPCR class A and class B, respectively. Moreover, the success rates of HF-Multistate are also higher than those of AF-Multistate. We use C5AR1 and CALCR as examples to compare the failures of AF-Multistate that HF-Multistate successfully addresses. For C5AR1, AF-Multistate is error-prone as it predicts the wrong binding pocket intended for G proteins. For CALCR, while AF-Multistate identifies the correct pocket, it incorrectly orients the peptide, causing the N-terminal to face upwards.

To evaluate whether HF-Multistate scores are valid for peptide coarse filtering, we use the CXCR4-peptide affinity data curated in PPI-Affinity (Romero-Molina et al., 2022) to do another experiment against AF-Multistate. The results are shown in Figure 3. In details, we use average peptide pLDDT (pLLDT) and average GPCR-peptide pAE (pAE) as scores to filter peptide candidates. According to the physical meaning of pLDDT and pAE in AlphaFold-Multimer (Evans et al., 2021), the larger pLLDT indicates more plausible peptide sequence, maybe more nature and stable; the smaller pAE means higher confidence of binding since the alignment distances are more trustable. Moreover, we tested the multistate models using both active and inactive state features to check if the pAE and pLDDT scores differences of two states can help tell the real peptide functionality. Ideally, if the peptide is an inhibitor and the model is capable to distinguish active and inactive states, its ΔpAE ($\text{pAE}_{\text{active}} - \text{pAE}_{\text{inactive}}$) and ΔpLDDT ($\text{pLDDT}_{\text{inactive}} - \text{pLDDT}_{\text{active}}$) should be positive. In the ΔpAE-affinity plot of Figure 3e and ΔpLDDT-affinity

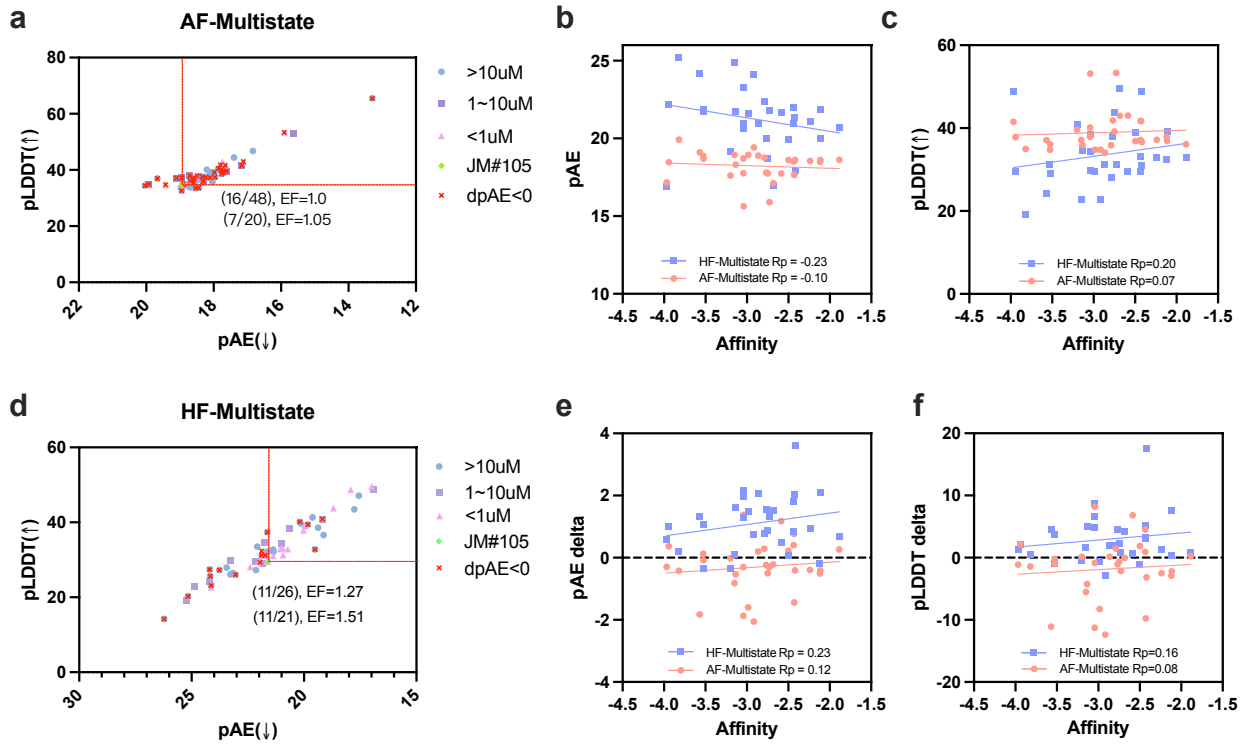

Figure 3. Comparing AF-Multistate and HF-Multistate on CXCR4-peptide affinity data.
We conduct coarse filtering using pLDDT and pAE scores of inhibitor peptides from the multistate model. Specifically, lower pAE scores and higher pLDDT scores are preferred. The values $\Delta\text{pAE} = \text{pAE}_{\text{active}} - \text{pAE}_{\text{inactive}}$ and $\Delta\text{pLDDT} = \text{pLDDT}_{\text{inactive}} - \text{pLDDT}_{\text{active}}$ reflect the confidence that a peptide acts as an inhibitor. Panels (**a, d**) display pAE and pLDDT scores from AF-Multistate and HF-Multistate, respectively. The green dot represents the JM#105 peptide, which has a Kd of 966 nM. The upper right area, delineated by red lines, contains peptides that the model scores higher than JM#105. In this region, the pink triangles (Kd < 1 μM) indicate correct predictions, while the red cross marker denotes samples predicted not to be inhibitors. When considering only pAE and pLDDT, the enrichment factor (EF) of HF-Multistate is 1.27, compared to AF-Multistate's EF of 1.0. After filtering out peptides marked as not inhibitors, the EF of HF-Multistate increases to 1.51, while AF-Multistate's EF is 1.05. Panels (**b, c, e, f**) present the Pearson correlation analysis of different multistate scores with peptide affinity. Notably, the ΔpAE and ΔpLDDT values (most > 0) for HF-Multistate indicate a tendency for peptides to act as inhibitors, whereas AF-Multistate does not demonstrate this trend.

plot of Figure 3f, we can see HF-Multisate scores follows this tendency and predict these peptides as inhibitor, while AF-Multistate cannot. In all Pearson correlation analysis (Figure 3(b,c,e,f)), HF-Multisate scores continue outperform AF-Multisate, and the coarse filtering process in Figure 3(a,d), also show that HF-Multisate achieves better **enrichment factor** (EF) when try to screen inhibitor peptides with affinity better than JM#105 peptide. In addition, when applying non-inhibitor peptide filter $\Delta\text{pAE} < 0$, EF of HF-Multisate can further improve to 1.51 from 1.27. However, EF of AF-Multistate only has subtle improvement from 1.0 to 1.05.

Results on GHSR

In this work, we first tested our pipeline on Growth Hormone Secretagogue Receptor (GHSR), which is a G-protein-coupled receptor that playing a crucial role in various physiological processes, primarily related to growth hormone release and appetite regulation (Kojima et al., 2001). Since the mechanism of GHSR is relatively clear and both agonist Ghrelin (Qin et al., 2022) and antagoinst LEAP2 (M'Kadmi et al., 2018) are known, we conduct peptide design using these reference peptides and take mechanism as the knowledge behind structural dynamics evaluation. Specifically, in GHSR, there is a WFF cluster ($\text{W}276^{6.48}$, $\text{F}279^{6.51}$, $\text{F}312^{7.42}$) (green in Figure 4c) exhibiting obvious conformational displacements upon ghrelin (light brown in Figure 4c) binding. This is because of the strong hydrophobic interaction between N-terminal of ghrelin and the WFF cluster. Meanwhile, the rearrangle of cascade motifs (purple in Figure 4c) leads to the expanding the intracellular end of TM6. After analysis this mechanism, we design four measurements that can be used to evaluate the state transition during molecular

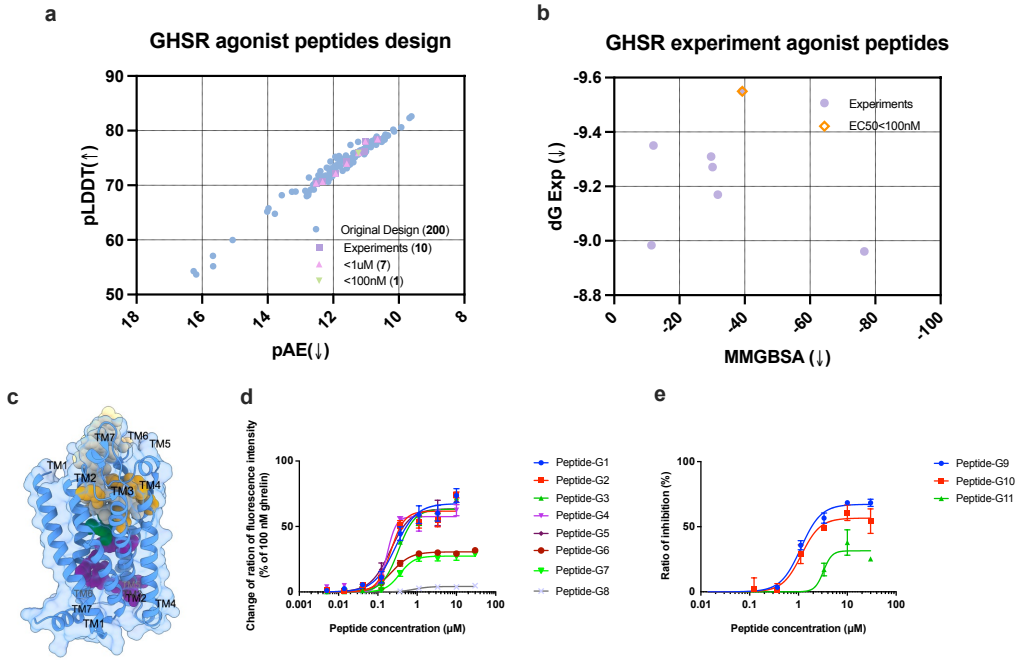

Figure 4. Peptides filtering for GHSR and affinity test results
For GHSR, we utilize state-specific scores from HF-Multistate, such as pAE interaction and pLDDT of peptides, to filter candidates for MMGBSA simulation. MMGBSA provides the ranking to select peptides for synthesis and affinity testing. The designed peptides, experimental peptides, and high-affinity peptides are highlighted using different colors.

dynamics simulation:

- dw: width of WFF cluster.
- TM6 RMSD: TM6 RMSD against active (RMSD_a) and inactive states (RMSD_i).
- d: the distance of intracellular end of TM6 towards active (d_a) and inactive (d_i) states.
- θ: the swing angle of intracellular end of TM6 from active (d_a) and inactive (d_i) states.

Theoretically, if the state of GHSR state transition happends during simulation, it indicates the predicted GHSR-peptide structure is unstable, and the peptide either has reverse functionality or doesn't have affinity. Before conducting such simulation, we first analysis the influence of membrane on these measurements. As a results in Figure S4, the membrane does not have effect on the change direction, but on the amplitude of the change. Considering the large computation resources increase when simulate with membrane, we remove it then do the simulation for 10ns for each peptide design.

In addition to structural dynamics evaluation, we also applied confidence score analysis and MMGBSA free energy analysis. The filtering process and experimental results are presented in Figure 4b. For brevity, we only demonstrate the filtering process for agonist peptides here. Most of the peptides selected for experiments demonstrated EC50 <1μM, with peptide G4 showing EC50 <100nM. Notably, we achieved these results with a design pool of only 200 peptides, using just two reference backbones and some fixed key interaction sites extracted by PLIP tool. Moreover, three inhibitor peptides are found in our experimental peptides.

Results on APJ

To test the generalizability of our pipeline on peptide designs targeting GPCRs, we conducted an experiment on a new target, the Apelin Receptor (APJ). APJ is a multifunctional GPCR that plays significant roles in cardiovascular health, fluid balance, metabolism, and angiogenesis (O'Carroll et al., 2013). Its broad physiological impact and potential for therapeutic targeting make it a protein of considerable interest in biomedical research, aligning with our research objectives. Unlike GHSR, we do not know the detailed activation mechanism for APJ. Therefore, we treat

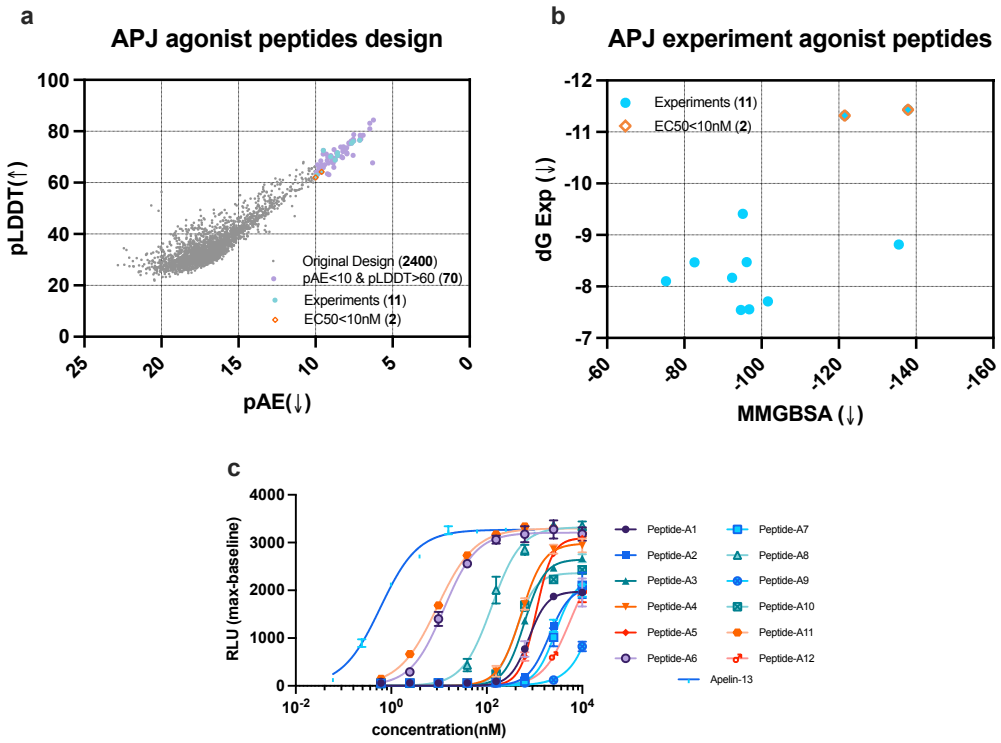

Figure 5. Peptides filtering for APJ and affinity test results
Multistate model scores (pAE interaction (pAE) and pLDDT of peptide (pLDDT)) distribution on original designs, coarse filtered, experimental and high-affinity peptides. Comparing to GHSR (Figure 4), we use de-novo design and generated a much larger pool for screening.

it as a de novo design problem. Specifically, we use FoldSeek (Van Kempen et al., 2024) to find candidate peptide backbones by searching for peptides and protein fragments that interact with a protein having a similar structure to APJ. For each backbone, we apply the generation pipeline used in the GHSR experiments (Figure 1a). Because the initial backbones are much larger, we can design a more diverse set of peptides, totaling 2400, which is over 10 times the number designed for GHSR. Then we evaluate them using HF-Multistate. The scores distribution is in Figure 5a. Next 70 peptides with pLDDT > 60 and pAE < 10 are selected for MMGBSA simulation. Top-ranked peptides are then selected for synthesis and affinity testing. The results in Figure 5b illustrate that for experimental peptides with affinity, MMGBSA can reflect the real affinity to some extent. Peptide-A6 and peptide-A11 show good EC50 <10nM (Figure 5c).

Discussion

The development of state-specific peptides targeting G protein-coupled receptors (GPCRs) represents a significant advancement in the field of drug discovery and peptide therapeutics. GPCRs, as key mediators of cellular signaling, have long been prime targets for pharmaceutical interventions. Recent successes in GPCR-targeted peptide therapeutics underscore the potential of this approach. For instance, the FDA approval of semaglutide, a GLP-1 receptor agonist peptide for type 2 diabetes and obesity treatment, has demonstrated the clinical efficacy of peptides in modulating GPCR activity (Knudsen and Lau, 2019; Bucheit et al., 2020). Similarly, the development of dasiglucagon, a glucagon receptor agonist peptide for severe hypoglycemia, showcases the potential of engineered peptides to improve upon natural ligands in terms of stability and efficacy (Pieber et al., 2021). However, these successes also illuminate the challenges in peptide design, particularly in accounting for the dynamic nature of GPCRs. Traditional peptide design methods often fail to consider the multiple conformational states of GPCRs, which can significantly influence ligand binding and receptor activation. Our computational pipeline addresses this critical gap by explicitly incorporating GPCR activation state into the peptide design process. By leveraging diversified seed peptide backbones, inverse folding models for sequence design, and state-specific refolding and scoring for filtering, our approach offers a more nuanced and potentially more effective method for designing GPCR-targeting peptides.

The key components of our computational pipeline address the multifaceted challenges of state-specific peptide design for GPCRs. By utilizing diversified seed peptides as demonstrated in the APJ experiment, we expand the initial search space, allowing for the exploration of a broader range of potential ligands. This approach is particularly valuable given the complex conformational landscape of GPCRs and the unclear structure mechanism. The incorporation of an inverse folding model for peptide design represents a significant advancement over traditional sequence-based methods. This model allows us to generate peptide sequences that are more likely to adopt the desired three-dimensional structure for optimal GPCR binding. Crucially, our pipeline integrates an optimized multistate model, HelixFold-Multistate, which can predict state-specific complex structures more accurately and provide more correlated confidence scores than AlphaFold-Multistate in a significantly faster manner. Furthermore, our sophisticated peptide evaluation and filtering process dramatically decreases the size of the design pool required to find high-affinity peptides targeting GPCRs. The culmination of these components is the evaluation of state-specific scores for peptide designs, providing a comprehensive assessment of a peptide's potential to interact with a GPCR in its various conformational states. This multi-faceted approach allows for the design of peptides that are not only likely to bind to the target GPCR but also to do so in a state-specific manner, potentially leading to more selective and efficacious therapeutic agents.

The experimental validation of our computational pipeline on two distinct GPCRs, the Growth Hormone Secretagogue Receptor (GHSR) and the Apelin Receptor (APJ), demonstrates its effectiveness and versatility. Our approach successfully identified agonist peptides for both receptors with remarkable affinity, achieving binding constants below 100nM. This high affinity is particularly noteworthy given the relatively small number of in silico sequences (200-3000) generated for each receptor. When the GPCR state-transition mechanisms are well studied, our pipeline applys it to filter peptide candidates, which dramatically reduces the size of the design pool. This targeted approach allows us to achieve remarkable results with an exceptionally small number of in silico sequences - in some cases, fewer than 200 per receptor like GHSR. The efficiency of our pipeline is further highlighted by the fact that fewer than 20 peptides needed to be synthesized and tested experimentally to achieve these results. This represents a significant reduction in the time and resources typically required for peptide drug discovery. Moreover, the designed peptides showed less than 50% sequence identity to the reference peptides, indicating that our approach can explore novel regions of the peptide sequence space while maintaining high binding affinity and function. When using diverse peptide seeds backbones, like APJ, our pipeline can explore larger design space and find high-affinity below 10nM. This ability to generate diverse, high-affinity peptides from a small initial pool provides numerous advantages in peptide drug development. It offers more options for further modification to increase potency and half-life, which is particularly crucial in peptide therapeutics where optimizing pharmacokinetic properties is often a key challenge. The pipeline's performance in designing antagonist peptides, while more challenging, still yielded promising results. The identification of a GHSR inhibitor with IC50 3.3μM, an APJ inhibitor with 20.3μM, although modest compared to the agonist results, demonstrates the potential of our approach in addressing the more complex task of antagonist design. These results collectively validate the effectiveness and efficiency of our state-specific peptide design pipeline, showcasing its potential to significantly accelerate the discovery of novel GPCR-targeting peptide therapeutics with desirable properties for drug development.

Our state-specific peptide design approach for GPCRs addresses several limitations of existing methods and offers unique advantages. Unlike conventional approaches that rely on natural peptides, phage display, or single-state computational methods, our pipeline explicitly accounts for the dynamic conformational states of GPCRs. This is a significant advancement over recent rational design methods like EvoBind (Bryant and Elofsson, 2022), EvoPlay (Wang et al., 2023), or PepMLM (Chen et al., 2023), which do not consider GPCR state transitions. While approaches like AlphaFold-Multistate (Heo and Feig, 2022) attempt to address GPCRs activate states, our experiments revealed its limitations in predicting GPCR-peptide complex structures accurately. Our pipeline overcomes this challenge by combining diversified seed peptides, inverse folding models for sequence generation, and a specifically optimized GPCR-peptide structure prediction model. This integrated approach allows us to design unconstrained, state-specific functional peptides, addressing a gap left by methods like HelixDiff (Xie et al., 2024), which focuses solely on structure-stable helix peptides. However, we acknowledge that the performance difference between agonist and antagonist design highlights an area for future refinement, particularly in designing peptides that must prevent receptor activation.

The success of our state-specific peptide design pipeline for GPCRs opens up exciting avenues for future research and development in the field of peptide therapeutics. While our current work has focused primarily on optimizing binding affinity, future iterations of this approach should aim to simultaneously optimize multiple pharmacological properties. Specifically, enhancing potency and extending half-lives are crucial next steps in developing peptide therapeutics with improved clinical efficacy. This could involve incorporating additional scoring functions or filters that account for factors affecting in vivo stability and receptor activation efficiency. Moreover, the versatility of our

approach presents an opportunity to design peptides targeting multiple GPCR targets simultaneously, similar to the multi-target drug tirzepatide. Such multi-target peptides could offer synergistic therapeutic effects and potentially address complex disorders involving multiple GPCR pathways. Another promising direction is the integration of our pipeline with advanced peptide modification techniques, such as cyclization or the incorporation of non-natural amino acids, to further enhance the pharmacokinetic properties of the designed peptides. Additionally, expanding the application of our method to less well-characterized GPCRs could significantly broaden the scope of peptide-based drug discovery. To facilitate this, future work should focus on improving the accuracy of GPCR state prediction for receptors with limited structural data. Lastly, the incorporation of diffusion model techniques to predict all-atoms structures directly leads to fewer MSA requirements (Abramson et al., 2024), and peptides may benefit from this. By addressing these future directions, our state-specific peptide design approach has the potential to revolutionize GPCR-targeted therapeutics, leading to the development of more effective and safer peptide drugs with optimized pharmacological profiles.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, pages 1–3.

Alavi, S. E., Cabot, P. J., and Moyle, P. M. (2019). Glucagon-like peptide-1 receptor agonists and strategies to improve their efficiency. *Molecular pharmaceutics*, 16(6):2278–2295.

Basu, S. and Wallner, B. (2016). Dockq: a quality measure for protein-protein docking models. *PloS one*, 11(8):e0161879.

Brown, M. C., Calvete, J. J., Staniszewska, I., Walsh, E. M., Perez-Liz, G., Del Valle, L., Lazarovici, P., and Marcinkiewicz, C. (2007). Vegf-related protein isolated from vipera palestinae venom, promotes angiogenesis. *Growth Factors*, 25(2):108–117.

Bryant, P. and Elofsson, A. (2022). Evobind: in silico directed evolution of peptide binders with alphafold. *bioRxiv*, pages 2022–07.

Bucheit, J. D., Pamulapati, L. G., Carter, N., Malloy, K., Dixon, D. L., and Sisson, E. M. (2020). Oral semaglutide: a review of the first oral glucagon-like peptide 1 receptor agonist. *Diabetes technology & therapeutics*, 22(1):10–18.

Chang, L., Mondal, A., and Perez, A. (2022). Towards rational computational peptide design. *Frontiers in Bioinformatics*, 2: 1046493.

Chen, T., Pertsemlidis, S., Watson, R., Kavirayuni, V. S., Hsu, A., Vure, P., Pulugurta, R., Vincoff, S., Hong, L., Wang, T., et al. (2023). Pepmlm: Target sequence-conditioned generation of peptide binders via masked language modeling. *ArXiv*.

Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R. J., Milles, L. F., Wicky, B. I., Courbet, A., de Haas, R. J., Bethel, N., et al. (2022). Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378(6615):49–56.

Davenport, A. P., Scully, C. C., de Graaf, C., Brown, A. J., and Maguire, J. J. (2020). Advances in therapeutic peptides targeting g protein-coupled receptors. *Nature Reviews Drug Discovery*, 19(6):389–413.

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2021). Protein complex prediction with alphafold-multimer. *biorxiv*, pages 2021–10.

Fang, X., Gao, J., Hu, J., Liu, L., Xue, Y., Zhang, X., and Zhu, K. (2024). Helixfold-multimer: Elevating protein complex structure prediction to new heights. *arXiv preprint arXiv:2404.10260*.

Hauser, A. S., Kooistra, A. J., Munk, C., Heydenreich, F. M., Veprintsev, D. B., Bouvier, M., Babu, M. M., and Gloriam, D. E. (2021). Gpcr activation mechanisms across classes and macro/microscales. *Nature structural & molecular biology*, 28(11): 879–888.

Heo, L. and Feig, M. (2022). Multi-state modeling of g-protein coupled receptors at experimental accuracy. *Proteins: Structure, Function, and Bioinformatics*, 90(11):1873–1885.

Hill, T. A., Shepherd, N. E., Diness, F., and Fairlie, D. P. (2014). Constraining cyclic peptides to mimic protein structure motifs. *Angewandte Chemie International Edition*, 53(48):13020–13041.

Hsu, C., Verkuil, R., Liu, J., Lin, Z., Hie, B., Sercu, T., Lerer, A., and Rives, A. Learning inverse folding from millions of predicted structures. In *International conference on machine learning*, pages 8946–8970. PMLR, (2022).

Knudsen, L. B. and Lau, J. (2019). The discovery and development of liraglutide and semaglutide. *Frontiers in endocrinology*, 10:155.

Kojima, M., Hosoda, H., Matsuo, H., and Kangawa, K. (2001). Ghrelin: discovery of the natural endogenous ligand for the growth hormone secretagogue receptor. *Trends in Endocrinology & Metabolism*, 12(3):118–122.

Li, Q., Vlachos, E. N., and Bryant, P. (2024). Design of linear and cyclic peptide binders of different lengths only from a protein target sequence. *bioRxiv*, pages 2024–06.

Martins, P. M., Santos, L. H., Mariano, D., Queiroz, F. C., Bastos, L. L., Gomes, I. d. S., Fischer, P. H., Rocha, R. E., Silveira, S. A., de Lima, L. H., et al. (2021). Propedia: a database for protein–peptide identification based on a hybrid clustering algorithm. *BMC bioinformatics*, 22:1–20.

Matsubayashi, Y. and Sakagami, Y. (2006). Peptide hormones in plants. *Annu. Rev. Plant Biol.*, 57:649–674.

Mygind, P. H., Fischer, R. L., Schnorr, K. M., Hansen, M. T., Sönksen, C. P., Ludvigsen, S., Raventós, D., Buskov, S., Christensen, B., De Maria, L., et al. (2005). Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. *Nature*, 437(7061):975–980.

M'Kadmi, C., Cabral, A., Barrile, F., Giribaldi, J., Cantel, S., Damian, M., Mary, S., Denoyelle, S., Dutertre, S., Péraldi-Roux, S., et al. (2018). N-terminal liver-expressed antimicrobial peptide 2 (leap2) region exhibits inverse agonist activity toward the ghrelin receptor. *Journal of Medicinal Chemistry*, 62(2):965–973.

Nevola, L. and Giralt, E. (2015). Modulating protein–protein interactions: the potential of peptides. *Chemical Communications*, 51(16):3302–3315.

Omidfar, K. and Daneshpour, M. (2015). Advances in phage display technology for drug discovery. *Expert opinion on drug discovery*, 10(6):651–669.

O'Carroll, A.-M., Lolait, S. J., Harris, L. E., and Pope, G. R. (2013). The apelin receptor apj: journey from an orphan to a multifaceted regulator of homeostasis. *J Endocrinol*, 219(1):R13–35.

Pacesa, M., Nickel, L., Schmidt, J., Pyatova, E., Schellhaas, C., Kissling, L., Alcaraz-Serna, A., Cho, Y., Ghamary, K. H., Vinue, L., et al. (2024). Bindcraft: one-shot design of functional protein binders. *bioRxiv*, pages 2024–09.

Pándy-Szekeres, G., Caroli, J., Mamyrbekov, A., Kermani, A. A., Keserű, G. M., Kooistra, A. J., and Gloriam, D. E. (2023). Gpcrdb in 2023: state-specific structure models using alphafold2 and new ligand resources. *Nucleic acids research*, 51(D1): D395–D402.

Pieber, T. R., Aronson, R., Hövelmann, U., Willard, J., Plum-Mörschel, L., Knudsen, K. M., Bandak, B., and Tehranchi, R. (2021). Dasiglucagon—a next-generation glucagon analog for rapid and effective treatment of severe hypoglycemia: results of phase 3 randomized double-blind clinical trial. *Diabetes Care*, 44(6):1361–1367.

Pokuri, V. K., Fong, M. K., and Iyer, R. (2016). Octreotide and lanreotide in gastroenteropancreatic neuroendocrine tumors. *Current oncology reports*, 18:1–9.

Qin, J., Cai, Y., Xu, Z., Ming, Q., Ji, S.-Y., Wu, C., Zhang, H., Mao, C., Shen, D.-D., Hirata, K., et al. (2022). Molecular mechanism of agonism and inverse agonism in ghrelin receptor. *Nature Communications*, 13(1):300.

Romero-Molina, S., Ruiz-Blanco, Y. B., Mieres-Perez, J., Harms, M., Munch, J., Ehrmann, M., and Sanchez-Garcia, E. (2022). Ppi-affinity: A web tool for the prediction and optimization of protein–peptide and protein–protein binding affinity. *Journal of proteome research*, 21(8):1829–1841.

Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I., et al. (2017). A comprehensive map of molecular drug targets. *Nature reviews Drug discovery*, 16(1):19–34.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489.

Toivanen, P. I., Nieminen, T., Laakkonen, J. P., Heikura, T., Kaikkonen, M. U., and Ylä-Herttuala, S. (2017). Snake venom vegf vammin induces a highly efficient angiogenic response in skeletal muscle via vegfr-2/nrp specific signaling. *Scientific reports*, 7(1):5525.

Van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with foldseek. *Nature biotechnology*, 42(2):243–246.

Vázquez Torres, S., Leung, P. J., Venkatesh, P., Lutz, I. D., Hink, F., Huynh, H.-H., Becker, J., Yeh, A. H.-W., Juergens, D., Bennett, N. R., et al. (2024). De novo design of high-affinity binders of bioactive helical peptides. *Nature*, 626(7998):435–442.

Wang, L., Wang, N., Zhang, W., Cheng, X., Yan, Z., Shao, G., Wang, X., Wang, R., and Fu, C. (2022). Therapeutic peptides: current applications and future directions. *Signal Transduction and Targeted Therapy*, 7(1):48.

Wang, Y., Tang, H., Huang, L., Pan, L., Yang, L., Yang, H., Mu, F., and Yang, M. (2023). Self-play reinforcement learning guides protein engineering. *Nature Machine Intelligence*, 5(8):845–860.

Weiss, D. R., Karpiak, J., Huang, X.-P., Sassano, M. F., Lyu, J., Roth, B. L., and Shoichet, B. K. (2018). Selectivity challenges in docking screens for gpcr targets and antitargets. *Journal of medicinal chemistry*, 61(15):6830–6845.

Wilson, A. C., Vadakkadath Meethal, S., Bowen, R. L., and Atwood, C. S. (2007). Leuprolide acetate: a drug of diverse clinical applications. *Expert opinion on investigational drugs*, 16(11):1851–1863.

Xie, X., Valiente, P. A., Kim, J., and Kim, P. M. (2024). Helixdiff, a score-based diffusion model for generating all-atom α-helical structures. *ACS Central Science*.

Yamazaki, Y., Matsunaga, Y., Tokunaga, Y., Obayashi, S., Saito, M., and Morita, T. (2009). Snake venom vascular endothelial growth factors (vegf-fs) exclusively vary their structures and functions among species. *Journal of biological chemistry*, 284(15): 9885–9891.

Zambaldi, V., La, D., Chu, A. E., Patani, H., Danson, A. E., Kwan, T. O., Frerix, T., Schneider, R. G., Saxton, D., Thillaisundaram, A., et al. (2024). De novo design of high-affinity protein binders with alphaproteo. *arXiv preprint arXiv:2409.08022*.

Supplementary Information

A. HelixFold-Multistate

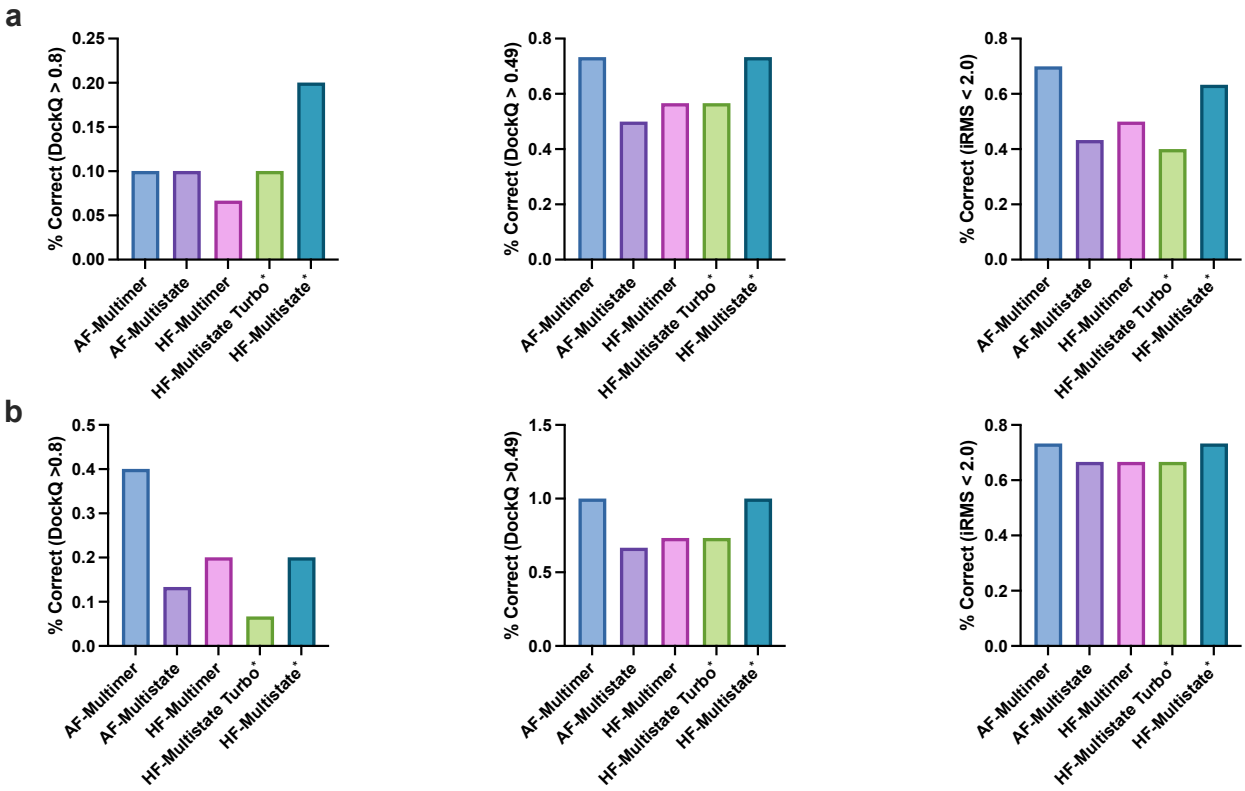

Figure S1. Comparing activation state prediction performance for GPCR class A on TM3 and TM6 domains. The bar plot shows the percentage of accuracy under different metrics: DockQ > 0.8, DockQ > 0.49, and iRMS < 2.0Å on recent GPCR class A-peptide **(a)** and GPCR class B-peptide **(b)** PDBs.

B. Molecular Dynamic (MD) simulation settings

All molecular dynamics simulations were conducted using GROMACS (Version 2021) with the Amber03 force field for modeling the interactions within the protein-peptide complex. The system was immersed in a rectangular water box with dimensions approximately 8.5×9.5×10.5 nm³ and periodic boundary conditions were applied, resulting in a total system size of about 83,000 atoms comprising a single protein-peptide molecule and around 25,000 water molecules. To neutralize the overall charge, sodium ions were introduced into the system. The system underwent an energy minimization process through steepest descent steps, with a maximum force of 10.0 kJ·mol^{-1}·nm^{-1} and up to 5,000 steps. Hydrogen-containing bonds were constrained using the LINCS algorithm to enable a time step of 2 fs. Van der Waals interactions were truncated at 1.0 nm, with a gradual switch to zero from 1.0 nm to 1.2 nm for the Lennard-Jones potential. Short-range electrostatic interactions within a cutoff distance of rc = 10 Å were treated as Coulombic interactions. Long-range electrostatic interactions were managed using the particle mesh Ewald (PME) algorithm. The simulations were performed in the NVT ensemble with a Noṡe-Hoover thermostat at 300 K. Pressure coupling was achieved using the Parrinello-Rahman algorithm with a coupling time of τ = 2 ps. Subsequently, the system was equilibrated in the NVT ensemble at 300 K for 100 ps, followed by equilibration in the NPT ensemble at a pressure of 1 bar for an additional 100 ps. The final structure obtained from the NPT equilibration was used as the initial structure for independent molecular dynamics simulations, each lasting 100 ns. Subsequently, the last 500 frames were collected for each system at 2 ps intervals for subsequent MM/GBSA calculations. In this research, the MM/GBSA calculations utilized one GB models (GB-OBC1). we noted that the choice of interior (solute) dielectric constant could have a significant impact on the predictive accuracy of MM/GBSA. Following the MM/GBSA benchmark results of protein-peptide system, interior dielectric constants 1 were employed for these calculations. The exterior (solvent) dielectric constant was maintained at 80.

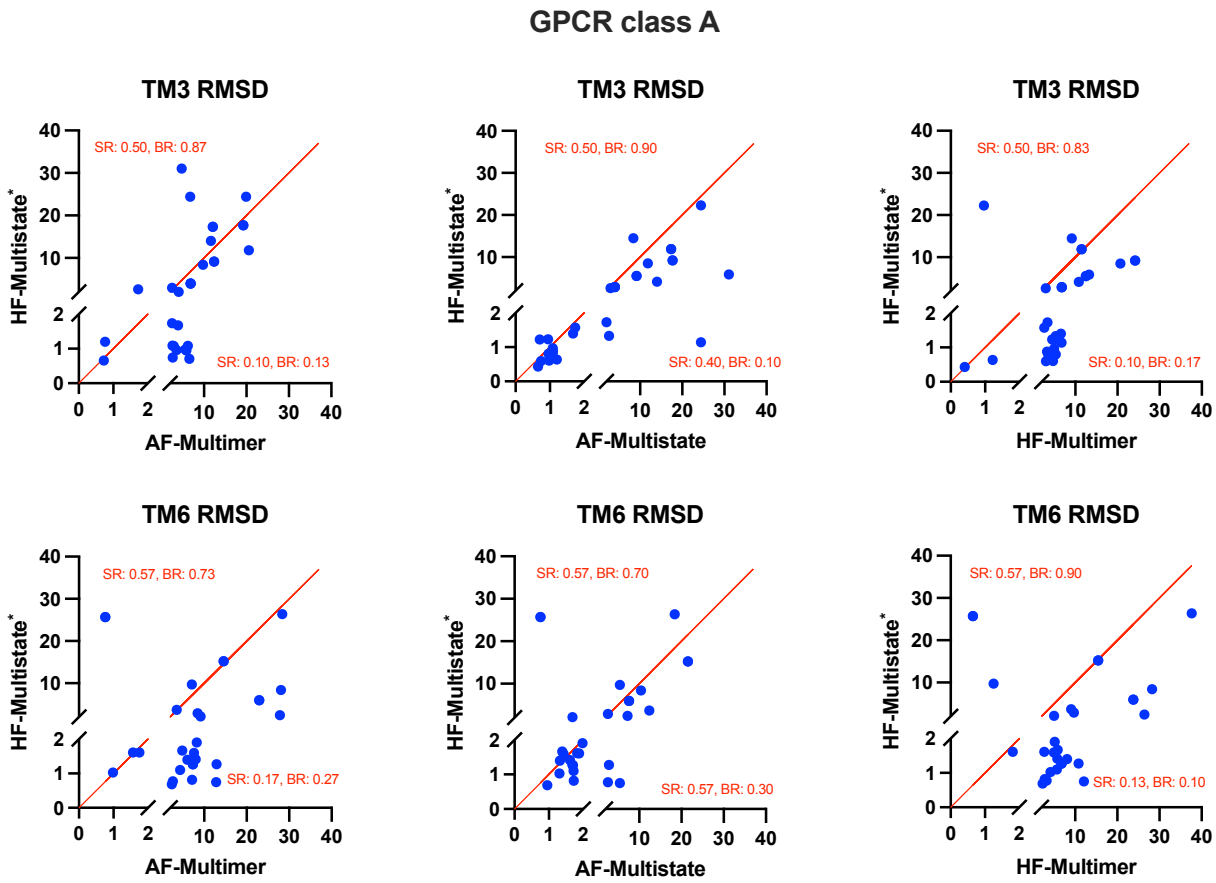

Figure S2. Comparing activation state prediction performance for GPCR class A on TM3 and TM6 domains.
For GPCR class A proteins, their activation states are mainly determined by TM3 and TM6 domains. The corresponding RMSD metrics reflect that our proposed HF-Multistate model outperforms baselines in terms of SR (success rate) and BR (beat rate).

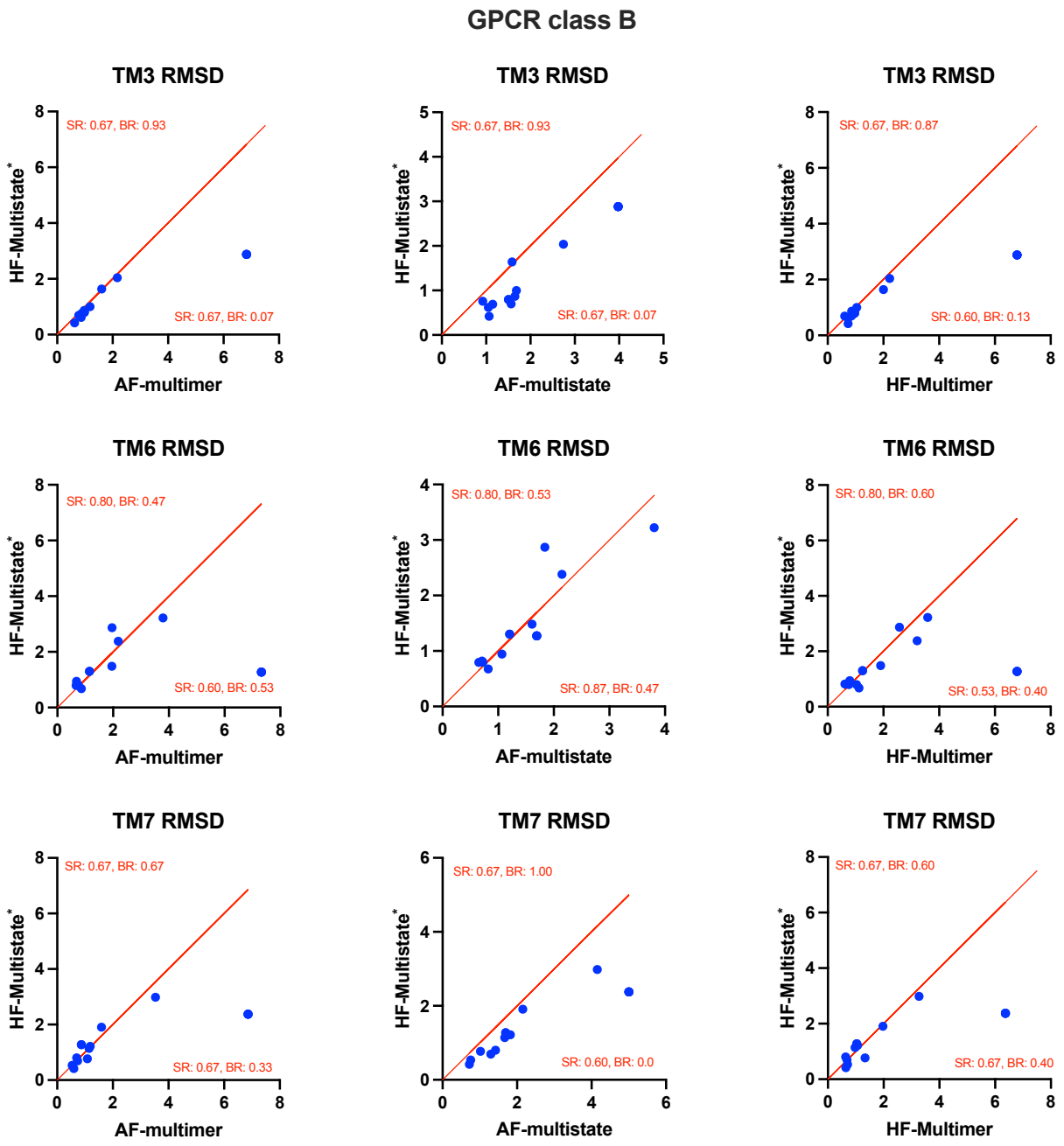

Figure S3. Comparing activation state prediction performance for GPCR class B on TM3, TM6 and TM7 domains. For GPCR class B proteins, their activation states are mainly determined by TM3, TM6 and TM7 domains. Unlike GPCR class A (Figure S2), there is less PDB data for class B, and the margin between HF-Multistate and baseline models is smaller.

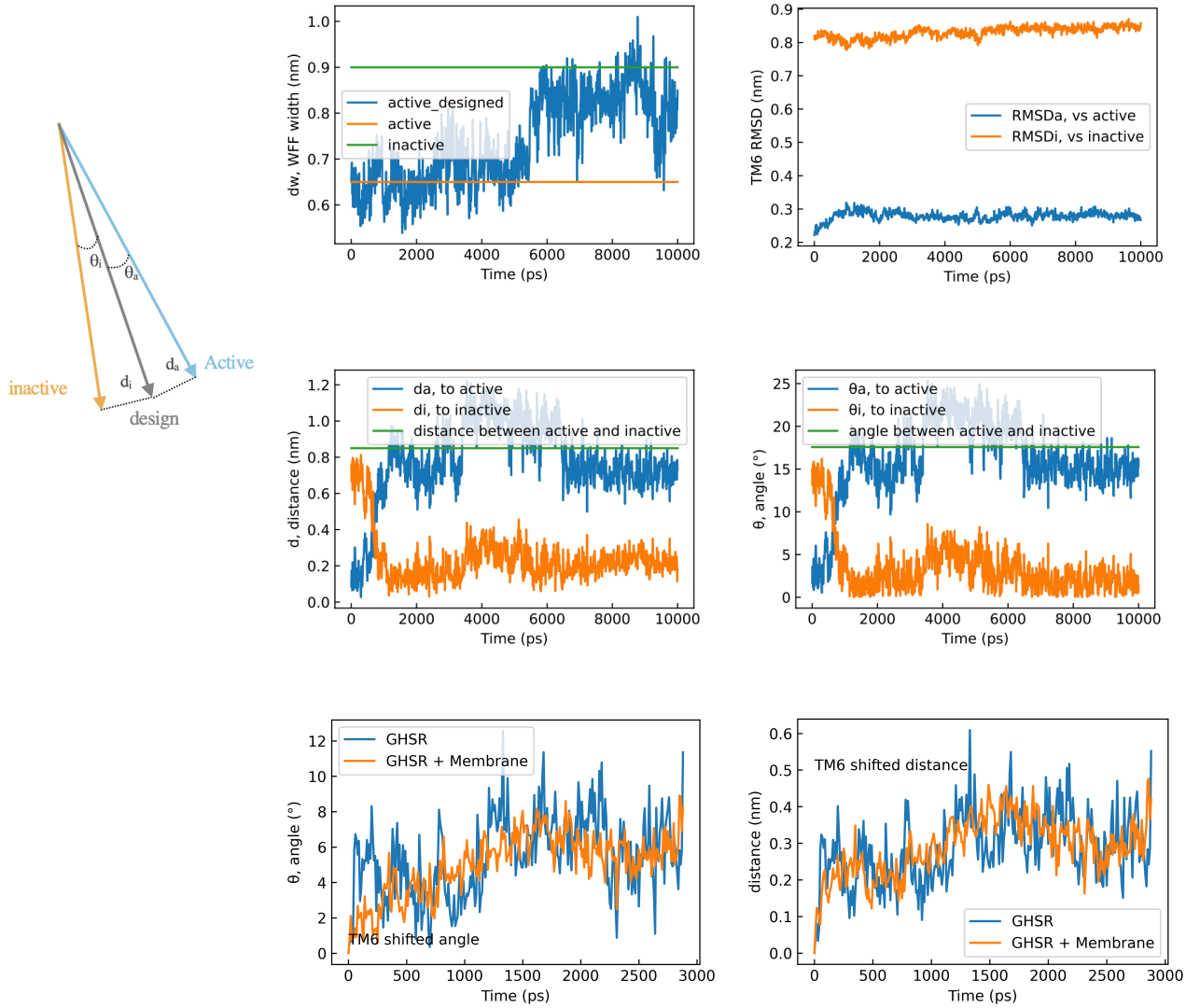

Figure S4. Comparing activation state prediction performance for GPCR class B on TM3, TM6 and TM7 domains. For GPCR class B proteins, their activation states are mainly determined by TM3, TM6 and TM7 domains. Unlike GPCR class A (Figure S2), there is less PDB data for class B, and the margin between HF-Multistate and baseline models is smaller.