YDS-Ternoplex: Surpassing AlphaFold 3-Type Models for Molecular Glue-Mediated Ternary Complex Prediction

The Statistical Mechanics Inference Team

Correspondence: Xing Che, xche@yds-pharmatech.com

Abstract

Molecular glues represent an innovative class of drugs that enable previously impossible protein-protein interactions, but their rational design remains challenging, a problem that accurate ternary complex modeling can significantly address. Here we present YDS-Ternoplex, a novel computational approach that enhances AlphaFold 3-type models by incorporating enhanced sampling inductive bias during inference to accurately predict molecular glue-mediated ternary complex structures. We demonstrate YDS-Ternoplex's capabilities across five diverse test cases, including both E3 ligase-based systems (VHL:CDO1 and CRBN complexes with mTOR-FRB, NEK7, and VAV1-SH3c) and non-E3 ligase complexes (FKBP12:mTOR-FRB). The model achieves remarkable accuracy with RMSD values as low as 1.303 Å compared to experimental structures and successfully predicts novel protein-protein interfaces not present in training data. Notably, in the FKBP12:mTOR-FRB case, YDS-Ternoplex correctly predicts a novel interface configuration instead of defaulting to known interactions present in training data, demonstrating strong generalization capabilities. Our results suggest that strategic enhancement of the inference process through inductive bias can significantly improve ternary complex prediction accuracy, potentially accelerating the development of molecular glue therapeutics for previously undruggable targets.

Introduction

Drug discovery has traditionally been limited by our ability to target only proteins with well-defined binding pockets suitable for small molecules. This constraint has left many disease-relevant proteins, particularly those operating through protein-protein interactions, beyond therapeutic reach. However, molecular glues—an innovative class of drugs—are transforming this paradigm by acting as molecular matchmakers that enable previously impossible protein-protein interactions[1].

Molecular glues operate through different mechanisms of action. Classical molecular glues, exemplified by rapamycin, act as molecular "adhesives" by stabilizing specific protein-protein interactions that would not occur spontaneously, such as the FKBP12-mTOR complex that leads to mTOR inhibition. Another important class of molecular glues triggers protein degradation, harnessing cellular protein degradation machinery by facilitating ternary complex formation between an E3 ubiquitin ligase (such as VHL or CRBN) and a target protein. These molecular recognition events involve complex cooperative binding, where the molecular glue often induces conformational changes in one protein that create new interaction surfaces to attract another protein. While this intricate interplay enables molecular glues to modulate previously "undruggable" proteins, it also makes their rational design particularly challenging[2].

Computational modeling is a common strategy for exploring rational drug design. However, traditional methods, originally developed for binary interactions such as ligand-protein or protein-protein complexes, face significant limitations in accuracy. When applied in combination to model ternary interactions, these methods often compound their inaccuracies, resulting in substantially reduced reliability. Moreover, tandem use fails to account for cooperative binding and the intricate conformational changes required, further diminishing their effectiveness.

State-of-the-art models, such as AlphaFold 3[3]-type models[4567], represent a paradigm shift in ternary complex modeling. These models leverage end-to-end deep learning architectures capable of integrating multiple modalities, including protein sequences, DNA, RNA, and ligand information. This holistic approach allows for the prediction of complex biomolecular interactions with remarkable precision, capturing the intricate conformational change that traditional methods cannot. However, even with these advancements, these models still struggle to accurately model these molecular glue-mediated interactions.

Here we present YDS-Ternoplex, which was built upon commercially available AlphaFold 3-type models. Given AlphaFold 3's demonstrated capabilities as a structure prediction foundation model, despite its limitations in handling molecular glue-mediated ternary complexes accurately, we developed an enhanced approach. By incorporating enhanced sampling inductive bias into the AlphaFold 3-type model inference process, YDS-Ternoplex

effectively predicts molecular glue-mediated ternary complex structures. We demonstrate YDS-Ternoplex's capabilities across five diverse test cases of molecular glue-mediated ternary complexes. Our approach accurately predicts structures for both E3 ligase-based systems (VHL:CDO1 and CRBN with mTOR-FRB, NEK7, and VAV1-SH3c) and non-E3 ligase complexes (FKBP12:mTOR-FRB), achieving RMSD values as low as 1.303 Å compared to experimental structures. Notably, YDS-Ternoplex successfully predicts all five novel protein-protein interfaces not present in training data. This capability is particularly evident in the FKBP12:MG:mTOR-FRB complex, where an alternative protein-protein interface exists in the training data that could potentially bias predictions toward this known interaction. YDS-Ternoplex's ability to correctly predict the novel interface instead of defaulting to the known interaction demonstrates its strong generalization capabilities rather than mere memorization of training data.

Results

Case 1: Ternary structure prediction of VHL:MG:CDO1

The first case is to predict the ternary complex of VHL:MG:CDO1, which has a solved structure in PDB (PDB ID: 8VL9). We specifically selected this case because no ternary complex structures of the VHL-CDO1 protein pair were included in the training set. This deliberate choice helps avoid suspicion of model memorization, making it an ideal case for evaluating the predictive capabilities of our method[8].

Von Hippel-Lindau (VHL) functions as the substrate recognition component of the cullin-RING E3 ligase complex, playing a critical role in protein degradation via the ubiquitin-proteasome pathway. The protein of interest (POI) in this complex is cysteine dioxygenase type 1 (CDO1), an enzyme responsible for cysteine metabolism and regulation of oxidative stress. Dysregulation of CDO1 has been implicated in various cancers, particularly as a tumor suppressor. The biological significance of this case lies in demonstrating how molecular glues (MGs) can mediate interactions between metabolic enzymes like CDO1 and E3 ligases, potentially offering novel therapeutic strategies for targeting cancer-associated metabolic pathways[8].

From our ensemble of sampled structures, we selected the model with the highest interface-predicted TM-score (ipTM), which reached 0.865. When compared to the experimentally determined crystal structure (PDB ID: 8VL9), this model showed excellent agreement with an RMSD of 1.303 Å. As shown in the structural alignment (Figure 1a), YDS-Ternoplex effectively predicted the overall architecture of the ternary complex, and binding conformation of the MG.

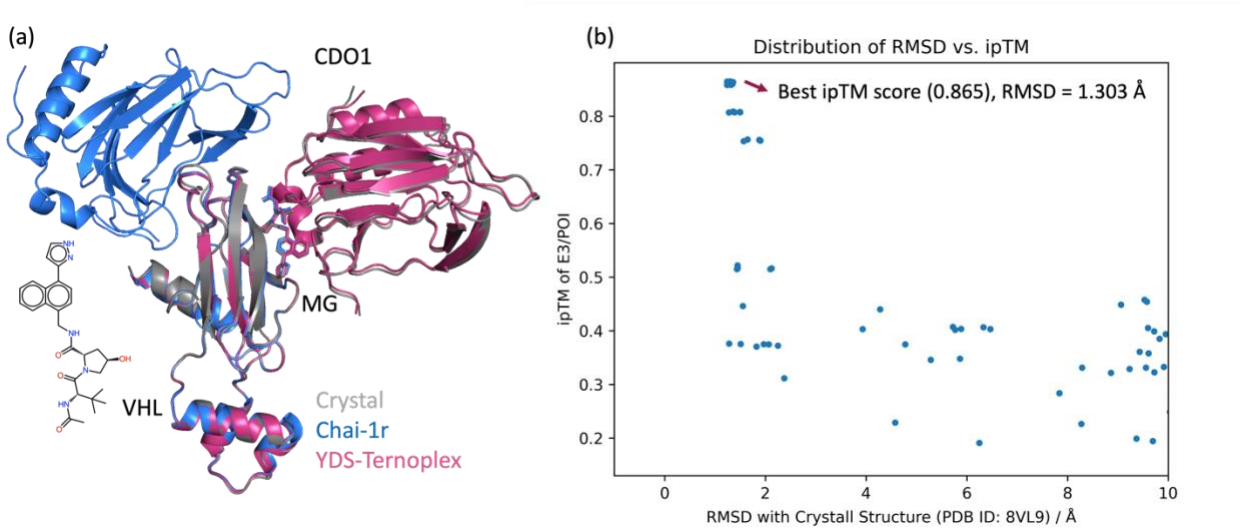

Figure 1. YDS-Ternoplex prediction of the VHL:MG:CDO1 ternary complex. (a): Structural alignment of the crystal structure (gray), YDS-Ternoplex prediction (pink), and Chai-1r prediction (blue). Right: Distribution of interface-predicted TM-score (ipTM) versus RMSD for YDS-Ternoplex sampled structures compared to the crystal structure (PDB ID: 8VL9). The best model, selected based on having the highest ipTM score (0.865) among all samples, shows an RMSD of 1.303 Å (indicated by arrow). RMSD was calculated using all heavy atoms.

Case 2: Ternary structure prediction of CRBN:MG:mTOR-FRB

The second case explores the prediction of the ternary complex comprising CRBN, the mTOR-FRB domain, and a MG compound. mTOR (mammalian target of rapamycin) is a central regulator of cellular growth, metabolism, and survival, functioning as part of two distinct complexes, mTORC1 and mTORC2. Its FRB (FKBP12-rapamycin binding) domain plays a critical role in binding regulatory proteins and small molecules, making it a key target in therapeutic applications. CRBN (Cereblon) acts as the substrate receptor in the $CRL4^{CRBN}$ E3 ubiquitin ligase complex, facilitating the degradation of neosubstrate proteins upon ternary complex formation[9]. The YDS-Ternoplex prediction was selected from an ensemble of sampled structures with its highest inter-protein TM-score (ipTM) of 0.785.

Though there is no structure in PDB, recent research from Monte Rosa Therapeutics[9] has disclosed a detailed analysis on their structure data. Analysis of the helical G-loop within the mTOR-FRB domain provided key insights into the ligand-mediated stabilization of the ternary complex. The YDS-Ternoplex model accurately captured critical interactions

between mTOR-FRB residues (M2089, K2090, and S2091) and CRBN residues (N351, H357, and W400), forming a network of hydrogen bonds that align closely with experimentally validated structural data.

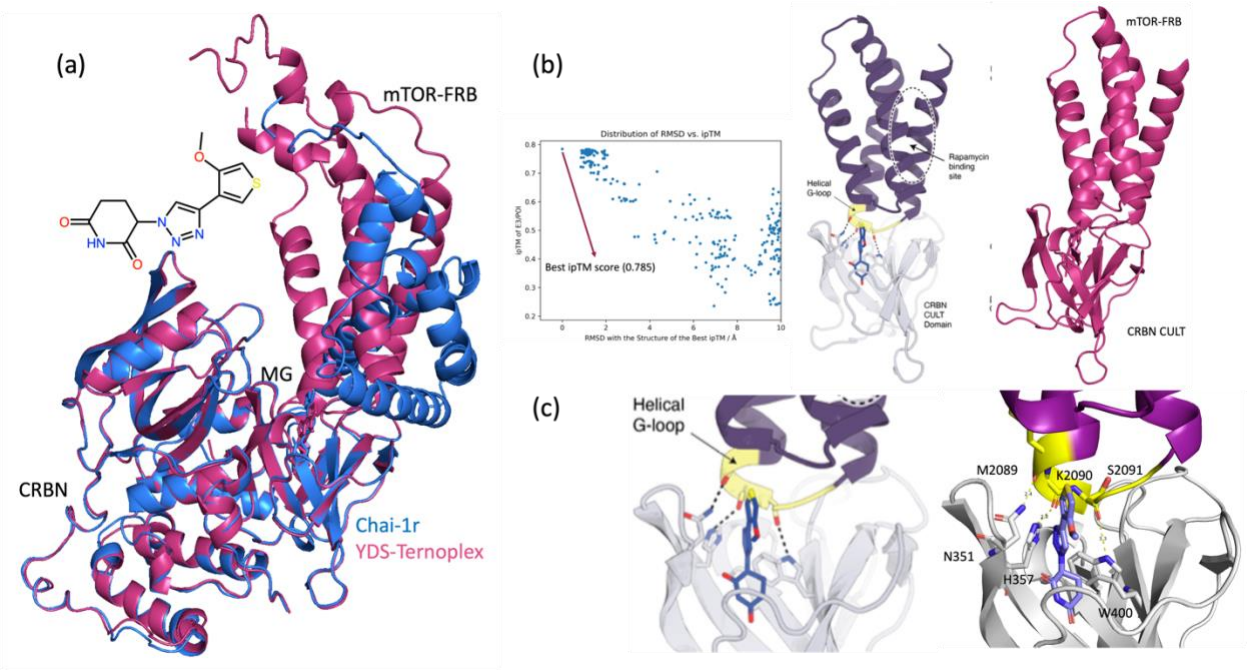

Figure 2. Predicted CRBN:MG:mTOR-FRB ternary complex and comparison with experimental analysis. (a) Overlay of predicted models YDS-Ternoplex (pink) and Chai-1r (blue). (b) RMSD versus ipTM scatterplot of YDS-Ternoplex sampled structures, with the highest scoring model (YDS-Ternoplex) achieving an ipTM of 0.785. The reference structure for the RMSD calculation was the structure with the highest ipTM score. And separated views of CRBN CULT and mTOR-FRB domains, showing ligand-mediated interaction regions. (c) Detailed view of the mTOR-FRB helical G-loop interacting with CRBN, showing ligand-mediated stabilization through interactions involving key residues in mTOR-FRB (M2089, K2090, S2091) and CRBN (N351, H357, W400), consistent with experimental data.

Case 3: Ternary structure prediction of CRBN:MG:NEK7

This work presents the accurate prediction and experimental validation of the CRBN:NEK7 ternary complex facilitated by a molecular glue compound, emphasizing the reliability and consistency of the YDS-Ternoplex model. NEK7 (NIMA-related kinase 7) is a serine/threonine kinase essential for microtubule organization and mitotic spindle assembly during cell division. Its dysregulation is associated with inflammatory responses and cancer, making it a target of interest[9]. The YDS-Ternoplex prediction was selected from an ensemble of sampled structures with its highest inter-protein TM-score (ipTM) of 0.797.

Though there is no structure in PDB, recent research from Monte Rosa Therapeutics[9] has disclosed a detailed analysis on their structure data. The YDS-Ternoplex model successfully predicted the key interactions between NEK7 residues (L54, L55, D56) and CRBN residues (N351, H357, and W400), forming a precise network of hydrogen bonds in strong agreement with experimental results.

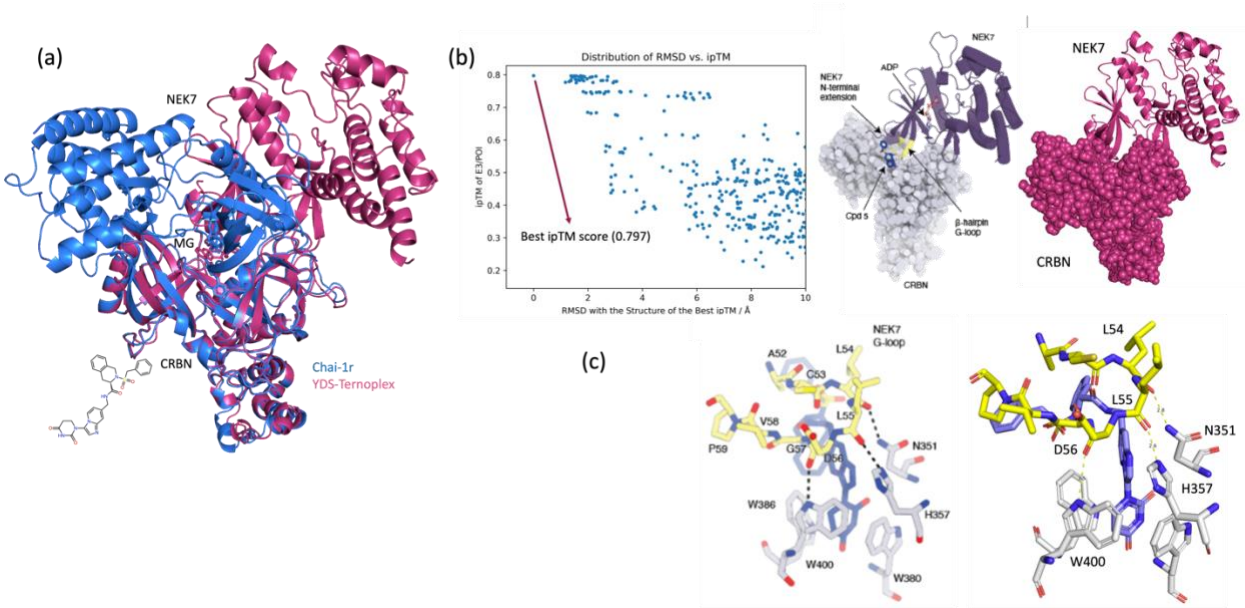

Figure 3**:** Predicted CRBN:MG:NEK7 ternary complex and comparison with experimental analysis. (a) Overlay of predicted models YDS-Ternoplex (pink) and Chai-1r (blue). (b) RMSD versus ipTM scatterplot of YDS-Ternoplex sampled structures, with the highest scoring model (YDS-Ternoplex) achieving an ipTM of 0.797. The reference structure for the RMSD calculation was the structure with the highest ipTM score. Separated views of CRBN CULT and NEK7 domains highlight the ligand-mediated interaction regions. (c) Detailed view of the NEK7 β-hairpin G-loop interacting with CRBN, showing ligand-mediated stabilization through interactions involving key residues in NEK7 (L54, L55, D56) and CRBN (N351, H357, W400), consistent with experimental data.

Case 4: Ternary structure prediction of CRBN:MG:VAV1-SH3c

In the fourth case, the POI is the SH3c domain of VAV1, a guanine nucleotide exchange factor (GEF) that plays a critical role in T-cell activation and cytoskeletal dynamics. The SH3c domain is crucial for mediating protein-protein interactions necessary for VAV1's function in immune signaling. By forming a ternary complex with VAV1's SH3c domain, molecular glues provide a unique mechanism to modulate immune pathways, which could have significant implications for immunotherapy[9].

As illustrated by Figure 4, the structural prediction and experimental validation of the CRBN:VAV1 SH3c ternary complex mediated by a molecular glue compound demonstrate the accuracy of the YDS-Ternoplex model. The predicted structure with the highest inter-protein TM-score (ipTM) of 0.785, selected from an ensemble of sampled structures, shows a close match to experimentally validated binding poses. Though there is no structure in PDB, recent research from Monte Rosa Therapeutics[9] has disclosed a detailed analysis on their structure data. Detailed views of YDS-Ternoplex structure reveal that the VAV1 SH3c domain engages CRBN through the RT-loop, forming key hydrogen bonds between VAV1 residues (R796, D797, and S799) and CRBN residues (W400, H357, and N351). These interactions, as well as the molecular surface mimicry of the VAV1 RT-loop, are consistent with experimental observations, underscoring the predictive power of the YDS-Ternoplex.

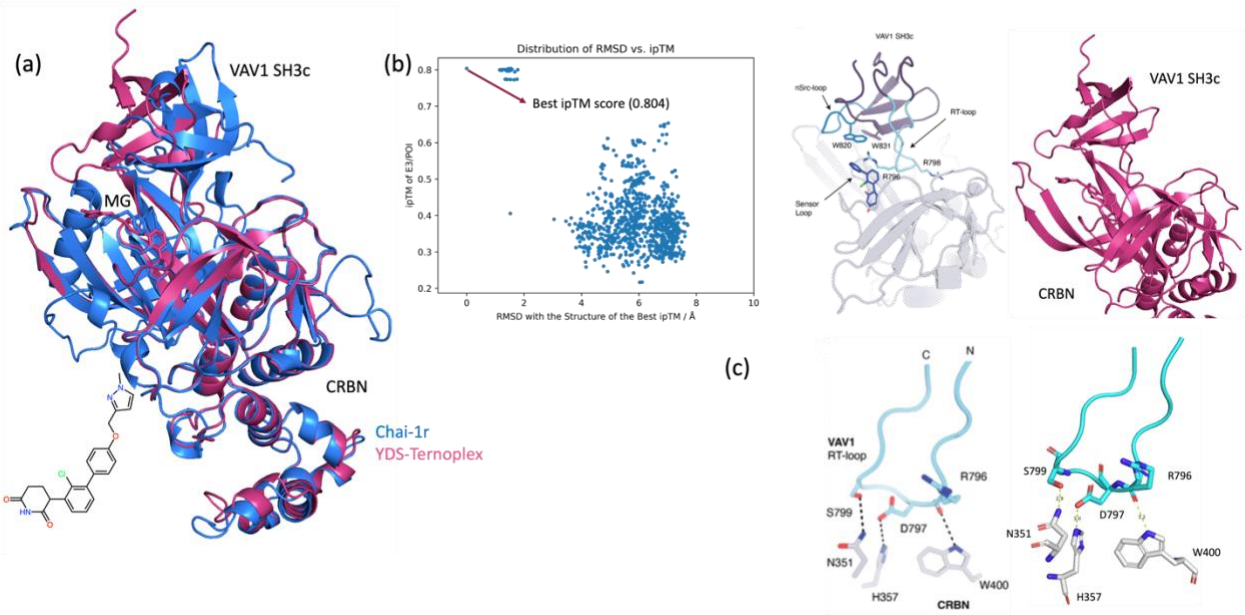

Figure 4. Predicted CRBN:MG:VAV1-SH3c ternary complex and comparison with experimental data. (a) Overlay of predicted models YDS-Ternoplex (pink) and Chai-1r (blue). (b) RMSD versus ipTM scatterplot of YDS-Ternoplex sampled structures, with the highest scoring model (YDS-Ternoplex) achieving an ipTM of 0.785. The reference structure for the RMSD calculation was the structure with the highest ipTM score. Separated views of CRBN and VAV1-SH3c domains highlight the ligand-mediated interaction regions. (c) Detailed view of the VAV1 RT-loop interacting with CRBN, showing ligand-mediated stabilization through interactions involving key residues in VAV1 (R796, D797, and S799) and CRBN (W400, H357, and N351), consistent with experimental data.

Case 5: Ternary structure prediction of FKBP12:MG:mTOR-FRB

The final case is the FKBP12:MG:mTOR-FRB ternary complex, which involves FKBP12 and the FRB domain of mTOR. Unlike the other cases, FKBP12 is not an E3 ligase but a prolyl isomerase that binds to immunosuppressive drugs such as rapamycin. There are a total of 11 PDB structures that include FKBP12, mTOR-FRB, and molecular glues. Remarkably, 10 of these structures share similar ternary configurations and protein-protein interfaces, whereas the most recent structure, PDB ID: 8PPZ, exhibits a completely different ternary structure and protein-protein interaction (PPI) pattern compared to the others[10]. Since 8PPZ was not included in the training set and the model never see this PPI during training, this case provides a unique opportunity to evaluate whether the model can correctly predict an unseen ternary configuration by learning potential atomic-level interactions.

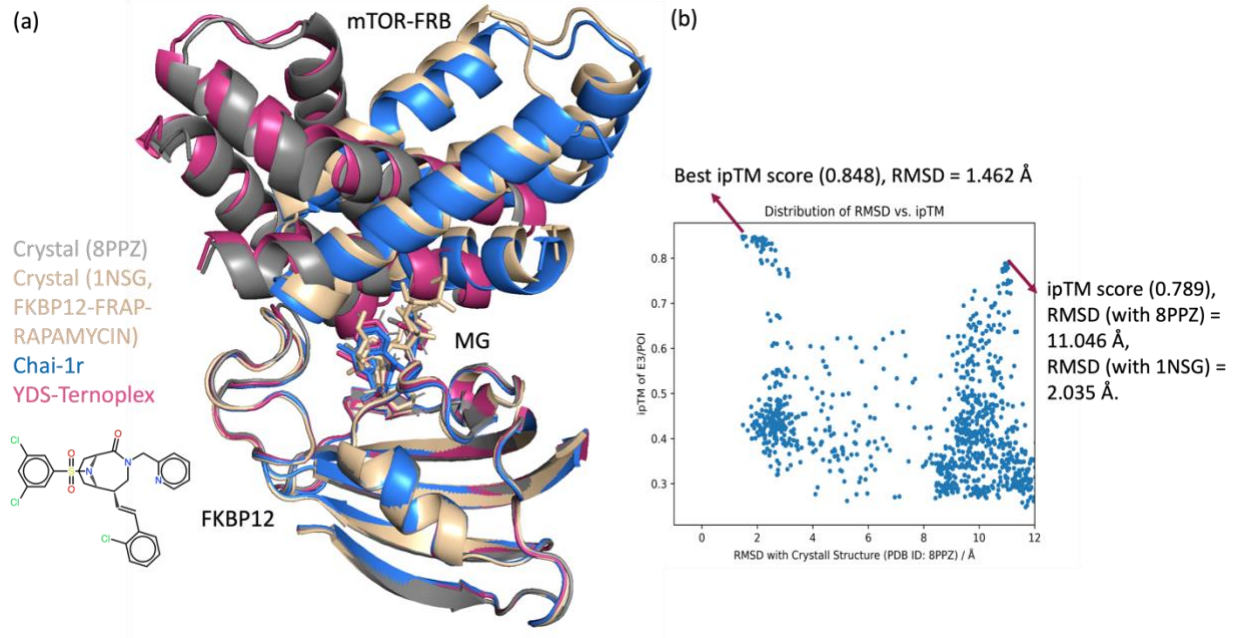

Figure 5. Predicted FKBP12:MG:mTOR-FRB ternary complex and comparison with experimental structures. (a) Overlay of the crystal structure (PDB ID: 8PPZ) with predicted models YDS-Ternoplex (pink) and Chai-1r (blue), and another molecular glue rapamycin indued ternary structure 1NSG. (b) RMSD versus ipTM scatterplot of YDS-Ternoplex sampled structures. The reference structure for RMSD calculation was 8PPZ.

As illustrated by Figure 5, the YDS-Ternoplex model accurately predicts the FKBP12:MG:mTOR-FRB ternary complex. The structure with the highest inter-protein TM-score (ipTM) of 0.848 was selected from the ensemble of sampled models and aligns well with the experimental crystal structure (PDB ID: 8PPZ), achieving a root-mean-square deviation (RMSD) of 1.462 Å. Without incorporating enhanced sampling inductive bias into the inference process, models like Chai-1r (blue) tend to predominantly sample protein-protein interactions (PPIs) that have been seen during training, as shown by the strong overlap of the Chai-1r prediction with the 1NSG structure. In our sampling, this type of PPI is associated with an ipTM around 0.789. The YDS-Ternoplex model demonstrates its ability to generalize beyond seen PPIs by sampling and accurately predicting a novel ternary configuration, showcasing its capability to infer unseen PPIs and novel molecular glue-mediated interactions.

Discussion

The YDS-Ternoplex model demonstrates substantial advancements in predicting molecular glue-mediated ternary complexes by incorporating enhanced sampling inductive bias. By building on AlphaFold 3-type models, YDS-Ternoplex not only achieves higher predictive accuracy but also overcomes the limitations of existing approaches that tend to overfit to seen protein-protein interaction patterns or fail to capture cooperative binding central to molecular glue mechanisms.

Our results highlight the model's capability to generalize across diverse ternary complexes, including both E3 ligase-based (e.g., VHL:CDO1, CRBN:NEK7) and non-E3 ligase systems (e.g., FKBP12:mTOR-FRB). In each case, the model accurately captured critical interaction interfaces and ligand-mediated stabilization, evidenced by low RMSD values and strong alignment with experimental data. Notably, the FKBP12:mTOR-FRB case exemplifies YDS-Ternoplex's ability to infer entirely novel configurations, distinguishing it from baseline models that default to memorized interaction patterns.

The success of YDS-Ternoplex in achieving state-of-the-art performance by adding enhanced sampling inductive bias exclusively during the inference process underscores a significant implication: the Pairformer module has effectively learned the interaction landscape. This suggests that the underlying latent space is adept at capturing the foundational features of high dimensional interactions enabling robust predictions with only inference-stage adjustments. This efficiency not only validates the design of the

Pairformer architecture but also highlights its potential for broader applications in mapping interaction to function.

The integration of enhanced sampling inductive bias is central to this success. By drawing on principles from statistical mechanics, the model efficiently navigates high-dimensional conformational landscapes, overcoming energy barriers that limit traditional sampling approaches. This targeted exploration of rare but biologically significant configurations ensures robust predictions, even for complexes with limited structural precedence.

Moreover, the theoretical framework of inference scaling laws[11] underpins our approach, emphasizing that indiscriminate increases in computational resources do not guarantee performance gains. Instead, strategic enhancements to the inference process, such as our inductive bias mechanism, drive meaningful improvements. This insight underscores the importance of pairing computational resources with methodological innovations to achieve scalable and efficient predictions.

In conclusion, YDS-Ternoplex represents a significant leap forward in computational modeling of molecular glue-mediated ternary complexes. By addressing the intricate interplay of cooperative binding, it opens new avenues for rational drug design targeting previously undruggable proteins. These advancements pave the way for more precise and effective therapeutic strategies, expanding the scope of molecular glue applications in drug discovery.

Reference:

1. Garber, K. The glue degraders. *Nat Biotechnol* **42**, 546–550 (2024).

2. Konstantinidou, M. & Arkin, M. R. Molecular glues for protein-protein interactions: Progressing toward a new dream. *Cell Chemical Biology* **31**, 1064–1088 (2024).

3. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

4. Wohlwend, J. *et al.* Boltz-1 Democratizing Biomolecular Interaction Modeling. Preprint at https://doi.org/10.1101/2024.11.19.624167 (2024).
5. Chai Discovery *et al.* Chai-1: Decoding the molecular interactions of life. Preprint at https://doi.org/10.1101/2024.10.10.615955 (2024).
6. Qiao, Z. *et al.* NeuralPLexer3: Accurate Biomolecular Complex Structure Prediction with Flow Models. Preprint at https://doi.org/10.48550/arXiv.2412.10743 (2024).
7. Protenix_Technical_Report.
8. Tutter, A. *et al.* A small molecule VHL molecular glue degrader for cysteine dioxygenase 1. Preprint at https://doi.org/10.1101/2024.01.25.576086 (2024).
9. Petzold, G. *et al.* Mining the CRBN Target Space Redefines Rules for Molecular Glue-induced Neosubstrate Recognition. Preprint at https://doi.org/10.1101/2024.10.07.616933 (2024).
10. Deutscher, R. C. E. *et al.* Discovery of fully synthetic FKBP12-mTOR molecular glues. Preprint at https://doi.org/10.26434/chemrxiv-2023-4vb0m-v2 (2024).
11. Wu, Y., Sun, Z., Li, S., Welleck, S. & Yang, Y. Inference Scaling Laws: An Empirical Analysis of Compute-Optimal Inference for Problem-Solving with Language Models. Preprint at https://doi.org/10.48550/arXiv.2408.00724 (2024).