Supporting Information

DSSA-PPI: Enhancing binding affinity change prediction upon protein mutations using disentangled structure-sequence aware attention

Juhong Wu[1], Jiehui Sun[1], Tian Liang[1], Yongqi Zhang[1], Han Zhang[1], Tianyi Zhang[4], Xianmin Feng[4], Ping Gao[1], Peng Xu[2,*], Jinyu Li[1,3,5,*]

[1] College of Chemistry, Fuzhou University, Fuzhou 350116, Fujian, China

[2] College of Biological Science and Engineering, Fuzhou University, Fuzhou 350116, Fujian, China

[3] College of Biological and Pharmaceutical Engineering, Jilin Agricultural Science and Technology University, Jilin 132101, China

[4] Jilin Medical University, Jilin 132013, China

[5] Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, Xiamen 361005, Fujian, China

Corresponding authors:

[*]Peng Xu, email: pengxu@fzu.edu.cn

[*]Jinyu Li, email: j.li@fzu.edu.cn

Table of Contents

S1 Supplementary figures

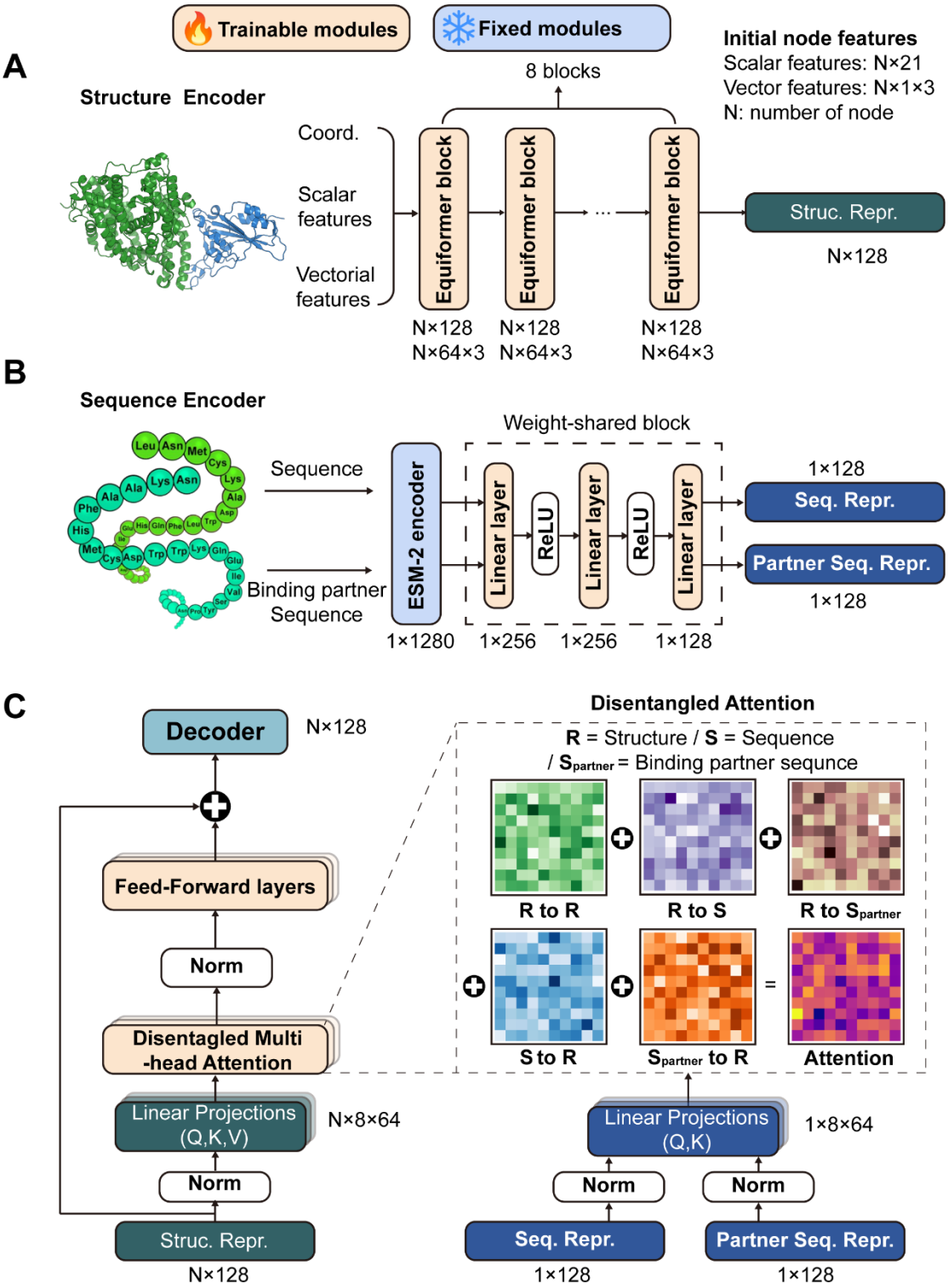

Figure S1. Schematic of DSSA architecture. (A) Structure encoder processes the 3D structure of an input protein-protein interface graph. Initial node features, comprising scalar (N×21) and vectorial (N×1×3) attributes for N residues, are fed into a series of eight trainable Equiformer blocks to produce structural representations. (B) Sequence encoder utilizes a fixed, pre-trained ESM-2 model, followed by a weight-shared block to produce

sequence representations. (C) Disentangled multi-head attention mechanism that explicitly computes attention scores for five attention terms.

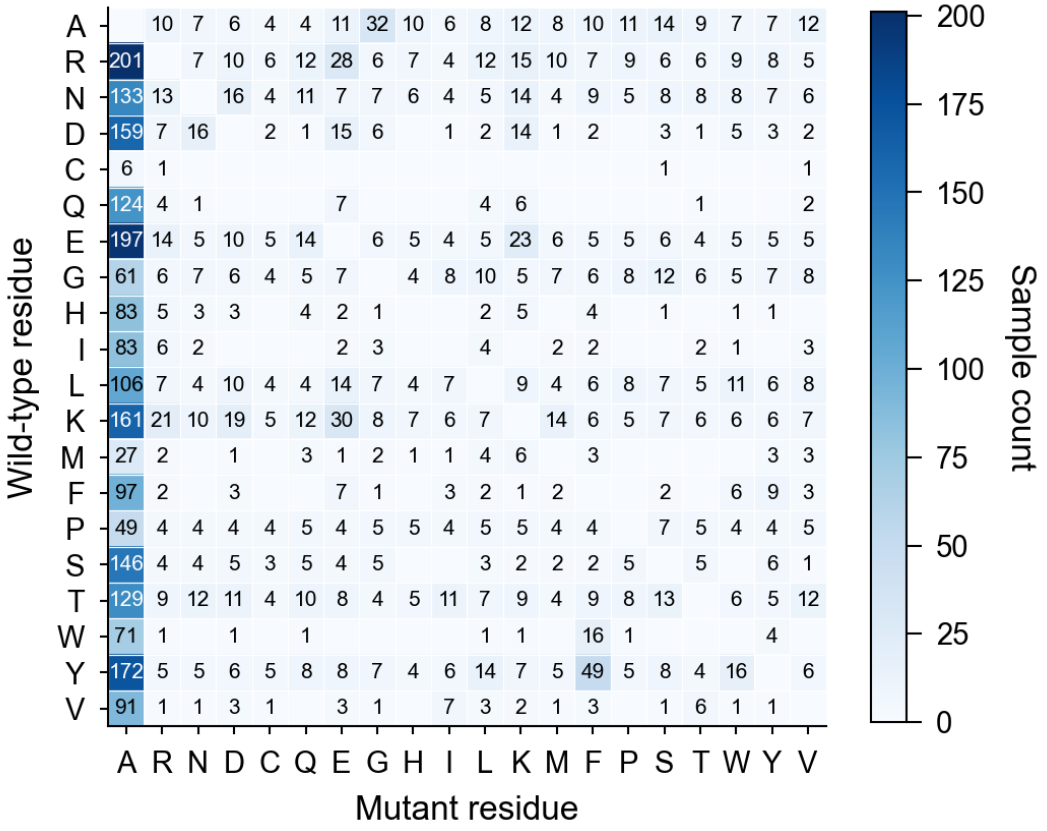

Figure S2. Distribution of mutation samples across amino acid substitutions in the S4169 dataset.

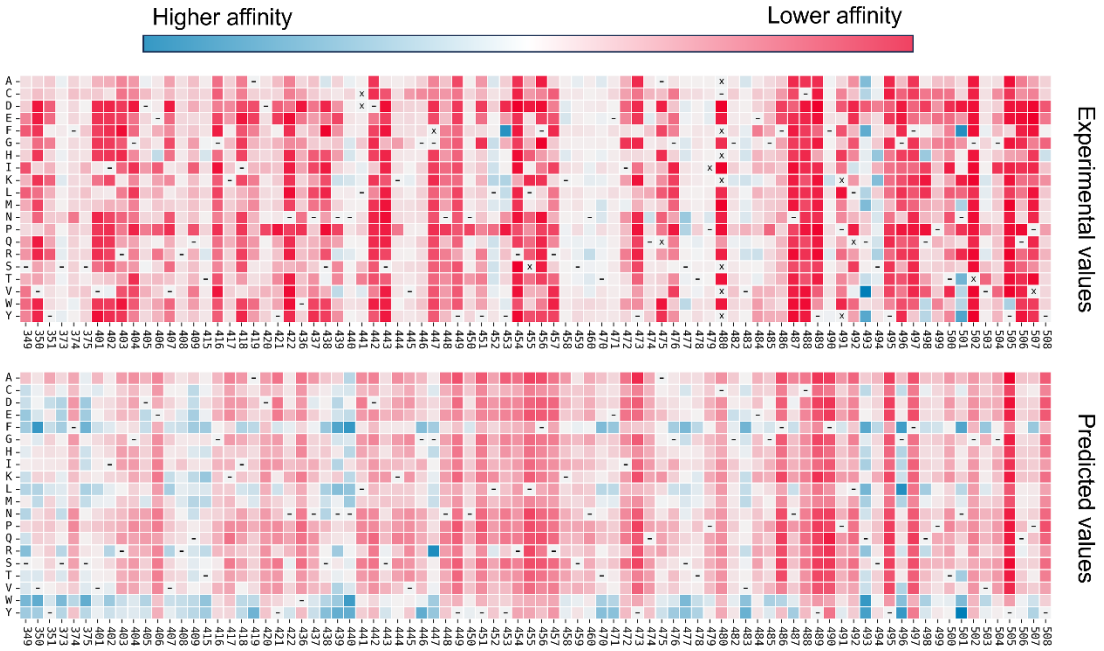

Figure S3. Heatmaps illustrate the experimental measured (top panel) and the predicted (bottom panel) ACE2-binding affinity landscapes of single mutations in the sequence of Wuhan SARS-CoV-2 RBM. The experimental values are measured by deep mutational scanning[1]. Missing data (NaN) are denoted by black '×' symbols.

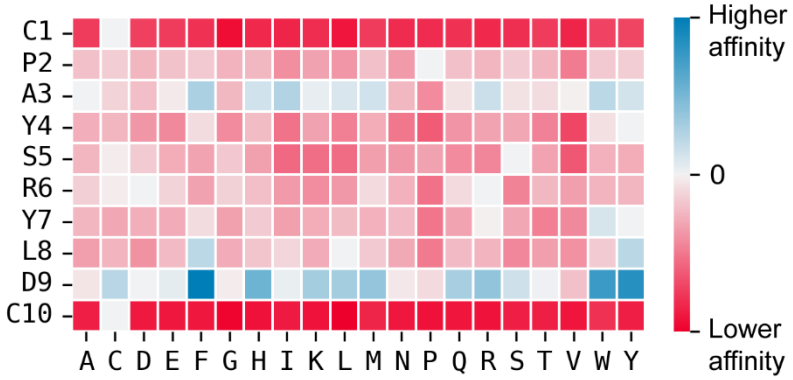

Figure S4. Heatmap of predicted single-point mutation effects on mupain-1 binding to FXIa. Blue indicates mutations predicted to enhance affinity, while red denotes mutations predicted to reduce affinity.

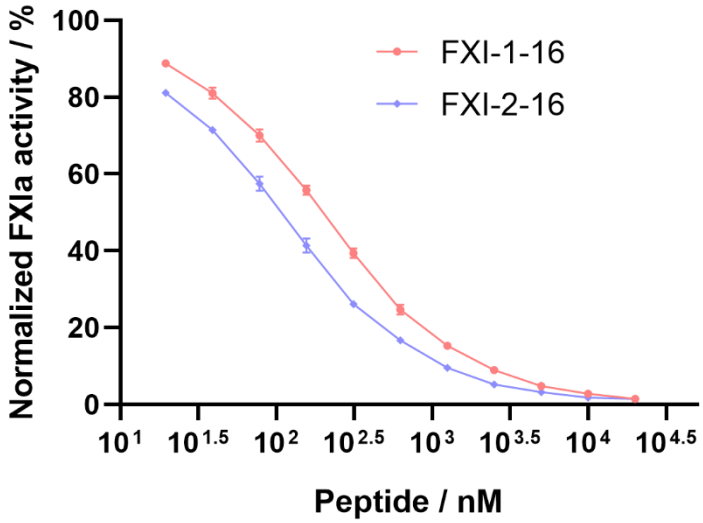

Figure S5. Curve of DSSA-PPI–designed peptides' inhibition against FXIa's proteolytic activity determined by chromogenic substrate S-2366 (n=3).

S2 Supplementary tables

Table S1. Performance comparison on the S8338 benchmark dataset under leave-one-complex-out cross-validation.

Methods	S8338	
	PCC ↑	RMSE ↓
DSSA-PPI (1 epoch)	0.65	1.68
DSSA-PPI (2 epochs)	0.74	1.42
DSSA-PPI (3 epochs)	0.77	1.35
TopNetTree[a]	0.80	1.15
MutaBind2[b]	0.74	1.37
mCSM-PPI2[b]	0.75	1.30

[a] Metrics for TopNetTree from the original publication[2]

[b] Metrics for MutaBind2 and mCSM-PPI2 from the MutaBind2[3]

Table S2. Performance of DSSA-PPI across 10 folds under mutation-level and complex-level evaluations.

Folds	Mutation-level				Complex-level			
	PCC↑	Spear.↑	RMSE↓	AUROC↑	PCC↑	Spear.↑	RMSE↓	AUROC↑
1	0.76	0.72	1.52	0.85	0.77	0.70	1.44	0.85
2	0.79	0.76	1.51	0.86	0.62	0.54	2.33	0.84
3	0.78	0.73	1.64	0.85	0.48	0.41	2.08	0.65
4	0.78	0.72	1.57	0.83	0.64	0.59	1.67	0.74
5	0.74	0.68	1.62	0.84	0.73	0.67	1.61	0.72
6	0.75	0.71	1.58	0.84	0.66	0.62	1.60	0.76
7	0.78	0.75	1.43	0.85	0.48	0.50	2.12	0.73
8	0.80	0.76	1.38	0.85	0.65	0.60	1.64	0.73
9	0.74	0.69	1.59	0.83	0.41	0.41	1.77	0.66
10	0.82	0.77	1.39	0.87	0.54	0.29	1.82	0.66

Tabel S3. Comparison of model performance on the benchmark datasets S1131, S4169 and M1707 under 10-fold mutation-level cross-validation.[a] Performance of DSSA-PPI is reported as mean ± SD (n=10).

Method	S1131		S4169		M1707	
	PCC ↑	RMSE ↓	PCC ↑	RMSE ↓	PCC ↑	RMSE ↓
DSSA-PPI	0.82 ± 0.06	1.59 ± 0.14	0.75 ± 0.04	1.24 ± 0.09	0.82 ± 0.04	1.46 ± 0.07
MpbPPI[4]	**0.87**	**1.25**	**0.80**	**1.07**	–	–
GeoPPI[5]	0.85	–	0.78	–	0.89	–

TopNetTree[6]	0.85	1.55	0.79	1.13	–	–
PerSpect-EL[7]	0.85	1.30	–	–	–	–
PTA-EL[8]	**0.87**	1.22	0.77	1.15	–	–
Profile-score + FoldX[9]	0.74	1.70	–	–	–	–
BindProfX[9]	0.74	1.70	–	–	–	–
mCSM-PPI2[10]	–	–	0.76	1.19	–	–
FoldX[8]	0.46	–	0.33	–	–	–

[a] Bold values indicate the best results. The dash sign indicates the results of the corresponding methods are not available. Baseline results were taken directly from the related publications.

Table S4. Comparison of model performance in RMSE metric on the benchmark datasets S1131, S4169 and M1707. Metrics are reported as mean ± standard deviation over five runs. [a]

Methods	S1131	S4169	M1707
DSSA-PPI	**1.72 ± 0.04**	**1.55 ± 0.02**	2.54 ± 0.05
PPIformer	1.84 ± 0.01	1.67 ± 0.06	2.87 ± 0.04
ESM-2-FFN	1.97 ± 0.05	1.80 ± 0.08	3.19 ± 0.11
MT-TopLap	1.89 ± 0.04	1.70 ± 0.02	–
MuToN	1.73 ± 0.04	1.56 ± 0.02	–
MpbPPI	2.45 ± 0.08	1.64 ± 0.01	**2.52 ± 0.04**
ProBASS	1.90 ± 0.02	1.58 ± 0.01	2.65 ± 0.04
SAAMBE-SEQ	2.31 ± 0.01	1.84 ± 0.05	–
FoldX	2.66	2.26	3.72

[a] Bold values indicate the best results. The dash sign indicates the results of the corresponding methods are not available. Baseline results were taken directly from the related publications.

Tabel S5. Comparison of DSSA-PPI with physics-based methods on the S1131, S4169, and M1707 benchmark datasets under five-fold complex-level cross-validation.

Method	S1131		S4169		M1707	
	PCC ↑	RMSE ↓	PCC ↑	RMSE ↓	PCC ↑	RMSE ↓
DSSA-PPI	0.71 ± 0.01	1.59 ± 0.14	0.58 ± 0.01	1.24 ± 0.09	0.42 ± 0.02	1.46 ± 0.07
FoldX[a]	0.42	2.66	0.33	2.26	-	-
Flex ddG[a]	0.22	3.51	0.14	4.49	-	-
MM/GBSA[a]	0.30	6.52	0.16	4.96	-	-

[a] Performance values for FoldX, Flex ddG, and MM/GBSA were obtained from MpbPPI under identical evaluation settings.

Table S6. Architectural comparison of DSSA-PPI and representative machine-learning based models.

Method	Structural Encoder	Sequence Encoder	Fusion Strategy	Predictor
DSSA-PPI	PPIFormer	ESM-2	Disentangled cross attention	FFN[a] (Softmax)
PPIFormer	Equiformer	—	None	FFN (Softmax)
MT-TopLap	Topology-based features	ESM-2 + auxiliary features	Concatenation	FFN
ESM-2-FFN	—	ESM-2	None	FFN
ProBASS	ESM-IF1	ESM-2	Concatenation	GBT[b]
SAAMBE-SEQ	Physicochemical descriptors	Evolutionary profiles (PSSM)	Concatenation	GBT

[a] FFN refers to a feed-forward network.

[b] GBT refers to a gradient boosted tree.

Table S7. Performance comparison of DSSA-PPI, DSSA-PPI-PF (with PPIFormer-initialized structure encoder), and baseline PPIFormer across three benchmark datasets. Metrics are reported as mean ± standard deviation over five runs.

Methods	S1131		S4169		M1707	
	PCC ↑	AUROC ↑	PCC ↑	AUROC ↑	PCC ↑	AUROC ↑
DSSA-PPI	0.71 ± 0.01	0.77 ± 0.02	0.58 ± 0.01	0.70 ± 0.02	0.42 ± 0.02	0.72 ± 0.02
DSSA-PPI-PF	0.70 ± 0.01	0.77 ± 0.02	0.56 ± 0.01	0.69 ± 0.01	0.39 ± 0.03	0.67 ± 0.02
PPIFormer	0.68 ± 0.01	0.72 ± 0.02	0.54 ± 0.01	0.66 ± 0.01	0.31 ± 0.03	0.61 ± 0.01

Table S8. Evaluation on receptor binding motif (RBM) of SARS-CoV-2 interacting with ACE2.

Method	PCC ↑	RMSE ↓	AUROC ↑
DSSA-PPI	0.31	1.42	0.63
PPIFormer	0.15	1.69	0.63
ESM-2-FFN	-0.01	2.23	0.56
MuToN	0.01	1.40	0.50
MpbPPI	0.08	3.28	0.54
ProBASS	0.00	1.39	0.46
SAAMBE-SEQ	-0.19	2.64	0.53
FoldX	-0.52	6.11	0.24

S3 Supplementary methods

Pre-training dataset. For the pre-training task, we utilized the PPIRef50K dataset curated by Bushuiev et al.[11]. This dataset comprises 45,553 structurally distinct 3D protein-protein interfaces extracted from the Protein Data Bank (PDB). PPIRef50K has demonstrated superior performance in structure-based pre-training tasks, making it an ideal choice for initializing our model.

SKEMPI v2.0 dataset. The SKEMPI v2.0 is a manually curated dataset that includes binding affinity changes upon mutations for structurally solved PPIs and currently contains 7,085 mutations across 345 PPIs[12]. It is a common-use dataset for binding affinity change prediction benchmarking, containing single and multiple mutations. Moreover, for a comprehensive evaluation, we split the entire dataset into five sub-datasets: support (SUP), core (COR), rim (RIM), surface (SUR), and interior (INT), based on the location of mutations. These annotations were directly adopted from the SKEMPI v2 dataset, which follows the classification scheme and threshold criteria originally defined by Levy et al.[13] Binding affinity changes were measured as changes in Gibbs free energy (ΔΔG), calculated from the equilibrium dissociation constant (K_D) using the formula:

$$\Delta G = RT\ln K_D \quad (1)$$

where R is the ideal gas constant, and T is the temperature in Kelvin. The change in binding affinity upon mutation was then computed as follows:

$$\Delta\Delta G(\mathrm{mut}-\mathrm{wt}) = \Delta G(\mathrm{mut}) - \Delta G(\mathrm{wt}) \quad (2)$$

These datasets were used to train and evaluate our proposed method.

Pre-training with structural masking. S Analogous to the masked language modeling (MLM) objective in BERT, we pre-trained our model using a structural masking strategy tailored for protein complexes. For each complex in the pre-training dataset, a fixed proportion of residues (50%) from one protein partner was randomly selected for masking. Following the standard MLM convention, 80% of the selected residues were maksed, 10% were substituted with a random amino acid, and the remaining 10% were left unchanged.

The model was then trained to predict the original identities of the masked residues based on the structural context. Masked residues were dynamically reselected at each training step, enabling the model to learn generalized representations across different regions of the protein structure.

To optimize learning, the model minimized the recovery error using a cross-entropy (CE) loss function. Given the uneven distribution of amino acid types at protein complex interfaces, we employed a class-weighted cross-entropy loss to mitigate the impact of residue imbalance. Specifically, the relative frequency of each amino acid type was calculated across the pre-training dataset, and a corresponding weighting factor was incorporated into the loss function for each type. The overall loss function for the masked structure modeling (MSM) task is formulated as:

$$L_{MSM} = \sum_{i \in Mask} -w_{a_i} \cdot \log p\left(\hat{a}_i = a_i \mid a \setminus Mask\right) \quad (3)$$

Here, $p(\hat{a}_i = a_i \mid a \setminus Mask)$ denotes the probability of predicting the actual amino acid type a_i of residue $a_i \in Mask$, w_{a_i} is a weighting factor for the native amino acid type a_i. We pre-trained our model on four NVIDIA 4090 graphics processing units (GPUs) in a distributed data-parallel (DDP) mode. To maximize GPU memory utilization, we set the batch size to 10. The initial learning rate was configured at 5×10^{-4}, and we employed the Noam learning rate scheduler, which initially increases the learning rate linearly over a set number of warmup steps before decaying proportionally to the inverse square root of the step number. Specifically, we allocated 50 epochs for the warmup phase to ensure gradual and stable convergence. For optimization, we used the AdamW optimizer with default parameters ($\beta_1 = 0.9,\ \beta_2 = 0.999$). The complete pretraining stage required approximately 40 hours using four NVIDIA 4090 GPUs.

Prediction for downstream PPI task. Following the pre-training phase, the model is fine-tuned for the downstream task of predicting the binding affinity change of protein complexes between wild-type (wt) and mutated (mut) states. In the fine-tuning phase, the model predicts the binding affinity change by approximating the Gibbs free energy equation (equation 1) using the log-odds of the amino acid type probabilities derived from the output of the model. The predicted binding affinity change is calculated as:

$$\Delta\Delta G(\text{mut-wt}) = f_{scale} \times (logp_{\text{wt}} - logp_{\text{mut}}) \quad (4)$$

Where p_{wt} and p_{mut} are the predicted probabilities for wild-type and mutant residues, respectively. A learnable scaling parameter f_{scale} was introduced to capture the proportionality constant related to RT. A key aspect of this method is the enforcement of a symmetry constraint in $\Delta\Delta G$ prediction, ensuring:

$$\Delta\Delta G(\text{mut-wt}) = -\Delta\Delta G(\text{wt-mut}) \quad (5)$$

This symmetry property is critical for accurate predictions but has been overlooked in some prior models. The mean squared error (MSE) severed as loss objective at the fine-tuning stage. DSSA-PPI is highly efficient during inference stage. For example, a full scan of all single-point mutations within the SARS-CoV-2 receptor-binding motif (~7,000 variants) can be completed in 10 seconds on a single GPU. Similarly, evaluating single or double substitutions for the 10-mer peptide mupain-1 requires only a few seconds, enabling rapid mutational profiling and high-throughput computational protein engineering.

Evaluation metrics. The evaluation metrics used in this study include Pearson's correlation coefficient (PCC), Spearman's rank correlation coefficient, root mean squared error (RMSE), and the area under the receiver operating characteristic curve (AUROC), calculated based on the sign of mutation-induced changes in binding affinity. In addition, we introduce the hit-rate metric (Top10Hit), which evaluates the model ability to correctly identify top-ranking mutations that contribute significantly to changes in binding affinity. This metric measures the hit rate of experimentally validated top 10 ranked mutants within a fraction of model predictions. It simulates the scenario where the model prioritizes the most promising mutations for experimental validation.

Complex structure preparation for FXIa/mupain-1 and PK/mupain-1. The FXIa/mupain-1 complex was modeled by structurally aligning the crystal structure of FXIa (PDB ID: 4CRG) to the uPA/mupain-1 complex (PDB ID: 4X1N), followed by transferring the peptide coordinates. Similarly, the PK/mupain-1 complex was generated by superimposing the crystal structure of PK (PDB ID: 5F8X) onto the same uPA/mupain-1 template. The resulting modeled complexes were used as structural inputs in DSSA-PPI.

Experimental materials. Recombinant catalytic domain of human FXIa was expressed in *Pichia pastoris* strain X-33 as described before[14]. Harvested recombinant protein was purified with cation exchange column (SP Sepharose Fast Flow, SPFF, GE Healthcare, Uppsala, Sweden), followed by size exclusion chromatography (Superdex 75 HR 10/30, Pharmacia Biotech). Peptides were synthesized by Wuxi PharmaTech (Shanghai, China), with non-natural amino acid, Fmoc-L-3-(*N*-amidino-4-piperidyl)alanine, prepared as described in previous work[15]. Chromogenic substrate S-2366 (pyro-Glu-Pro-Arg-p-nitroanilide) for FXIa was purchased from Chromogenix (MA, USA).

Determination of K_i Values. For determination of K_i values for the inhibition of the enzyme by peptides under steady state inhibition conditions, 2 nM purified enzyme was preincubated in a total volume of 200 µL of 0.01 M Hepes, 0.14 M NaCl, pH 7.4 with 0.1% BSA at 37 °C, with various concentrations of peptides for 15 min prior to the addition of chromogenic substrate in concentrations approximately equal to the K_M value for the enzyme[16]. The initial reaction velocities were monitored at an absorbance of 405 nm. The inhibition constants (K_i) were subsequently determined from the nonlinear regression analyses of plots of V_i/V_0 versus $[I]_0$, using equation (6), derived under assumption of competitive inhibition:

$$\frac{V_i}{V_0} = \frac{K_i(K_M+[S]_0)}{K_i[S]_0+K_M(K_i+[I]_0)} \quad (6)$$

V_i and V_0 are the reaction velocities in the presence and absence of inhibitor, respectively; $[S]_0$ and $[I]_0$ are the substrate and inhibitor concentrations, respectively; K_M is K_M for substrate hydrolysis by the protease. In equation (6), it is assumed that $[S] \approx [S]_0$ and $[I] \approx [I]_0$. These conditions were fulfilled, as less than 10% of the substrate was converted to product in the assays and as the assay typically contained a final concentration of the protease of 2 nM and inhibitor concentrations of >5 nM.

S4 Brief review of baseline models

We evaluate the performance of DSSA-PPI against seven baseline methods: PPIFormer, ESM-2, TopNetTree, MT-TopLap, MuToN, MpbPPI, ProBASS, SAAMBE-SEQ, and FoldX.

To ensure fair and consistent comparison, all baseline models were re-implemented using their publicly available standalone codebases. Each model was evaluated under the same five-fold complex-level cross-validation protocol as used for DSSA-PPI. Default settings and hyperparameters provided by the original authors were used. Below we briefly describe each method:

PPIFormer[11]: A geometric deep learning model designed to predict changes in protein–protein interaction (PPI) binding free energy (ΔΔG) due to amino acid mutations at the interface. PPIFormer is pretrained on a structurally diverse dataset of 3D protein–protein interfaces and estimates ΔΔG by computing the log-odds probability of wild-type (WT) and mutant (MUT) amino acids. Standard alone code at https://github.com/anton-bushuiev/PPIformer

ESM-2[17]: A large-scale protein language model based on the Transformer architecture, trained on millions of protein sequences to learn complex sequence patterns. ESM-2 employs masked language modeling (MLM) to predict missing amino acids. Standard alone code at https://github.com/facebookresearch/esm

MT-TopLap[18,19]: A hybrid deep learning framework that integrates persistent Laplacian topological descriptors, physicochemical features, and embeddings from pretrained protein language models. MT-TopLap is pre-trained in a multitask setting across eight deep mutational scanning datasets and fine-tuned for ΔΔG prediction. Standard alone code at https://github.com/ExpectozJJ/MT-TopLap

MuToN[20]: A geometric deep learning model that captures residue-level structural perturbations caused by mutations. MuToN constructs contact graphs for WT and MUT

protein–protein complexes, where nodes encode residue features derived from a protein language model and one-hot amino acid encoding. ΔΔG is estimated by computing the difference between WT and MUT representations. Standard alone code at https://github.com/zpliulab/MuToN

MpbPPI[4]: A multi-task pretraining-based equivariant model designed to predict the impact of amino acid mutations on PPIs. It is first pretrained on four distinct tasks to learn generalizable protein representations. The final prediction of ΔΔG is made using a gradient boosting tree (GBT)-based decoder applied to the difference between WT and MUT complex representations. Standard alone code at https://github.com/arantir123/MpbPPI

ProBASS[21]: A hybrid neural network that integrating both sequence- and structural information by leveraging pre-trained model ESM-2 and ESM-IF1. Then it concatenates these embeddings and feed them into GBT-based decoder for predicting. Standard alone code at https://github.com/sagagugit/ProBASS

SAMMBE-SEQ[22]: A sequence-based model that combines hand-crafted features—including position-specific scoring matrices (PSSMs), physicochemical properties, and evolutionary conservation metrics—with a GBT regressor to predict ΔΔG. Standard alone code at http://compbio.clemson.edu/saambe_webserver/indexSEQ.php#started

FoldX[23]: An empirical force field-based tool that estimates ΔΔG by modeling the energetic effects of mutations using a predefined scoring function. FoldX is widely used in protein stability and affinity predictions due to its efficiency.

S5 Dataset splitting in SKEMPI v2

We provide detailed information on the data splitting strategy used for the evaluation on the SKEMPI v2 dataset. Specifically, we employed a leave-one-complex-out (LOCO) cross-validation approach. To this end, LOCO evaluation was conducted on the S8338 dataset, which comprises both forward and reverse mutations derived from the S4169 benchmark set. In each fold, all mutations associated with a single protein–protein complex were withheld as the test set, while the model was trained on the remaining complexes. This process was repeated across all 319 protein–protein complexes, following established protocols described in Zhang *et al*.[3]

In addition, we applied a 10-fold cross-validation strategy using two distinct splitting schemes: mutation-level and complex-level splitting. In mutation-level splitting, individual mutation entries were randomly partitioned into ten folds, ensuring an even distribution of samples. In contrast, complex-level splitting grouped all mutations originating from the same wild-type protein–protein complex structure and assigned them entirely to either the training or test set. This prevents data leakage across structural contexts and better reflects the generalization capability of the model across unseen complexes.

References

(1) Starr, T. N.; Greaney, A. J.; Hilton, S. K.; Ellis, D.; Crawford, K. H. D.; Dingens, A. S.; Navarro, M. J.; Bowen, J. E.; Tortorici, M. A.; Walls, A. C.; King, N. P.; Veesler, D.; Bloom, J. D. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020**, *182*, 1295-1310.e20.

(2) Chen, J.; Wang, R.; Wang, M.; Wei, G.-W. Mutations Strengthened SARS-CoV-2 Infectivity. *J. Mol. Biol.* **2020**, *432*, 5212–5226.

(3) Zhang, N.; Chen, Y.; Lu, H.; Zhao, F.; Alvarez, R. V.; Goncearenco, A.; Panchenko, A. R.; Li, M. MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions. *iScience* **2020**, *23*, 100939.

(4) Yue, Y.; Li, S.; Wang, L.; Liu, H.; Tong, H. H. Y.; He, S. MpbPPI: A Multi-Task Pre-Training-Based Equivariant Approach for the Prediction of the Effect of Amino Acid Mutations on Protein–Protein Interactions. *Brief. Bioinform.* **2023**, *24*, bbad310.

(5) Liu, X.; Luo, Y.; Li, P.; Song, S.; Peng, J. Deep Geometric Representations for Modeling Effects of Mutations on Protein-Protein Binding Affinity. *PLOS Comput. Biol.* **2021**, *17*, e1009284.

(6) Wang, M.; Cang, Z.; Wei, G.-W. A Topology-Based Network Tree for the Prediction of Protein–Protein Binding Affinity Changes Following Mutation. *Nat. Mach. Intell.* **2020**, *2*, 116–123.

(7) Wee, J.; Xia, K. Persistent Spectral Based Ensemble Learning (PerSpect-EL) for Protein–Protein Binding Affinity Prediction. *Brief. Bioinform.* **2022**, *23*, bbac024.

(8) Liu, X.; Feng, H.; Lü, Z.; Xia, K. Persistent Tor-Algebra for Protein–Protein Interaction Analysis. *Brief. Bioinform.* **2023**, *24*, bbad046.

(9) Xiong, P.; Zhang, C.; Zheng, W.; Zhang, Y. BindProfX: Assessing Mutation-Induced Binding Affinity Change by Protein Interface Profiles with Pseudo-Counts. *J. Mol. Biol.* **2017**, *429*, 426–434.

(10) Rodrigues, C. H. M.; Myung, Y.; Pires, D. E. V.; Ascher, D. B. mCSM-PPI2: Predicting the Effects of Mutations on Protein–Protein Interactions. *Nucleic Acids Res.* **2019**, *47*, W338–W344.

(11) Bushuiev, A.; Bushuiev, R.; Kouba, P.; Filkin, A.; Gabrielova, M.; Gabriel, M.; Sedlar, J.; Pluskal, T.; Damborsky, J.; Mazurenko, S.; Sivic, J. Learning to Design Protein-Protein Interactions with Enhanced Generalization. *Twelfth Int. Conf. Learn. Represent.* **2024**.

(12) Jankauskaitė, J.; Jiménez-García, B.; Dapkūnas, J.; Fernández-Recio, J.; Moal, I. H. SKEMPI 2.0: An Updated Benchmark of Changes in Protein–Protein Binding Energy, Kinetics and Thermodynamics upon Mutation. *Bioinformatics* **2019**, *35*, 462–469.

(13) Levy, E. D. A Simple Definition of Structural Regions in Proteins and Its Use in Analyzing Interface Evolution. *J. Mol. Biol.* **2010**, *403*, 660–670.

(14) Jiang, L.-G.; Yuan, C.; Chen, H.-W.; Wang, Y.; Zhao, B.-Y.; Zhang, X.; Huang, M.-D. Preparation and Structure of a New Coagulation Factor XI Catalytic Domain for Drug Discovery. *Chin. J. Struct. Chem.* **2011**, *30*, 1021–1029.

(15) Hosseini, M.; Jiang, L.; Sørensen, H. P.; Jensen, J. K.; Christensen, A.; Fogh, S.; Yuan, C.; Andersen, L. M.; Huang, M.; Andreasen, P. A.; Jensen, K. J. Elucidation of

the Contribution of Active Site and Exosite Interactions to Affinity and Specificity of Peptidylic Serine Protease Inhibitors Using Non-Natural Arginine Analogs. *Mol. Pharmacol.* **2011**, *80*, 585–597.

(16) Xu, P.; Xu, M.; Jiang, L.; Yang, Q.; Luo, Z.; Dauter, Z.; Huang, M.; Andreasen, P. A. Design of Specific Serine Protease Inhibitors Based on a Versatile Peptide Scaffold: Conversion of a Urokinase Inhibitor to a Plasma Kallikrein Inhibitor. *J. Med. Chem.* **2015**, *58*, 8868–8876.

(17) Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. *Science* **2023**, *379*, 1123–1130.

(18) Wee, J.; Chen, J.; Xia, K.; Wei, G.-W. Integration of Persistent Laplacian and Pre-Trained Transformer for Protein Solubility Changes upon Mutation. *Comput. Biol. Med.* **2024**, *169*, 107918.

(19) 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*, 6676–6683.

(20) Li, P.; Liu, Z.-P. MuToN Quantifies Binding Affinity Changes upon Protein Mutations by Geometric Deep Learning. *Adv. Sci.* **2024**, *11*, 2402918.

(21) Gurusinghe, S. N. S.; Wu, Y.; DeGrado, W.; Shifman, J. M. ProBASS—a Language Model with Sequence and Structural Features for Predicting the Effect of Mutations on Binding Affinity. *Bioinformatics* **2025**, *41*, btaf270.

(22) Li, G.; Pahari, S.; Murthy, A. K.; Liang, S.; Fragoza, R.; Yu, H.; Alexov, E. SAAMBE-SEQ: A Sequence-Based Method for Predicting Mutation Effect on Protein–Protein Binding Affinity. *Bioinformatics* **2021**, *37*, 992–999.

(23) Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX Web Server: An Online Force Field. *Nucleic Acids Res.* **2005**, *33*, W382–W388.