OXFORD

Antibody Therapeutics, 2026, 9(4), 674–686
https://doi.org/10.1093/abt/tbag053
Advance access publication date 13 October 2026
Research Article

PPBind-3D/1D: protein–protein binding affinity prediction for structural and sequence data

Peiyi Chen[1,2,‡], Huaqing Liu[1,2,3,‡], Shuxian Zhou[1,2], Lunji Wang[1,2], Mao Pang[4], Jing Zhao[5], Guoxin Fan[6,*], Lanqing Han[1,2,7,*]

[1]Artificial Intelligence Innovation Center, Research Institute of Tsinghua, Pearl River Delta, Guangzhou, Guangdong, 510700, China
[2]Cyagen Biomodels (Guangzhou) Co., Ltd, Guangzhou, Guangdong, 510700, China
[3]School of Pharmacy, Faculty of Medicine & Laboratory of Drug Discovery from Natural Resources and Industrialization, Macau University of Science and Technology, Macau, 999078, China
[4]Department of Spine Surgery, The Third Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, 510630, China
[5]Scientific Research Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, Guangdong, 518107, China
[6]National Key Clinical Pain Medicine of China, Shenzhen Nanshan People's Hospital, Shenzhen University Medical School, Shenzhen, Guangdong, 518052, China
[7]Cyagen Biosciences (Suzhou) Inc., Suzhou, Jiangsu, 215000, China

*Corresponding authors. Guoxin Fan, No. 89 Taoyuan Road, Shenzhen 518052, China. E-mail: fanguoxin@email.szu.edu.cn; Lanqing Han, No. 98 Xiangxue Ba Road, Guangzhou 510735, China. E-mail: hanlance@tsinghua-gd.org
‡Peiyi Chen and Huaqing Liu contributed equally to this work.

Abstract

Background: One key step of protein drug development is the screening of protein–protein binding (PPB) affinity. Various deep learning methods have been proposed to predict PPB affinity, but they are often challenged by data leakage, data scarcity, data heterogeneity, and structure-sequence gap.
Methods: To address these challenges in PPB affinity prediction, we propose a comprehensive framework by developing prediction models, PPBind-3D and PPBind-1D, to predict PPB affinity. PPBind-3D leverages structural information near the PPB interface to make its predictions. Additionally, PPBind-1D was developed using sequence data to address the lack of structural data in practical applications.
Results: PPBind-3D, equipped with monotonic neural network-constrained multi-task learning and Self-and-Mutual Attention, demonstrated state-of-the-art performance in PPB prediction across diverse datasets under both strict and random data partition settings. PPBind-1D, trained by aligning to PPBind-3D, overcame structural data scarcity to deliver robust sequence-based predictions while maintaining high accuracy, enabling practical deployment where structural data is unavailable. To facilitate real-world application, PPBind-3D was integrated into an online tool, which is available at https://rddc.tsinghua-gd.org/tool/dg-prediction. Finally, we demonstrated three application cases of our artificial intelligence models in the virtual screening of protein drugs, further confirming the framework's utility and scalability.
Conclusions: Comparative and ablation studies demonstrate that PPBind-3D surpasses conventional benchmark approaches. Utilizing solely sequence information, PPBind-1D closely approximates the performance of PPBind-3D, thereby providing a viable alternative when structural data are inaccessible. Validation across three virtual screening case studies confirms the efficacy of both models, illustrating that our models can significantly facilitate high-throughput screening.

Statement of Significance
The study developed a comprehensive prediction framework (PPBind-3D/1D) to deal with the challenges like data leakage, data scarcity, data heterogeneity, and structure-sequence gap in predictions of PPB affinity. PPBind-3D surpasses conventional benchmark approaches, while PPBind-1D closely approximates the performance of PPBind-3D.

Keywords: artificial intelligence; protein–protein interaction; PPB-affinity dataset; binding affinity; virtual screening

Introduction

A critical challenge in the engineering of protein drugs is to assess the strength of binding between the protein drug and the target protein, known as protein–protein binding (PPB) affinity. The therapeutic effect of protein drugs typically relies on their ability to bind to specific target proteins. Protein drugs with high PPB affinity can bind more effectively to target proteins, thereby exerting a therapeutic effect. On the other hand, protein drugs with high PPB affinity can bind more specifically to the target proteins, reducing the impact on non-target proteins. This helps to reduce the side effects of the drug and enhance the safety of treatment.

High-throughput PPB affinity screening can accelerate the development of protein drugs. In recent years, technologies such as protein microarrays [1] and the Octet system [2] have been developed. Although these experimental methods are accurate, they require cumbersome experimental operations, strict experimental conditions, and expensive equipment and consumables. Therefore, algorithm-based PPB affinity prediction is a more promising paradigm for high-throughput screening. However, current computational approaches face significant hurdles. The scarcity of large, high-quality, experimentally validated affinity datasets limits the training and generalization of deep learning models [3, 4]. Furthermore, the inherent heterogeneity in

Received: August 4, 2025. **Revised:** June 24, 2026. **Accepted:** July 17, 2026
© The Author(s) 2026. Published by Oxford University Press on behalf of the Chinese Antibody Society.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

affinity data generated from diverse experimental methodologies (e.g., surface plasmon resonance [SPR], isothermal titration calorimetry [ITC], and enzyme-linked immunosorbent assay [ELISA]) introduces substantial noise and distributional shifts, complicating model training. Existing algorithms often struggle to effectively integrate or reconcile these disparate data sources, leading to suboptimal performance when applied to novel protein pairs or data from different experimental origins. Additionally, while structural information provides critical insights into binding mechanisms, the practical application of structure-based affinity prediction is severely hampered by the limited availability of high-resolution 3D structures for many protein complexes of interest. This necessitates robust methods capable of leveraging abundant sequence data when structural data is absent, without sacrificing predictive accuracy. To addressing the challenges of data scarcity, data heterogeneity, and structure-unavailability, it is crucial to develop reliable, widely applicable PPB affinity prediction tools to accelerate therapeutic discovery.

In order to address the aforementioned challenges, this work makes three contributions. First, the largest protein affinity dataset to date, PPB-Affinity [5], comprising 12 062 samples, was employed in the development of a geometric deep learning model, PPBind-3D, which predicts PPB affinity based on structural features near the binding interface of protein–protein complexes. We also trained the model by integrating the heterogeneous affinity data, especially deep mutation screening (DMS) [6] data, through a monotonic neural network module [7, 8], thereby further enhancing the model's generalization performance. Second, we proposed a more rigorous method for clustering protein complex structures. In previous artificial intelligence studies of predicted binding affinity, there has always been data leakage of varying degrees due to the lack of rigorous data division, making it impossible to accurately assess the predictive performance of the model. To address this, we calculated the features of protein complex structures in a SE(3)-Invariant manner using the iDist algorithm [9] and then clustered the protein complex structure features based on graph partition algorithms [10], thus achieving a more rigorous data division. Finally, we developed a sequence model PPBind-1D based on our innovative "Feature Alignment" principle, which guided the sequence model through structural models to achieve the predictive performance of structural models. Additionally, a substantial number of authentic protein complex structures were employed, including unlabeled samples, to assist in training the PPBind-1D model to align with the PPBind-3D model.

Materials and methods

Data composition

Our data source is shown in Fig. 1A, which mainly consists of four parts: (i) PPB-Affinity Dataset, this is the largest protein affinity dataset to date, where each sample has a experimentally measured binding free energy ΔG, the three-dimensional structure of the wild-type complex, and mutation information, etc. (ii) Heterogeneous DMS Affinity Datasets, heterogeneous affinity datasets such as PBAD-AS [11], PDAD-SA [12], where the affinity measurements are not ΔG or dissociation constant K_d, but rather $K_{d,app}$ or log_2 enrichment ratio. Within the same set of experiments, these measurements are positively or negatively correlated with the affinity ΔG values, but they cannot be directly converted to ΔG values using known formulas. (iii) Database of Interacting Protein Structures-Plus (DIPS-Plus) [13], an enhanced, feature-rich protein complex structure dataset of 42 108 complexes for geometric deep learning of protein interfaces. (iv) Affinity validation data of nanobodies with different antigens, including Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4), Programmed Death Protein 1 (PD1), Programmed Death-Ligand 1 (PD-L1), and Hen Egg Lysozyme (HEL). The PPB-Affinity dataset and the heterogeneous DMS affinity datasets are used for the development and validation of PPBind-3D and PPBind-1D, while the DIPS-Plus dataset is used exclusively for the development of PPBind-1D. The validation case Dataset does not participate in model training.

Data partitioning

Data partition is usually used to verify the true performance of the model. For protein affinity data, however, traditional random partition is not reasonable because the same or similar protein complexes may appear in both the training set and the validation set, resulting in an inability to correctly evaluate the model's performance. Rotamer Density Estimator (RDE)-Network [14] proposed data partition based on Protein Data Bank (PDB) code, but there may be data leakage [9] due to the fact that protein complexes with different PDB codes may also be composed of homologous proteins (such as 2NU0, 1SGQ). In order to address this issue, we propose a novel data partitioning method based on Anton Bushuiev's SE (3) Protein-Protein Interaction (PPI) redundancy removal technique iDist [9]. The specific data division process is detailed in Supplementary S1. To explore the impact of data volume and data partitioning methods on model performance, we conducted additional experiments, details of which can be found in Supplementary S2 and S3.

PPBind-3D

We designed the network as illustrated in Fig. 1B. Firstly, it should be noted that despite the significant differences in protein length and conformation observed between different protein complexes, they all possess a binding interface that directly affects affinity. In order to concentrate the model on the area in close proximity to the binding interface, for each amino acid residue present in the receptor of the protein complex, if there is an amino acid residue present on the ligand and the distance between their C-alpha atoms is less than 10 Å, then this pair of residues is defined as the binding sites. The amino acid residues in the receptor and ligand in closest proximity to the binding site were extracted using the K nearest neighbors algorithm, which identified the visible patches of the model. We employ RDE [14] to extract and simulate the amino acid side chain potential conformational distribution information. To fully leverage the affinity data of various protein complex mutations without the necessity of inputting mutant structures, thereby significantly expanding the quantity of available data.

Subsequently, in order to fully leverage the information derived from the three-dimensional structure, we represented the residues as nodes and their pairs as edges. This allows us to represent the protein complex as a complete graph. Specifically, we define the node feature vector at the residue level as h, which includes the type of amino acid residue, physicochemical properties of amino acids, relative solvent-accessible surface area, types of dihedral angles, and types of side chain torsion angles. The edge feature vector is denoted as e, including the amino acids types, differences in relative solvent-accessible surface areas, relative positions, Euclidean distances, and virtual dihedral angles between the two connected residues.

The core of our architecture is the geometric encoder, inspired by DDG-Pred [15] and RDE-Network [14], which is an SE (3) invariant attention module. In the Geometric Encoder, two modes of feature updating based on attention, "SELF" and "MUTUAL" (S&M),

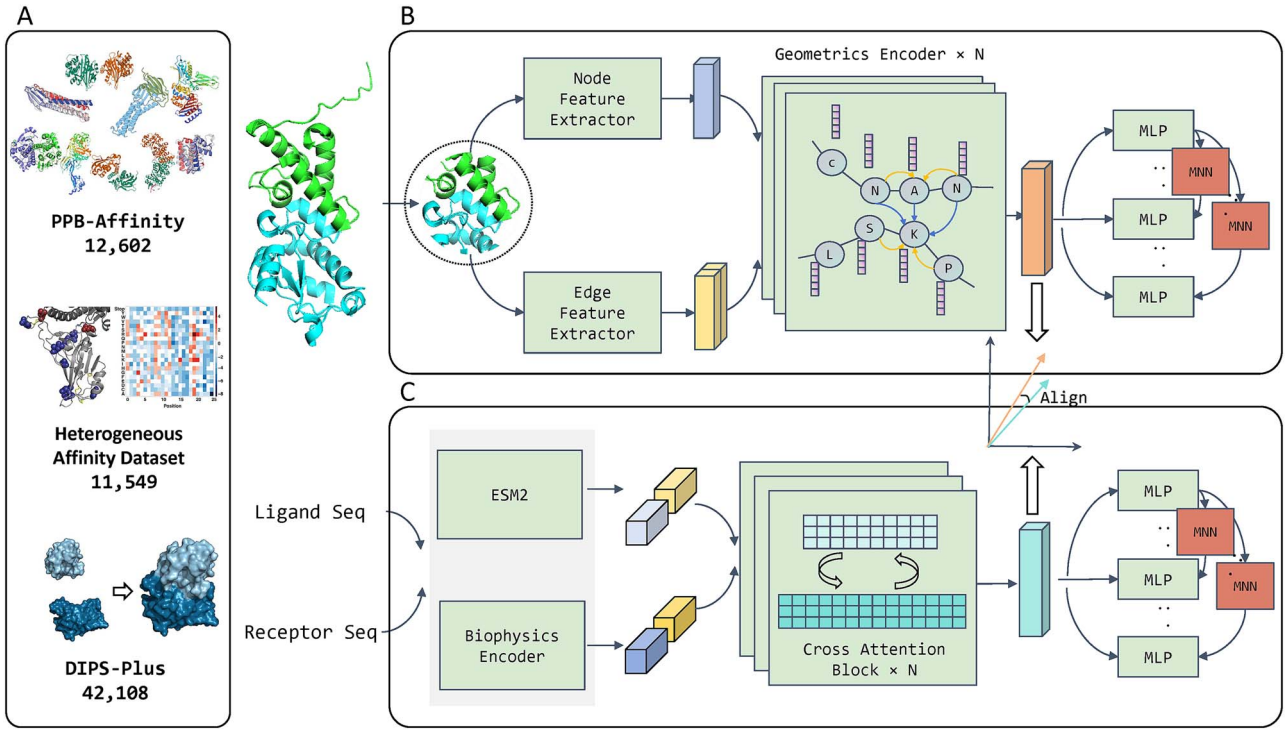


Figure 1. Overview of PPBind series model, include (A) the composition of dataset, (B) PPBind-3D, and (C) PPBind-1D architectures.

are designed. The "SELF" mode involves the ligand or receptor updating features based solely on its own structural information, while the "MUTUAL" mode involves the ligand or receptor updating features based on the structural information of the counterpart. A complete feature update is defined as a process that begins with "SELF" and then proceeds to "MUTUAL" once more. Specifically, for a L-layer model, the attention computation process in the l-th ($1 \le l \le L$) layer Geometric encoder can be represented as follows:

$$\alpha_{ij}^{h(l)} = \frac{1}{\sqrt{d}}\text{Linear}\left(h_i^{(l)}\right)\cdot\text{Linear}\left(h_j^{(l)}\right)^T \tag{1}$$

$$\alpha_{ij}^{e(l)} = \text{Linear}\left(e_{ij}^{(l)}\right) \tag{2}$$

$$\alpha_{ij}^{\text{spatial}(l)} = \gamma \parallel \left(R_i\text{Linear}\left(h_i^{(l)}\right)+T_i\right)-\left(R_j\text{Linear}\left(h_j^{(l)}\right)+T_j\right)\parallel_2 \tag{3}$$

$$\alpha_{ij}^{(l)} = \text{softmax}\left(\alpha_{ij}^{h(l)}+\alpha_{ij}^{e(l)}+\alpha_{ij}^{\text{spatial}(l)}\right) \tag{4}$$

Among them, R and T represent the rotation matrix and translation vector of the i-th residue transformed from the local coordinate system to the global coordinate system; h represents the node feature; γ is a learnable parameter, and $\alpha_{ij}^{(l)}$ is the weight of the l-th layer Geometric encoder attention. In the "SELF" mode, both i and j are residues in the ligand or receptor. If i and j are not homologous, the mode is "MUTUAL." The process of feature updating can be expressed as:

$$h_i^{(l)'} = \text{Concat}\left(\alpha_{ij}^{(l)}\text{Linear}\left(h_i^{(l)}\right), \sum_j \alpha_{ij}^{(l)}\text{Linear}\left(e_{ij}^{(l)}\right), R_i^{-1}\alpha_{ij}^{(l)}\right.$$
$$\left.\text{Linear}\left(h_i^{(l)}\right)-T_i\right) \tag{5}$$

Many studies on PPB affinities, such as the heterogeneous DMS affinity dataset we collected, did not measure the K_d or ΔG values directly, but measured values like ligand enrichment, which were more abundant. Although these values cannot be directly converted into K_d or ΔG values through a formula, they are positively or negatively correlated with K_d and ΔG. To leverage this valuable heterogeneous affinity wet-lab data, we referred to G. Wang's work [16] and introduced Monotonic Neural Networks into the prediction head, which we called it as monotonic neural network-constrained multi-task learning (MMTL). Specifically, affinity prediction is treated as multi-task learning, with each task corresponding to a distinct prediction head, all prediction head sharing a common backbone network. The primary prediction head is tasked with predicting ΔG values, while the other prediction heads predict various non-ΔG from different sources. Thus, the learning objective for PPBind-3D can be expressed as a minimization objective function as follows.

$$\underset{\theta,\theta_t}{\text{argmin}}\frac{1}{T*N}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(y_{t,i}-M_\theta(x_i)_t\right)^2$$
$$+\frac{1}{(T-1)*N}\sum_{t\neq 1}^{T}\sum_{i=1}^{N}\left(\lambda_t\cdot M_\theta(x_i)_1-F_{\theta_t}\left(\lambda_t\cdot M_\theta(x_i)_t\right)\right)^2 \tag{6}$$

Here, T represents the task (prediction head) index, $y_{t,i}$ represents the true value of the i-th sample for task t. M_θ is PPBind network used to predict values for different tasks. F_θ is a neural network that approximates the computation of integrals using the Crenshaw-Coulters method, thereby enhancing the accuracy of the integrals. It is capable of learning and integrating monotonically increasing functions. For more details on neural network F_θ, please refer to section "Integration of Heterogeneous Data". As F_θ is applicable solely to functions that control monotonically increasing functions, we employ the symbol λ_t to denote the monotonicity of task label values with respect to ΔG, where a value of 1 denotes monotonically increasing and -1 denotes monotonically decreasing. The term before the "+" is the mean squared error formula, which is used for training the model to predict affinity. The term following the "+" is used to train the model on the monotonicity between different affinity and ΔG values.

PPBind-1D

In protein complexes, there are often more than one chain of ligands and receptors, i.e. the ligand and receptor themselves might constitute a complex. Currently, the protein language models or other sequence models can only take monomeric sequences and a linker is commonly used to connect the complex sequences into a single entity to accommodate complex sequences to handle complex sequences. To simplify the problem, this study considers data where the number of receptor and ligand chains does not exceed two. Thus, the most complex protein complex situation addressed here is that both the ligand and receptor are dimers. A linker consisting of 25 Gly residues is used to connect the subcomplexes of the ligand or receptor.

We designed the network as illustrated in Fig. 1C. We used physicochemical properties of amino acids and protein language model Evolutionary Scale Modeling-2 (ESM2) [17] to extract the basic features of ligand sequences or receptor sequences, respectively. Next, we simulated the process of protein–protein interactions using a cross-attention mechanism to facilitate information transferring and updating between ligand and receptor. The cross attention mechanism can be represented by the following formula:

$$\beta_{\mathrm{li},rj} = \mathbf{softmax}\left(\frac{1}{\sqrt{d}}\mathbf{Linear}\left(s_{\mathrm{li}}\right)\cdot\mathbf{Linear}\left(s_{rj}\right)^{T}\right) \tag{7}$$

$$h'_{\mathrm{li}} = \beta_{\mathrm{li},rj}\mathbf{Linear}\left(s_{\mathrm{li}}\right) \tag{8}$$

Where s represents sequence features, $\beta_{\mathrm{li},rj}$represents the attention of the i-th residue in the ligand to the j-th residue in the receptor, and similarly, the attention of the i-th residue in the receptor to the j-th residue in the ligand can be expressed as $\beta_{ri,lj}$.

In order to enable the sequence model to learn structural information, we proposed a novel "Alignment" method for training the model that was more lightweight and also simplified the model training process, allowing the model to extract as much structure-related features as possible and to approach the data distribution of the latent vector in the structural model more closely. "Alignment" consist of cosine similarity and mean square error, defined as:

$$L_{\mathrm{align}} = \frac{\chi_{\mathrm{structure}}\cdot\chi_{\mathrm{sequence}}}{\max\left(\left|\chi_{\mathrm{structure}}\right|_2,\epsilon\right)\cdot\max\left(\left|\chi_{\mathrm{sequence}}\right|_2,\epsilon\right)} + \left(\chi_{\mathrm{structure}}-\chi_{\mathrm{sequence}}\right)^2 \tag{9}$$

Where χ is the feature vector before inputting into the multimodal prediction head. The purpose of this design is to ensure that the direction of the feature vectors is as uniform as possible and the magnitude of the modulus is close, with features extracted solely from sequence information aligning with those extracted from structural information, thereby enhancing the predictive performance of the sequence model. To provide further guidance to "Alignment," the architecture and weights of the multimodal prediction head of PPBind-3D are transferred to PPBind-1D. The learning objective of PPBind-1D can be defined as a minimization objective function as the sum of L_{align} term and Equation (6). DIPS-Plus is used only for the Alignment stage of PPBind1D. For samples without labels, the term corresponding to Equation (6) is set to zero, and only the Alignment loss Equation (9) is applied.

Integration of heterogeneous data

Monotonic Neural Network (MNN) are designed to represent monotonic mappings of the form$\mathbf{y} = F\left(x;\theta\right)$, where $x, y \in \mathbb{R}$. Such mappings strictly preserve the monotonic relationship between $\mathbf{x}$ and $\mathbf{y}$, typically enforcing a positive correlation.

Wehenkel and Louppe [18] proposed the Unconstrained Monotonic Neural Network (UMNN) architecture, in which a monotonic transformation is obtained by integrating a strictly positive derivative function $\mathbf{f}\left(\mathbf{t};\theta\right)$:

$$F\left(x;\theta\right) = \int_0^x f\left(t;\theta\right),\mathbf{dt} + F\left(0;\theta\right) \tag{10}$$

where $f\left(t;\theta\right) > 0$ for all t, and $F\left(0;\theta\right)$ is a constant offset. In practice, $f\left(t;\theta\right)$ is parameterized by a feedforward Multi-Layer Perceptron (MLP). To guarantee strict positivity, the final layer employs an Exponential Linear Unit (ELU) activation function, followed by addition of one. The integral in $F\left(x;\theta\right)$ is then evaluated numerically using the Clenshaw–Curtis quadrature method. The implementation details for both forward computation and backward differentiation are described in the official repository https://github.com/AWehenkel/UMNN.

Incorporating heterogeneous data into our model relies critically on the MNN architecture. By constraining the integrand $f\left(t;\theta\right)$ to remain strictly positive—implemented as $\mathbf{ELU}(x)+1$, we embed monotonicity directly into the network through numerical integration.

We next formalize how this principle is applied in our model. First, the monotonic relationship between quantities must be defined. For example, a smaller binding free energy ($\Delta\mathbf{G}$) indicates stronger binding affinity, while a higher enrichment ratio corresponds to higher affinity. To capture the direction of monotonicity, we introduce a coefficient $\lambda \in -1,+1$, where $\lambda = -1$ denotes a negative correlation between enrichment ratios and $\Delta\mathbf{G}$.

The PPBind model ends with an MLP that outputs predicted scalar values for different data types. For each heterogeneous dataset (e.g. each DMS dataset), we assign a separate MNN module $F\left(\cdot;\theta\right)$, which takes as inputs the predicted $\Delta\mathbf{G}$ value from PPBind $y\prime_{\Delta\mathrm{G}}$ and the corresponding prediction from the heterogeneous dataset $y\prime_{\mathbf{heter}}$. The MNN output is thus formulated as:

$$\hat{y}\prime_{\mathbf{heter}} = F\left(y\prime_{\mathbf{heter}};\theta\right) \tag{11}$$

where θ denotes the learnable parameters of the MNN.

Consistent with Equation (6) in the main paper, the total loss can be expressed as the sum of two mean squared errors (MSEs):

$$\mathcal{L}\mathrm{total} = \underbrace{\mathbf{MSE}\left(y',y\right)}_{\text{Regression task loss}} + \underbrace{\mathbf{MSE}\left(\lambda y,F\left(y\prime_{\mathbf{heter}};\theta\right)\right)}_{\text{Monotonic alignment loss}} \tag{12}$$

where the first term measures the prediction error of the PPBind model, and the second enforces consistency between the monotonic transformation of heterogeneous data and the primary affinity metric. This joint loss enforces both accurate prediction of absolute values and preservation of the monotonic relationship across heterogeneous data modalities.

Results

Compare different ways of partitioning data

We evaluated the partition performance of three methods, namely the proposed partition method, partitioning according to PDB codes, and partitioning according to sample randomization. The minimum, maximum and average Euclidean distances between each fold of data were calculated shown as Fig. 2.

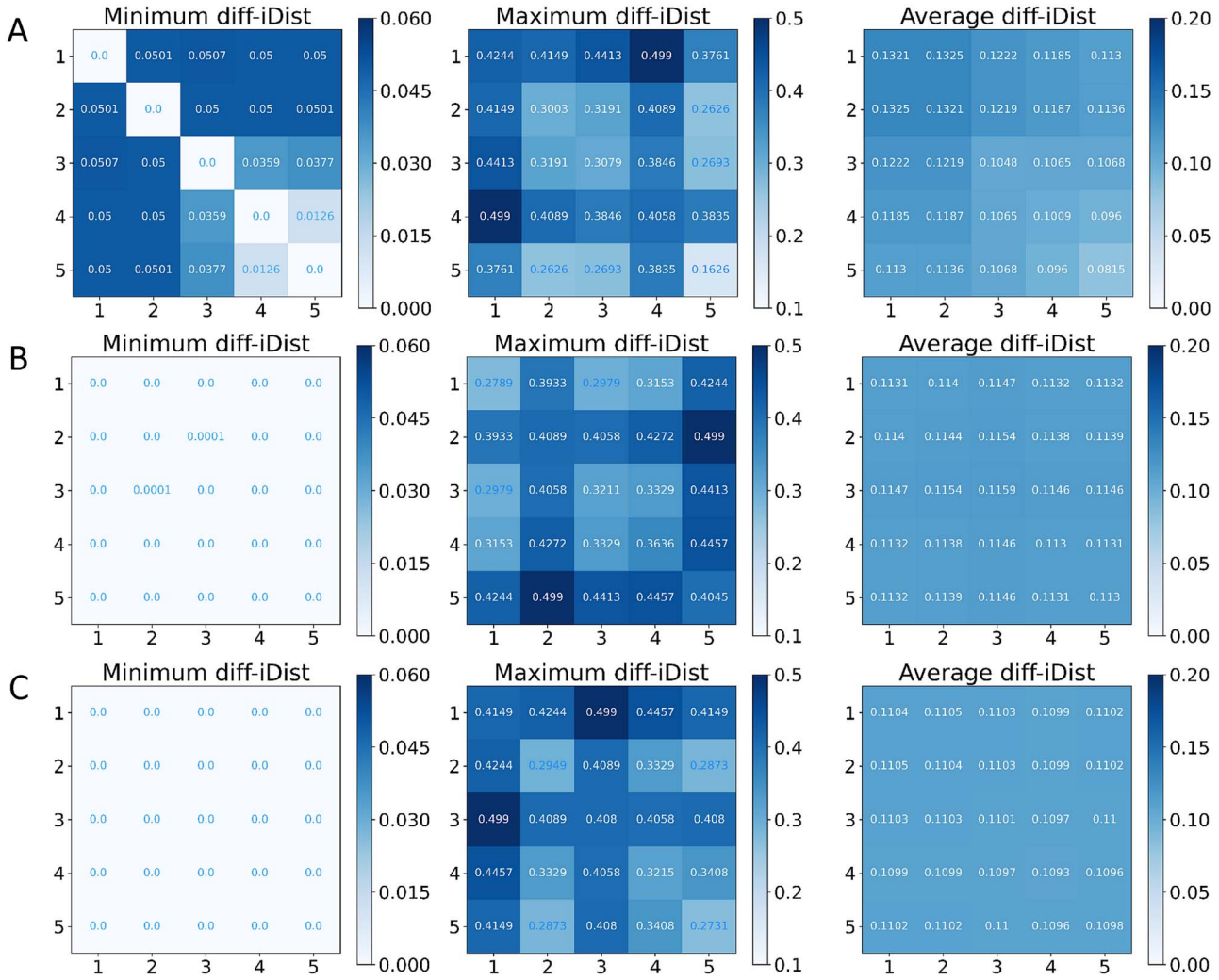


Figure 2. Our proposed strict partitioning method (A); randomly partitioning according to PDB (B); randomly partitioning according to samples (C).

Observing the distribution of the minimum, there were similar complexes between different folds in both randomized divisions. In our proposed strict division method, there were no similar complexes between each fold. From the average distance plot, the two randomized division methods were compared with our proposed method, which divides as many similar complexes as possible in the same fold, because the average distance per fold of the randomized division method is very close to the average distance per fold of the randomized division method, whereas the average distance per fold of our method is somewhat different, and the value is both large and small.

Model evaluation

We trained and tested PPBind-3D by the PPB-Affinity dataset and the DMS-Het dataset, where the DMS-Het dataset was for model training only and the PPB-Affinity dataset was for cross-validation. Under strict data partitioning, the five-fold cross-validation performance of PPBind-3D on the PPB-Affinity dataset was showed as Fig. 3A and B illustrates the performance of PPBind-3D when trained and tested at an 8:2 ratio with random partitioning.

Similarly, PPBind-1D has been validated using both strict and random partitioning, in a manner consistent with the validation of PPBind-3D. However, the training set for PPBind-1D additionally included DIPS-Plus. After a simple filtering of the PPB-Affinity dataset (as described in section "*PPBind-1D*"), PPBind-1D was trained based on the principle of sequence-structure-alignment. The performance of PPBind-1D was as follows in Fig. 3C and D.

The test metrics for random partitioning are significantly higher than those for strict partitioning. This is because the random partitioning introduces data leakage, which artificially boosts the test metrics. In contrast, strict partitioning avoids data leakage and provides a more accurate evaluation of the model's generalization performance. These results demonstrate the superiority of our proposed data partitioning method. Furthermore, the Pearson and Spearman correlations of our PPBind-3D and PPBind-1D models are both greater than 0.6 under the strict partitioning of data, indicating that our model architecture is preeminent.

To better evaluate the performance of our model, we used the following models as baseline comparisons: PRODIGY [19], which predicts affinity based on intermolecular contacts and properties derived from non-interface surfaces; DFIRE [20], which predicts affinity based on a potential function using the ideal gas state as a physical reference; CP_PIE [21], a mathematical programming-based approach for protein–protein docking filtering and scoring that utilizes residue contacts and overlap areas; ISLAND [22], which employs sequence-based features and a machine learning model to predict affinity; PPI-Affinity, using aggregation operators to extract spatially-informed features for affinity prediction [23];

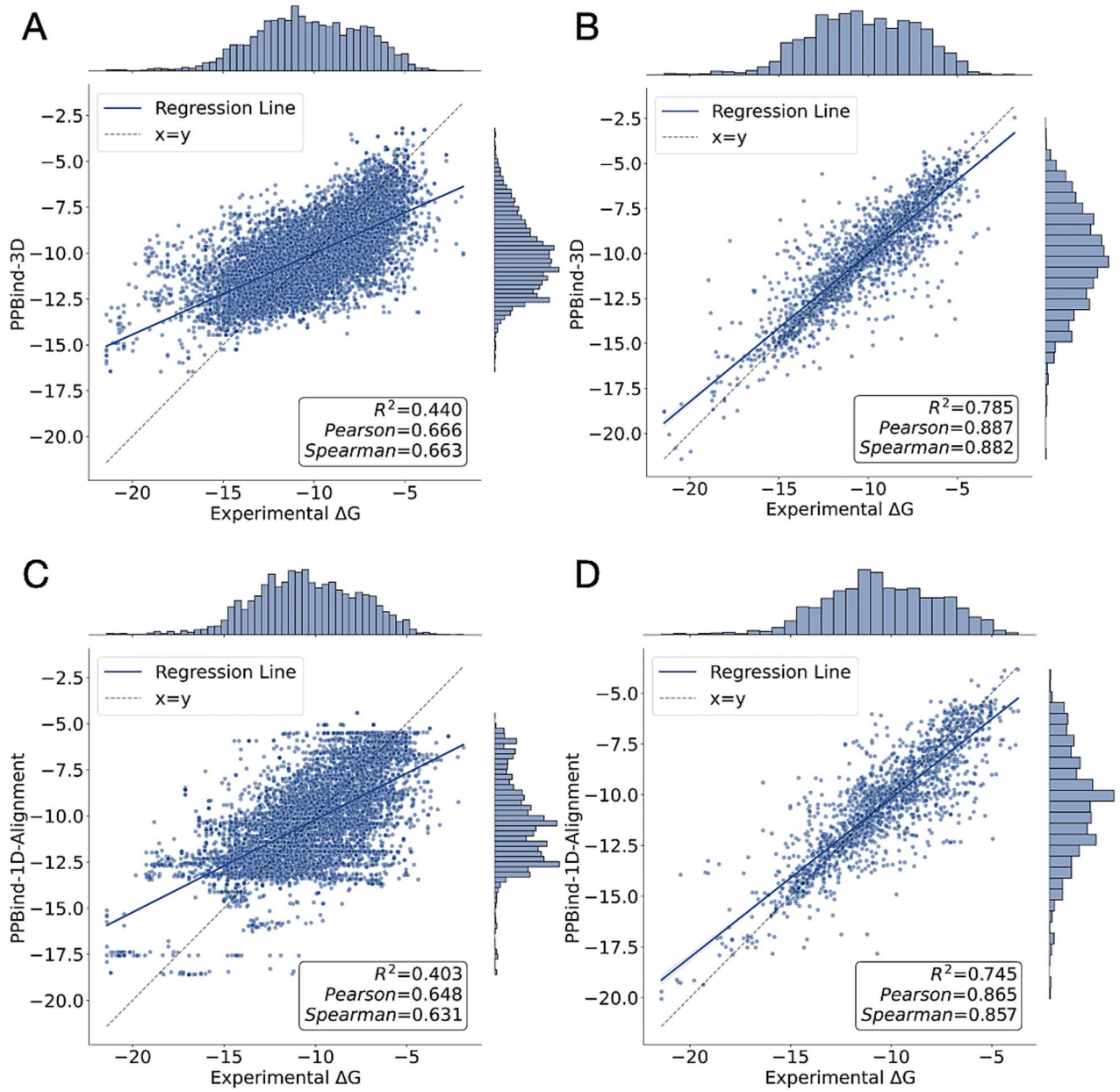


Figure 3. The performance of PPBind-3D under strict data partitioning (A) and under random data partitioning (B); the performance of PPBind-1D under strict data partitioning (C) and under random data partitioning (D).

and ProBAN [24], which utilizes complex structural data and a deep 3D convolutional neural network to predict affinity. The test data and baseline model metrics were sourced from ProBAN. The test data consists of two components: test set 1, which includes 126 samples, and test set 2, which includes 83 samples, with all samples in set 2 being protein complexes composed of two chains. Both sets are subsets of those in PDBbind v2020 [25]. Additionally, all PDB entries identified in the test data were excluded, and PPBind-3D was retrained. The resulting performance are illustrated in Table 1, from which it can be seen that PPBind-3D outperforms other algorithms in all aspects, demonstrating its superior performance.

Visualization of alignment

To observe the effectiveness of "Alignment," the feature representations are visualized by dimensionality reduction using the following steps: (i) Extract high-dimensional feature representations of the training samples using the PPBind-3D model, and fit a dimensionality reduction function F_U using the UMAP (Uniform Manifold Approximation and Projection) algorithm [26]. (ii) Extract high-dimensional feature representations of the validation samples using the PDBind-3D, PDBind-1D, and PPBind-1D-w/o Align models, respectively. (iii) Individually project the three sets of high-dimensional feature representations onto a 2D plane using the fitted F_U function and visualize them.

By comparing the three dimensionality reduction visualizations (Fig. 4A–C), it can be observed that the dimensionality reduction representation of the PPBind-1D model retains a similar data topological structure to that of the PPBind-3D model, whereas the PPBind-1D model without "Alignment" exhibits a scattered state. This indicates that the representations extracted by the PPBind-1D model are similar to those of the PPBind-3D model, suggesting that it is possible to enhance the prediction accuracy of

Table 1. Comparison between PPBind-3D and other models.

Method	Test set 1			Test set 2		
	Pearson	MAE (kcal/mol)	RMSE (kcal/mol)	Pearson	MAE (kcal/mol)	RMSE (kcal/mol)
PRODIGY[a]	–	–	–	0.28	2.47	3.52
DFIRE	–	–	–	0.08	25.05	29.17
CP_PIE	–	–	–	−0.1	10.9	11.27
ISLAND	–	–	–	0.28	2.3	2.85
PPI-Affinity	–	–	–	0.49	1.83	2.4
ProBAN	0.60	1.60	1.95	0.55	1.75	2.28
PPBind-3D(ours)	0.626	1.482	1.898	0.559	1.647	2.210

[a]Indicates non-machine learning method. The underline represents the optimal model.

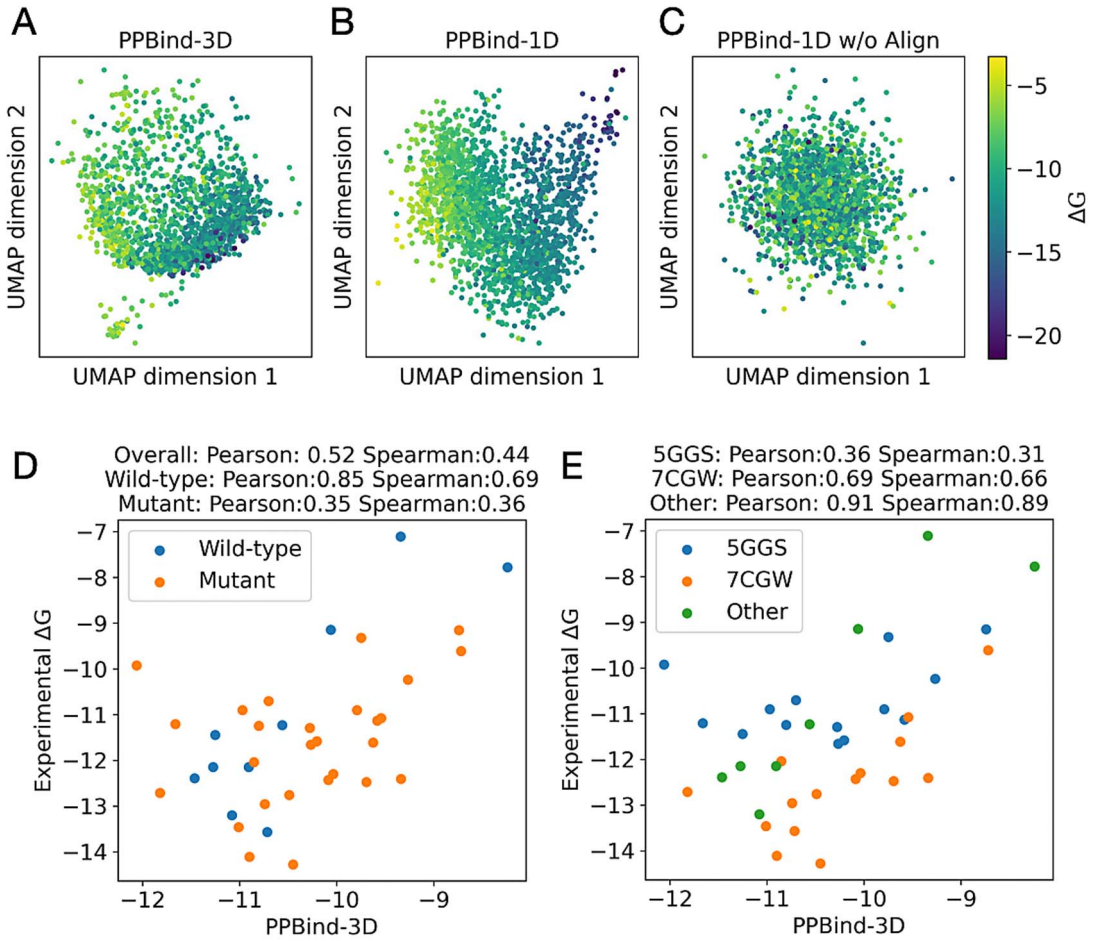


Figure 4. Visualize the representation of three models (A-C) and display the prediction results of the model under real structure based on the type (D) and based on the PDB code (E).

the PPBind-1D model to the level of the PPBind-3D model through our proposed "Alignment" method.

Virtual screening

To validate the performance of the model in virtual screening of protein affinity, a series of three case studies was conducted. In order to enhance the precision of the screening outcomes, this section employs the models that has been trained through the random partitioning of the data set. At the same time, we also compared three cases with the training data, including the Euclidean distance represented by iDist, PDB ID and its descriptive information. For details, please refer to Supplementary S4.

Case1: Based on PPBind-3D, predict affinity from real structure.

We have compiled a set of recent experimental data [27] on affinity and complex structures, which have not yet been included in the PPB-Affinity dataset. This dataset involves affinity K_d values and complex structures for various antibodies binding to antigens such as CTLA-4, PD-1, and PD-L1, as well as their mutants. PPBind-3D, was used to predict the affinities of these antibody–antigen complexes. The Fig. 4D is based on whether the affinity originates from a mutant or not, while the Fig. 4E is based on the PDB code. Overall, this case shows that PPBind-3D can be used for affinity prediction and virtual screening when real structural information is available.

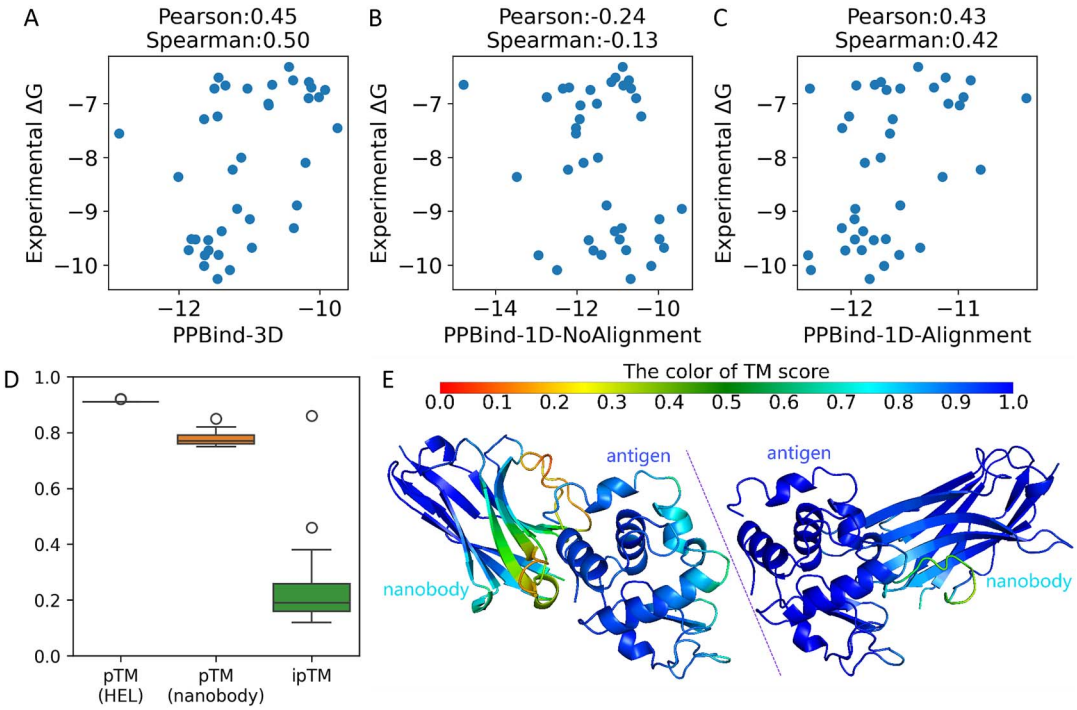


Figure 5. Correlation between the affinities predicted by (A) model PPBind-3D, (B) model PPBind-1D-NoAlignment, (C) model PPBind-1D-Alignment and the actual affinities in Case 2; (D) Box plot of the ipTM and pTM for the complex structures predicted by AlphaFold3 in Case 2; (E) Two structures predicted by AlphaFold3 in Case 2. The left side has lowest ipTM, and the right side has highest ipTM.

Case2: Anti-hen egg lysozyme antibodies affinity ranking

To assess the performance of the proposed models in virtual affinity screening without real structures, we conducted Case 2. Imitate MT-TopLap [28, 29] to predict protein structure through AlphaFold3 and further predict its affinity. We obtained a set of 38 K_d values for different nanobodies binding to HEL, as measured by Porebski et al. [30] through experiments. We then predicted the complex structures of each nanobody with HEL using AlphaFold3, followed by affinity predictions with PPBind-3D and PPBind-1D.

As shown in Fig 5A–C, the affinity predictions for structures using PPBind-3D based on AlphaFold3 were found to be of a comparable level to those using PPBind-1D-Align. Conversely, PPBind-1D-No-Align performed significantly less well than PPBind-1D-Align. It was observed that the predicted structures in this batch exhibited a general low ipTM (Fig. 5D), indicating potential inaccuracy in the structure prediction of the interface region. Furthermore, it was determined that distinct complexes exhibit disparate epitopes (Fig. 5E), which markedly influence the affinity strength and ultimately result in the failure of PPBind-3D prediction.

Case3: Anti-PD-L1 antibodies affinity ranking

The antibodies and the affinity values derived from Brzostek [31], Gao [32], Guan [33], He [34], Hong [35], Rajasekaran [36], Tan [37, 38] were used to validate ours model in the case. Similar to case2, only the sequences of the proteins are known. We also used AlphaFold3 to predict complex structures and compared the predictions of the three models.

As shown in Fig. 6A–C, the results of the predictive modeling demonstrate that PPBind-3D is the least effective; PPBind-1D-No-Align is the second-best performer, and PPBind-1D-Align is the most accurate. We found that even with high ipTM (Fig. 6D), docking posture and epitopes and paratopes varied between individuals (Fig. 6E), which we believe contributes to affinity prediction.

Case4: Anti-human TIGIT antibodies affinity ranking

Previous work by Agarwal [39] released a large volume of affinity data acquired via AlphaSeq measurements, along with 12 sets of affinity data characterized using biolayer interferometry (BLI). We fine-tuned the PPbind-1D model with the AlphaSeq dataset and validated the model against the experimental BLI data. The detailed procedures are presented in Supplementary Material Section S6: Evaluation on the AlphaBind Dataset. The PPbind-1D model achieved a prediction accuracy of 75% (predictions within a three-fold deviation from the true values), demonstrating strong generalization performance. The detailed prediction results are shown as Table 2 and the detailed procedures are presented in Supplementary S6.

Ablation study

To investigate the impact of data partitioning methods, network design, and training strategies on model performance, we conducted 16 ablation experiments on the PPB-Affinity dataset, as summarized in the Table 3. The overall metrics were derived from the complete PPB-Affinity dataset, while the Per-Structure metrics were obtained from samples with more than 10 mutants in the PPB-Affinity dataset. In addition, we explored another architecture, PPBind-Mix, which takes the receptor structure and the ligand sequence as inputs. In the ablation study, we primarily examined the impact of Alignment on PPBind-Mix, and also applied PPBind-Mix to Case 2 and Case 3 (Supplementary S5).

The structural model PPBind-3D outperforms the sequence model PPBind-1D

To rigorously assess the effects of strict versus random data partitioning on model performance, we performed 16 experiments. The strict partitioning employed five-fold cross-validation (details can be found in the Methods section under Data Partitioning), while

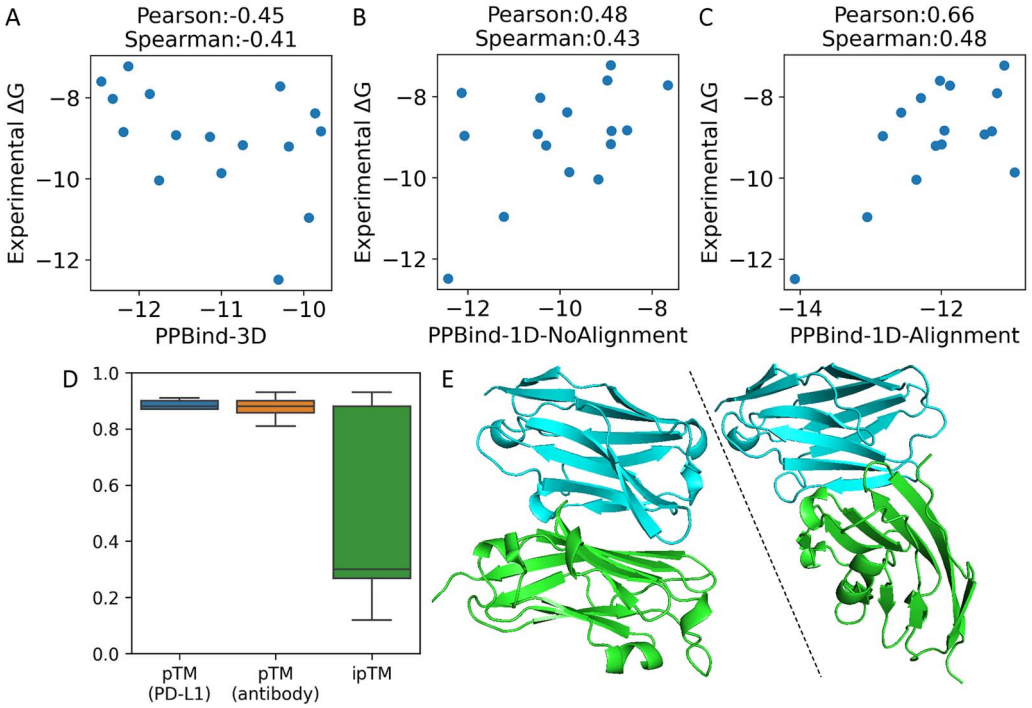


Figure 6. Correlation between the affinities predicted by (A) model PPBind-3D, (B) model PPBind-1D-NoAlignment, (C) model PPBind-1D-Alignment and the actual affinities in Case 3; (D) Box plot of the ipTM and pTM for the complex structures predicted by AlphaFold3 in Case 3; (E) The two structures for which AlphaFold3 predicted the highest ipTM scores in Case 3, separated by the dashed line. The upper chains are PD-L1, and the lower chains are antibody.

the random partitioning used an 80:20 split between the training and validation sets without cross-validation. In the case of random data partitioning, PPBind-3D significantly outperformed PPBind-1D, particularly in the Per-Structure metrics. This aligns with the intuition that structural information is more beneficial for predicting binding affinity than sequence information, especially in capturing the affinity differences induced by mutations. Conversely, under strict data partitioning, both models exhibited a notable decline in performance. However, PPBind-3D still maintained superior performance over PPBind-1D, particularly in the Per-Structure metrics. This comparative analysis suggests that random data partitioning likely introduces data leakage, resulting in inflated performance evaluations to some extent. Conversely, it demonstrates that our proposed strict data partitioning method can substantially mitigate the risk of data leakage.

MMTL enhances the models' generalization performance

A comparison between Experiment 4 and 5 reveals that employing MMTL (utilizing the DMS-Het dataset) results in a significant improvement in overall model performance. This indirectly supports the reliability of the non-dG affinity data. Considering that incorporating more data for training is beneficial for improving the model's generalization performance, we advocate for the use of MMTL.

S&M-attention outperforms all-attention

In the Geometric Module, we conducted ablation studies (Experiment 2, 5) to analyze the impact of different attention mechanisms. "All-attention" refers to a method that does not distinguish between the receptor and ligand, treating the complex structure as a whole for attention calculation and feature updating. In contrast, "S&M-attention" treats the receptor and ligand as individual entities. It first computes self-attention within each entity to update their features, followed by mutual attention to capture interactions between the receptor and ligand, further refining their respective feature representations. By comparing Experiment 2 and 5, it is clear that S&M-attention significantly outperforms All-attention.

The alignment mechanism enhances PPBind-1D

To validate the effectiveness of our proposed "Alignment" method for PPBind-1D, we conducted Experiment 6, 7, 8 and 12, 13, 14. In Experiment 6 and 12, the models were trained directly without using Alignment. Experiment 5 and 13 employed the Alignment method but did not utilize unlabeled samples. Experiment 6 and 7 aligned with Experiment 5, while Experiment 12 and 13 aligned with Experiment 11. Under strict data partitioning (Experiment 6, 7, and 8), it is evident that models using the Alignment mechanism outperform those trained directly across all metrics.

Additionally, incorporating unlabeled samples (DIPS-Plus Dataset) further improves the model's performance, bringing it closer to PPBind-3D. In contrast, Experiment 12, 13, 14 show that the model incorporating unlabeled samples for alignment performed the worst. This is due to data leakage between the PPB-Affinity training and test sets under random data partitioning, leading to inflated test set performance. Furthermore, this underscores the critical importance of strict data partitioning for the accurate assessment of model performance in PPB affinity prediction, while our model continues to demonstrate robust performance under rigorous partitioning protocols.

Discussion

To address the challenges of data leakage, data scarcity, data heterogeneity, and structure-unavailability in PPB affinity prediction, we herein propose a comprehensive framework. First,

Table 2. Detailed prediction performance of PPBind-1D after fine-tuning across 12 independent experimental datasets obtained via BLI measurements.

ID	PPBind-1D Prediction KD (nM)	KD (nM)
AAB-PP3115	1.776998	3.1
AAB-PP3117	2.141699	1.9
AAB-PP3120	1.898283	3.1
AAB-PP3121	1.955413	2.2
AAB-PP3124	3.663242	2.8
AAB-PP3127	3.446876	3.5
AAB-PP3129	9.627788	4
AAB-PP3130	485.231080	4.5
AAB-PP3132	114.488429	14
AAB-PP3133	676.462284	28
AAB-PP3135	211.447456	9
AAB-PP3136	1082.541540	377

a comprehensive dataset was constructed by integrating disparate PPB affinity data sources, and a partitioning strategy was introduced to significantly mitigate data leakage. Subsequently, MMTL was employed to integrate heterogeneous affinity data, thereby expanding the PPB training dataset and enhancing the generalization capability of the structure-based PPB affinity prediction model, PPBind-3D. Furthermore, structure-unavailability was addressed by developing a sequence-based affinity prediction model (PPBind-1D) through feature alignment techniques. Comparative and ablation studies demonstrate that PPBind-3D, trained under this framework, achieves superior prediction correlation and RMSE compared to conventional methods under both strict and random data partitioning settings. Notably, PPBind-1D closely approximates the predictive performance of PPBind-3D using exclusively protein sequences, offering a practical alternative when tertiary structures are unavailable. Validation across three virtual screening case studies confirms the efficacy of both models. Critically, this outcome indicates scalability within the training architectures of PPBind-3D and PPBind-1D. It is anticipated that incorporating additional experimental data will further enhance model performance.

One main finding of our study was that PPB affinity prediction models should not simply adopt random data partitioning, because different samples might share similar binding interface structures despite low sequence similarity. For instance, one research team adopted a leave-one-protein-out validation strategy to develop a model called PLNet [40], which seemed to outperform PPBind-3D in metrics like Pearson coefficient. However, this approach involves withholding one PDB structure as the validation set while using the remaining PDBs for training, which inherently risks significant data leakage. Similarly, another model named MT-TopLap also failed to actively mitigate data leakage in its partitioning of the SKEMPI dataset according to it's GitHub repository. Under random data splitting, PPBind-3D and PPBind-1D achieved Pearson coefficients of 0.887 and 0.865 on the full dataset, respectively, and 0.895 and 0.870 on the SKEMPI subset. Thus, our models are still comparable to MT-TopLap (Pearson coefficients: 0.88), although our model's purpose is different from MT-TopLap, which aimed to predict $\Delta\Delta G$. To more objectively demonstrate our models' performance, we employed clustering-based partitioning using iDist, further mitigating data leakage risks.

Another major finding was that MMTL could effectively integrate heterogeneous affinity data, enhancing the generalization capabilities of the training model by expanding the training set. Data volume is one of the primary factors influencing model performance. Most affinity data from wet-lab experiments are not presented as equilibrium dissociation constant (KD) values, but rather high-throughput metrics such as mean fluorescence intensity. To address this, we have collected as much affinity data with PDB structures (measured as KD or ΔG) as possible. While we expanded our dataset using DMS data, the core training data remains KD/ΔG values, for the following reasons: DMS typically quantifies affinity using metrics like enrichment rates or fluorescence intensity. Although DMS datasets are large, their quality is inferior to that of SPR or BLI data. DMS measurements are influenced by multiple factors, including experimental methods, conditions, and target proteins. For example, literature [30] reported a Spearman correlation of −0.697 between BLI and fluorescence intensity, highlighting discrepancies between DMS measurements and KD/ΔG values. While incorporating DMS data improves model performance, it cannot replace KD/ΔG data as the core training set. The volume of such wet-lab data is huge, yielding substantial benefits when integrated them into training via MMTL.

An evident benefit is that as the amount of data increases, the model's Pearson coefficient also improves, which was observed in a model called Graphinity [4] developed by huge amount of FoldX-generated data. Interestingly, we still found that the prediction performance of PPBind-3D was still comparable to Graphinity. For instance, with 12k FoldX-generated data, PPBind-3D achieved a Pearson coefficient of 0.504. Through log10-scaling and interpolation, we calculated Graphinity's Pearson value at 12k data points to be 0.459. Despite different task objectives, this comparison might imply PPBind-3D's superior architecture performance at comparable data scales. Furthermore, FoldX predictions exhibit systematic deviations from experimental $\Delta\Delta G$ values, as evidenced by multiple studies [14, 41] reporting Pearson coefficient below 0.4 between FoldX-predicted and experimentally measured values. Thus, cautions should be noted when using FoldX-generated data, which may imply the real-world application value of our models developed by experimental data.

The current study also has other interesting findings. For instance, compared to standard all-attention mechanisms, our proposed S&M attention mechanism accurately captures receptor-ligand interactions. It first updates monomeric protein features through self-attention, then refines them via cross-attention between binding partners, leading to superior affinity prediction. In addition, employing latent feature alignment, where PPBind-3D guides PPBind-1D model training, significantly improves the prediction accuracy of PPBind-1D. Lastly, through three-case application, we demonstrated that both PPBind-3D and PPBind-1D models effectively screened high-affinity binders from numerous candidate molecules targeting a given antigen.

It must be acknowledged that the current study exhibits limitations. Firstly, while the PPBind-3D model required a known complex structure for accurate affinity prediction, its accuracy might be compromised when the complex structure was predicted by AlphaFold 3. We advised that when the predicted interface pTM (iPTM) was high, PPBind-3D may be considered reliable for affinity prediction. Conversely, predicted structures under low iPTM should be treated with caution when using PPBind-3D model. However, with the increasing accuracy of structure prediction models like AlphaFold series, we could optimistically visualize that the workflow of PPBind-3D enabling affinity predictions even in the absence of experimental structures would be more often and reliable in the future. Another scenario in which PPBind-3D

Table 3. Results of ablation experiments.

Experiment	Data split	Network	Overall (dG)				Per-structure (dG)	
			Pearson	Spearman	R^2	MAE	Pearson	Spearman
1	Strict	PPBind-3D$^{\text{Ↄ}}$	0.582	0.593	0.288	1.876	0.378	0.343
2	Strict	PPBind-3D$^{\text{§, Δ, Ↄ}}$	0.617	0.618	0.374	1.779	0.383	0.343
3	Strict	PPBind-3D$^{\text{ↄ}}$	0.646	0.627	0.412	1.746	0.389	0.349
4	Strict	PPBind-3D$^{\text{§, ↄ}}$	0.647	0.639	0.406	1.773	**0.397**	**0.372**
5	Strict	PPBind-3D$^{\text{§, Δ, ↄ}}$	**0.666**	**0.663**	**0.440**	**1.684**	0.380	0.362
6	Strict	PPBind-1D$^{\text{§, Δ}}$	0.594	0.587	0.229	1.932	0.062	0.050
7	Strict	PPBind-1D$^{\text{§, Δ, ∧}}$	0.626	0.606	0.311	1.847	0.004	0.004
8	Strict	PPBind-1D$^{\text{§, Δ, ∧, ↓}}$	0.648	0.631	0.403	1.69	0.004	0.004
9	Strict	PPBind-Mix$^{\text{§, Δ}}$	0.627	0.616	0.384	1.837	0.116	0.115
10	Strict	PPBind-Mix$^{\text{§, Δ, ∧, ↓}}$	0.632	0.623	0.395	1.744	0.159	0.152
11	Random	PPBind-3D$^{\text{§, Δ, Ↄ}}$	**0.887**	**0.882**	**0.785**	**0.898**	**0.634**	**0.607**
12	Random	PPBind-1D$^{\text{§, Δ}}$	0.868	0.862	0.748	0.958	0.463	0.441
13	Random	PPBind-1D$^{\text{§, Δ, ∧}}$	0.876	0.866	0.763	0.908	0.443	0.412
14	Random	PPBind-1D$^{\text{§, Δ, ∧, ↓}}$	0.865	0.857	0.745	0.966	0.336	0.319
15	Random	PPBind-Mix$^{\text{§, Δ}}$	0.738	0.732	0.544	1.535	0.038	0.047
16	Random	PPBind-Mix$^{\text{§, Δ, ∧, ↓}}$	0.810	0.807	0.654	1.277	0.072	0.059

§: DMS Data, Δ: MMTL, Ↄ: All Attention, ↄ: SM Attention, ∧: Align, ↓: Unlabeled Sample, Bold indicates the best model under strict or random data splitting.

may also demonstrate its unique value is as follows: by leveraging RDE, PPBind-3D enables affinity prediction for mutant variants directly from wild-type structures and mutation annotations, bypassing the need for computationally expensive mutant structure generation. This capability is pivotal for high-throughput variant screening in antibody engineering. Secondly, although feature alignment enhanced the overall predictive performance of the PPBind-1D model, it failed to significantly improve per-structure prediction performance. Thirdly, it should be noted that there are still noticeable discrepancies between the experimental values and predicted values of our models. This is because accurate numerical prediction of binding affinity remains a major challenge in the field, due to factors such as data noise, variations in experimental conditions, and limited training data. In practical applications such as virtual screening and candidate molecule prioritization, however, we are more concerned with the model's ability to capture relative affinity (i.e. its capacity to correctly distinguish between high- and low-affinity molecules).

Acknowledgements

We thank Jax Luo from Neurological Institute of Cleveland Clinic in suggestions for figure layout and Li Ye from Macau University of Science and Technology for assistance in the last revision.

Author contributions

Peiyi Chen (Formal analysis [lead], Visualization [lead], Writing—original draft [lead]), Huaqing Liu (Investigation [lead], Methodology [lead], Validation [lead]), Shuxian Zhou (Data curation [lead]), Lunji Wang (Software [supporting]), Mao Pang (Writing—review & editing [equal]), Jing Zhao (Writing—review & editing [equal]), Guoxin Fan (Conceptualization [lead], Supervision [lead], Writing—review & editing [lead]), and Lanqing Han (Resources [lead], Funding acquisition [lead], Project administration [lead])

Supplementary material

Supplementary material is available at *Antibody Therapeutics* online.

Conflicts of interest

Cyagen Biomodels (Guangzhou) Co., Ltd. is a subsidiary of Cyagen Biosciences (Suzhou) Inc. Lanqing Han is the founder and CEO of Cyagen Biosciences (Suzhou) Inc. Peiyi Chen, Shuxian Zhou, Huaqing Liu and Lunji Wang are employees of Cyagen Biomodels (Guangzhou) Co., Ltd.

Funding

This work was supported by Guangzhou Key R&D Program Project [2023B03J1345] to L.H.; the Guangdong Basic and Applied Basic Research Foundation (2026A1515011835) to G.F. This work is also funded by Cyagen Biomodels (Guangzhou) Co., Ltd.

Data availability

Supplementary data to this article can be found online within the manuscript.

All data in this article are from the public database PPB-Affinity Dataset (https://zenodo.org/records/13054646).

Codes for PPBind-3D and PPBind-1D were available at https://github.com/ChenPy00/PPBind.

Meanwhile, the PPBind-3D model has also been integrated into a web server, accessible at https://rddc.tsinghua-gd.org/tool/dg-prediction.

Animal research statement

Not applicable.

Ethics and consent statement

Consent was not required for this study.

References

1. MacBeath G. Protein microarrays and proteomics. *Nat Genet* 2002;**32**:526–32. https://doi.org/10.1038/ng1037.
2. Cameron C, Peacor B, Oien N. *et al*. Octet® Potency Assay: Development, Qualification and Validation Strategies. Application

Note 26 (Sartorius, 2021). https://www.sartorius.com/download/789382/octet-potency-assay-development-validation-strategies-applic-2–data.pdf.
3. Kortemme T. De novo protein design—from new structures to programmable functions. *Cell* 2024;**187**:526–44. https://doi.org/10.1016/j.cell.2023.12.028.
4. Hummer AM, Schneider C, Chinery L. *et al.* Investigating the volume and diversity of data needed for generalizable antibody–antigen ΔΔG prediction. *Nat Comput Sci* 2025;**5**:635–47. https://doi.org/10.1038/s43588-025-00823-8.
5. Liu H, Chen P, Zhai X. *et al.* PPB-affinity: protein-protein binding affinity dataset for AI-based protein drug discovery. *Sci Data* 2024;**11**:1316. https://doi.org/10.1038/s41597-024-03997-4.
6. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. *Nat Methods* 2014;**11**:801–7. https://doi.org/10.1038/nmeth.3027.
7. Sill, J. Monotonic Networks. In: *Advances in Neural Information Processing Systems*, Vol. **10**, MIT Press, 1997. https://proceedings.neurips.cc/paper_files/paper/1997/file/83adc9225e4deb67d7ce42d58fe5157c-Paper.pdf.
8. Wang G, Liu X, Wang K. *et al.* Deep-learning-enabled protein-protein interaction analysis for prediction of SARS-CoV-2 infectivity and variant evolution. *Nat Med* 2023;**29**:2007–18. https://doi.org/10.1038/s41591-023-02483-5.
9. Bushuiev A, Bushuiev R, Kouba P. *et al.* Learning to design protein-protein interactions with enhanced generalization. arXiv 2024; arXiv:2310.18515.
10. Karypis G, Kumar V. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J Sci Comput* 1998;**20**:359–92. https://doi.org/10.1137/S1064827595287997.
11. Chan KK, Dorosky D, Sharma P. *et al.* Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* 2020;**369**:1261–5. https://doi.org/10.1126/science.abc0870.
12. Starr TN, Greaney AJ, Hannon WW. *et al.* Shifting mutational constraints in the SARS-CoV-2 receptor-binding domain during viral evolution. *Science* 2022;**377**:420–4. https://doi.org/10.1126/science.abo7896.
13. Morehead A, Chen C, Sedova A. *et al.* DIPS-plus: the enhanced database of interacting protein structures for interface prediction. *Sci Data* 2023;**10**:509.
14. Luo S, Su Y, Wu Z. *et al.* Rotamer Density Estimator is an Unsupervised Learner of the Effect of Mutations on Protein-Protein Interaction. bioRxiv 2023. https://doi.org/10.1101/2023.02.28.530137.
15. Shan S, Luo S, Yang Z. *et al.* Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization. *Proc Natl Acad Sci U S A* 2022;**119**. e2122954119.
16. Wehenkel A, Louppe G. Unconstrained Monotonic Neural Networks. CoRR 2019. https://doi.org/10.48550/arXiv.1908.05164.
17. Lin Z, Akin H, Rao R. *et al.* Language models of protein sequences at the scale of evolution enable accurate structure prediction. bioRxiv 2022. https://doi.org/10.1101/2022.07.20.500902.
18. Wehenkel A, Louppe G. Unconstrained Monotonic Neural Networks. In: *Advances in Neural Information Processing Systems*, Vol. **32**, Curran Associates, Inc., 2019. https://proceedings.neurips.cc/paper_files/paper/2019/file/2a084e55c87b1ebcdaad1f62fdbbac8e-Paper.pdf.
19. Xue LC, Rodrigues JP, Kastritis PL. *et al.* PRODIGY: a web server for predicting the binding affinity of protein–protein complexes. *Bioinformatics* 2016;**32**:3676–8. https://doi.org/10.1093/bioinformatics/btw514.
20. Liu S, Zhang C, Zhou H. *et al.* A physical reference state unifies the structure-derived potential of mean force for protein folding and binding. *Proteins* 2004;**56**:93–101. https://doi.org/10.1002/prot.20019.
21. Ravikant DVS, Elber R. PIE—efficient filters and coarse grained potentials for unbound protein–protein docking. *Proteins* 2010;**78**:400–19. https://doi.org/10.1002/prot.22550.
22. Abbasi WA, Yaseen A, Hassan FU. *et al.* ISLAND: in-silico proteins binding affinity prediction using sequence information. *BioData Min* 2020;**13**:20. https://doi.org/10.1186/s13040-020-00231-w.
23. Romero-Molina S, Ruiz-Blanco YB, Mieres-Perez J. *et al.* PPI-affinity: a web tool for the prediction and optimization of protein–peptide and protein–protein binding affinity. *J Proteome Res* 2022;**21**:1829–41. https://doi.org/10.1021/acs.jproteome.2c00020.
24. Bogdanova EA, Novoseletsky VN. ProBAN: neural network algorithm for predicting binding affinity in protein–protein complexes. *Proteins* 2024;**92**:1127–36. https://doi.org/10.1002/prot.26700.
25. Wang R, Fang X, Lu Y. *et al.* The PDBbind database: collection of binding affinities for protein−ligand complexes with known three-dimensional structures. *J Med Chem* 2004;**47**:2977–80. https://doi.org/10.1021/jm030580l.
26. McInnes L, Healy J, Melville J. *et al.* UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv 2020; arXiv:1802.03426.
27. Kang-Pettinger T, Walker K, Brown R. *et al.* Identification, binding, and structural characterization of single domain anti-PD-L1 antibodies inhibitory of immune regulatory proteins PD-1 and CD80. *J Biol Chem* 2023;**299**:102769. https://doi.org/10.1016/j.jbc.2022.102769.
28. Wee J, Wei G-W. Rapid response to fast viral evolution using AlphaFold 3-assisted topological deep learning. *Virus Evol* 2025;**11**:veaf026. https://doi.org/10.1093/ve/veaf026.
29. Wee J, Wei G-W. Evaluation of AlphaFold 3's protein–protein complexes for predicting binding free energy changes upon mutation. *J Chem Inf Model* 2024;**64**:6676–83. https://doi.org/10.1021/acs.jcim.4c00976.
30. Porebski BT, Balmforth M, Browne G. *et al.* Rapid discovery of high-affinity antibodies via massively parallel sequencing, ribosome display and affinity screening. *Nat Biomed Eng* 2024;**8**:214–32. https://doi.org/10.1038/s41551-023-01093-3.
31. Brzostek J, Gascoigne NR, Rybakin V. Cell type-specific regulation of immunological synapse dynamics by B7 ligand recognition. *Front Immunol* 2016;**7**:24. https://doi.org/10.3389/fimmu.2016.00024.
32. Gao H, Cai H, Liu J. *et al.* Structure of CTLA-4 complexed with a pH-sensitive cancer immunotherapeutic antibody. *Cell Discov* 2020;**6**:79. https://doi.org/10.1038/s41421-020-00202-9.
33. Guan J, Liu H, Chai Y. *et al.* Characterization of the high-affinity anti-CTLA-4 monoclonal antibody JS007 for immune checkpoint therapy of cancer. *MAbs* 2023;**15**:2153409. https://doi.org/10.1080/19420862.2022.2153409.
34. He M, Chai Y, Qi J. *et al.* Remarkably similar CTLA-4 binding properties of therapeutic ipilimumab and tremelimumab antibodies. *Oncotarget* 2017;**8**:67129–39. https://doi.org/10.18632/oncotarget.18004.
35. Hong Y, Feng Y, Sun H. *et al.* Tislelizumab uniquely binds to the CC' loop of PD-1 with slow-dissociated rate and complete PD-L1 blockage. *FEBS Open Bio* 2021;**11**:782–92. https://doi.org/10.1002/2211-5463.13102.
36. Rajasekaran N, Wang X, Ravindranathan S. *et al.* Toripalimab, a therapeutic monoclonal anti-PD-1 antibody with high binding affinity to PD-1 and enhanced potency to activate human T cells. *Cancer Immunol Immunother* 2024;**73**:60. https://doi.org/10.1007/s00262-024-03635-3.

37. Tan S, Zhang H, Chai Y. *et al.* An unexpected N-terminal loop in PD-1 dominates binding by nivolumab. *Nat Commun* 2017;**8**:14369. https://doi.org/10.1038/ncomms14369.
38. Tan S, Liu K, Chai Y. *et al.* Distinct PD-L1 binding characteristics of therapeutic monoclonal antibody durvalumab. *Protein Cell* 2018;**9**:135–9. https://doi.org/10.1007/s13238-017-0412-8.
39. Agarwal AA, Harrang J, Noble D. *et al.* AlphaBind, a domain-specific model to predict and optimize antibody-antigen binding affinity. *MAbs.* 2025;**17**:2534626. https://doi.org/10.1080/19420862.2025.2534626.
40. Xu X, Wang C, Wei GW. *et al.* PLNet: persistent Laplacian neural network for protein-protein binding free energy prediction. *Protein Sci* 2025;**34**:e70377. https://doi.org/10.1002/pro.70377.
41. Liu S, Zhu T, Ren M. *et al.* Predicting mutational effects on protein-protein binding via a side-chain diffusion probabilistic model. In: *Advances in Neural Information Processing Systems*, Vol. **36**, Curran Associates, Inc., 2023, 48994–9005.

© The Author(s) 2026. Published by Oxford University Press on behalf of the Chinese Antibody Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Antibody Therapeutics, 2026, 9(4), tbag053
https://doi.org/10.1093/abt/tbag053
Research Article