nature machine intelligence

Article

https://doi.org/10.1038/s42256-023-00691-9

Self-play reinforcement learning guides protein engineering

In the format provided by the authors and unedited

Supplementary Files

Table of Contents

Supplementary Notes

Sequence design evaluated by TAPE oracle

Action space and episode starting sequence

The action space for designing PAB1 and GFP sequences covers the entire length of the protein. In particular, it is 20^{75} for PAB1 and 20^{237} for GFP. We randomly select five starting sequences from the 30% of mutants with the lowest fitness for each dataset. We repeat the process five times for each starting sequence. The number of simulation times is set to 200.

Benchmarking methods

For this task, we have employed benchmark methods including Evolutionary-BO[1], Cbas[2] and DyNA-PPO[3] which were implemented in the FLEXS[1] environment based on TF-agents library. We started the generation process with a single protein sequence and divided it into 10 rounds, with 400 queries to the surrogate model in each round, resulting in a total of 4,000 queries across all 10 rounds. In each round, we updated the surrogate model by adding 1/10 of the training data. AdaLead[1] is a greedy search algorithm that utilizes model-guided evolution. At each iteration of batch queries, AdaLead performs a hill-climbing search on the learned landscape model and then queries the sequences with high predicted fitness. We implemented AdaLead using the FLEXS library, with a mutation rate set to 1/L, where L is the sequence length, and a threshold of 0.05. We also modified SAC (Soft Actor-Critic)[4] to a discrete version using the TF-agent library. In addition to having identical number of rounds and the same schemes to update surrogate as in other methods, the episode environmental setting is also the same as DyNA-PPO. If a sequence is repeatedly generated or the current move hits the same position in the action space within one episode, the episode ends. Moreover, if the fitness predicted by the surrogate no longer increases, the episode is terminated.

Information entropy evaluation

To assess the similarity between the natural and generated sequences of PAB1 and GFP, we put them into six ESM protein language models ('esm1b', 'esm1v1', 'esm1v2', 'esm1v3', 'esm1v4', 'esm1v5') to obtain the corresponding logit matrix. We then used formula (8) (Methods) to calculate the information entropy of each position, and formula (9) (Methods) to obtain the average information entropy. Finally, we computed the Spearman correlation of average information entropies between natural and generated sequences.

Distance-preserving scaling plot

To visualize the relationship between the edit distance and the DMS distance of generated sequences, we employ Distance-preserving scaling (DMS) dimensionality reduction from the sklearn package to transform one-hot encodings of generated sequences into a 2D array. Then, we use Linear Regression from the sklearn package to map the edit distance between sequence pairs to their 2D array DMS distance. By plotting the resulting Distance-preserving scaling pattern for all sequences (with two mutations as scale), we can visually inspect the clustering and distribution of the sequences. Furthermore, to highlight the discrepancy between the starting sequence and the generated sequences, we mark the starting sequence with different colors.

Peptide binder design using AlphaFold2 as surrogate

Settings for MCTS

In this implementation of EvoPlay, 1,000 generated sequences were obtained from both play move π_t and simulation move a_t. The counts of peptides at each stage of the move/simulation are presented in Supplementary Table 14. The AF2 surrogate model was only queried by newly-generated sequences to obtain state reward r_t; otherwise, the reward was selected directly from the historical list of generated sequences. In the original AlphaZero[5] paper by Silver et al., they described adding Dirichlet noise Dir(α) to the prior probabilities of actions from the root node in Monte Carlo Tree Search, with α values of {0.03,0.15,0.3} for Go, shogi, and chess, respectively, where α is inversely proportional to the amount of legal action space. Here, we used α=0.3 for the binder design task, since the legal action space of this peptide design task (20 peptide length, peptide length less than 30) is similar to that of chess ($1\times10^{\wedge50}$).

Alphafold2 with recycle 0

To speed up the peptide design process, we set the recycle parameter to 0 for AlphaFold 2.0. However, this results in poor prediction performance because the Evoformer multi-layer structure is incomplete (prev_pos, prev_msa_first_ros, and prev_pair are null when recycle is set to 0). To maintain the integrity of the Evoformer structure under recycle 0, we retain these empty parameters. This issue has been addressed in the latest version of AlphaFold 2.2.2. Setting recycle to 0 significantly reduces the inferring time for generating 1,000 peptides for 1ssc to approximately 2 hours, whereas with recycle 8, it takes 12 hours (see Supplementary Table 7).

Peptide frequency logos

To compare the physicochemical properties of interacting residues between natural and designed peptides, we create amino acid frequency logos, and calculate the number of each residue types at each position on the receptor interface, normalize it to the total number of counts at that position. The Logomaker package[6] in Python is used to create logos with bit information defined as:

$$K(X) = -\sum_{i=1}^{20} P(x_i) log P(x_i) \qquad (1)$$

Where $P(x_i)$ is the frequency of observing one type of amino acid in position i, and $K(X)$ is the total bit information[7].

Protein-peptide MD preparation

Native protein-peptide complex structures are retrieved from the RCSB protein data bank (https://www.rcsb.org/, PDB ID: 1ssc). We use PyMOL software (version 2.1) to remove ions, water molecules, and solvent molecules from the structure.

Analysis for molecular dynamics results of peptide design

We analyze the top 1% designs extracted from all runs and visualize contact sequence logos. The frequency distribution of amino acids in the designed peptides converged to a pattern similar to that of the native structure (Fig. 3h). We conduct molecular dynamics simulations (200 ns) for the peptide-protein complexes designed by EvoPlay and EvoBind, as well as the native complex. Structure of interfacial residues and electrostatic interactions of EvoPlay design are shown in Figure 3g. High-frequency lysine (K) in the designed peptides locates closest to the negatively charged

Glu-111 in the receptor, which may facilitate the formation of salt bridges via electrostatic interactions. Similarly, the presence of negatively charged glutamic acids (E) at positions 101-104 may be due to the nearby interface region where a positively charged Lys-104 is located. All of the results above demonstrate the capability of EvoPlay to capture the characteristics of peptide-protein interaction interfaces, which allows for stable binding.

Plasmid construction for 1ssc

The cDNA of bovine RNase1 (UniProt ID: P61823, residues: 27-138) was amplified by polymerase chain reaction with the primers 5'CATGCCATGGTTAAAGAAACCGCGGCGGC3', 5' CCGCTCGAGGCCTTCGCACGCCACAATAA3'. The PCR product was digested by NcoI-XhoI and then inserted into the vector pET-22b with a 6x-His tag at the C-terminus (Novagen). The reading frames of the above plasmids were verified by DNA sequencing.

Expression and purification of 1ssc

The cloned plasmid was transformed into *Escherichia coli* BL21 (DE3) cells (Novagen). Cultures of 1L LB medium with 100μg/ml ampicillin were grown at 37 °C for 4h and then incubated at 16°C overnight after adding 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). After centrifugation, the cell pellets were resuspended in lysis buffer (20mM Tris, pH 8.5, 250mM NaCl, 5% glycerol, 0.2% Triton X-100 and 2mM β-mercaptoethanol (βME)), sonicated and then centrifuged at 18,000 g at 4 °C for 30min to remove the insoluble lysate. The supernatant was mixed with Ni-NTA agarose (Qiagen) at 4 °C for 1h, washed with buffer containing 20mM Tris, pH 8.5, 250mM NaCl, 20mM imidazole and 2mM βME and then eluted with buffer containing 20 mM Tris, pH 8.5, 250 mM NaCl, 200 mM imidazole and 2 mM βME. Next, the protein was loaded onto an S-100 gel-filtration column (GE Healthcare) using running buffer (15mM sodium phosphate, pH 6.0, 100mM NaCl and 2mM dithiothreitol). The purity of the collected fraction was checked by SDS-PAGE and then the protein was concentrated with a 10kDa molecular-weight cut-off Amicon centrifugal ultrafiltration unit (Millipore). (Millipore). Finally, the purified protein at 15mg/ml with 5% glycerol was flash frozen with liquid nitrogen and stored at −80 °C.

Peptide Synthesis

Lyophilized powders of affinity peptides by EvoPlay's design were synthesized and purified by GenScript Co., Ltd. The quality of peptide (purity: above 98%) was verified by high-performance liquid chromatography and mass spectroscopy. Peptides were resuspended in 100% DMSO and diluted into PBS to achieve a stock concentration of 1 mM. Sonication and gentle warming were used to improve solubility in the event the peptides precipitated during dilution.

Monte Carlo Simulated Annealing optimization

We implemented the Monte Carlo algorithm proposed by Anishchenko, I. et al.[8], which optimizes sequences in a 20 × L sequence space (L represents the length of the sequence and 20 represents the number of amino acids). At each iteration, an amino acid is randomly selected from the sequence and replaced at a randomly selected position. The objective function to be optimized is the same as in EvoBind:

$$loss = \frac{(d_{peptide}+d_{receptor})\cdot\Delta COM}{pLDDT_{peptide}\times 2} \qquad (2)$$

and the starting sequences are the same as those used by the other two algorithms. Each move (replacement) is accepted based on the Metropolis criterion:

$$A_i = \min\left[1, exp\left(-\frac{(loss_i - loss_{i-1})}{T}\right)\right] \qquad (3)$$

$Loss_i$ and $loss_{i-1}$ represent the losses of the step i and $i-1$, respectively. The temperature T is initially set to 0.1 and halved every 200 steps. The move is accepted when the value of A_i is less than a uniformly random number.

EvoPlay-assisted directed evolution for GB1 and PhoQ

First-stage unsupervised clustering with GP-assisted sampling

We employed CLADE[9], which combines unsupervised clustering with Gaussian Process (GP)-assisted sampling, as a comparative approach to sample 384 mutants at the first stage and predict 96 mutants using the ftMLDE ensemble models[10] at the second stage. The resulting 480 mutants were evaluated using one-hot encoding after conducting an ablation study. The UCB acquisition function with a trade-off parameter β of 4 was used for the GP regression model in the CLADE implementation. The score was calculated as the sum of the mean and standard deviation. For further information, please refer to the CLADE[9] paper.

Second-stage supervised learning with ensemble models

For the second stage of our approach, we use the ftMLDE package[10] and the same 17 regression models employed in CLADE. However, we do not use hyperopt[11] for hyperparameter optimizations, as it would take another 30 minutes to complete the training on a single RTX6000 GPU for a single independent round (500 repeats would take an additional 10 days). The rest of the settings are the same as in CLADE, including 5-fold cross-validation on 384 sequences and selecting the top 96 predictions from three models.

GLuc designed by EvoPlay

Analysis for molecular dynamics results of GLuc-MT2 mutant

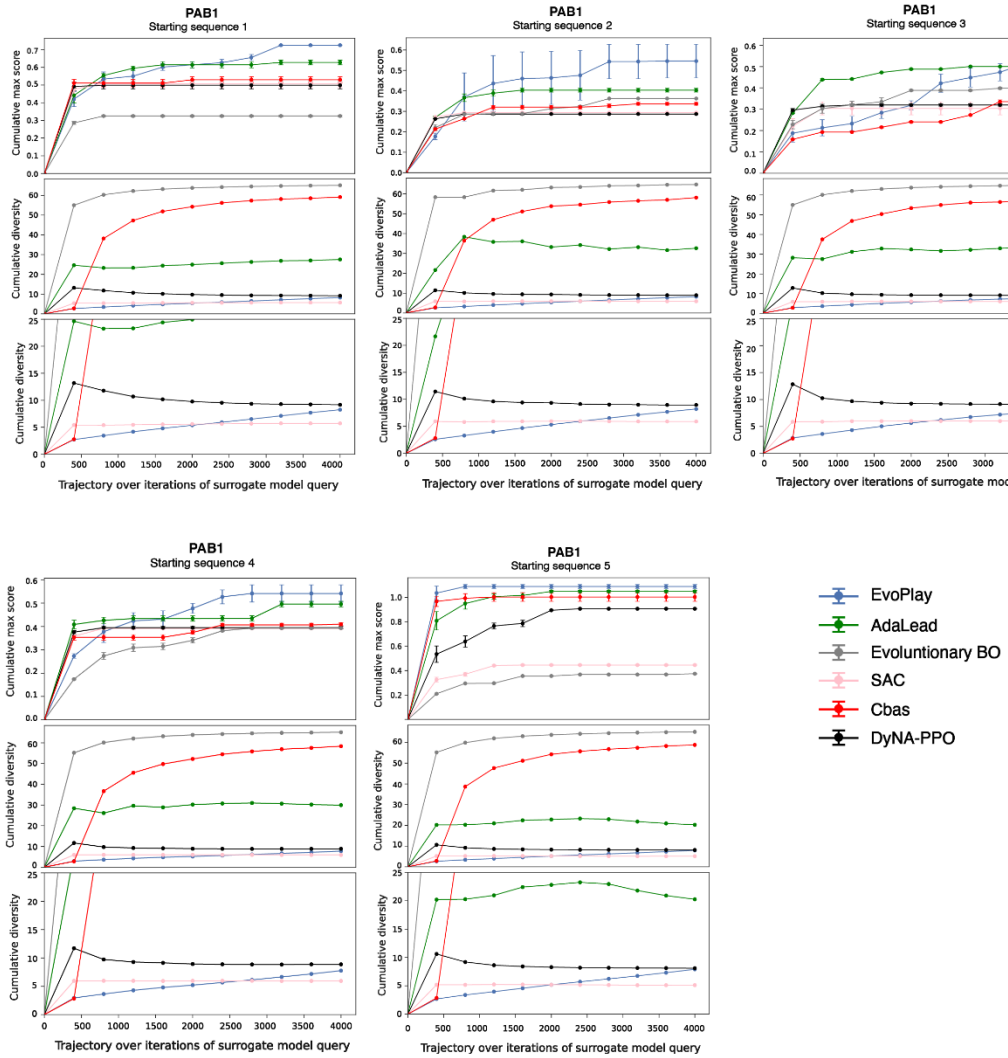
MD simulations of GLuc-MT2 reveal some critical factors driving catalytic activity. GLuc-MT2 form hydrogen bonds, π-π conjugates, hydrophobic interactions with CTZ compound to improve binding affinity and stability. His-62, Ile-63, His-78, Thr-79 are identified as key binding residues of GLuc-MT2 (Fig. 5c). N10S and V12A increase the adaptability of the enzyme cavity (Supplementary Fig. 6), leading to altered energy contribution of key binding residues (Ile-63, His-78, and Ala-87, Fig. 5c(ii)) and plausibly stronger binding affinity (-115.95±15.03 kJ/mol, 6.13-fold fitness of WT) (details in Supplementary Table 10).

Tools and settings for Molecular Dynamic Simulation

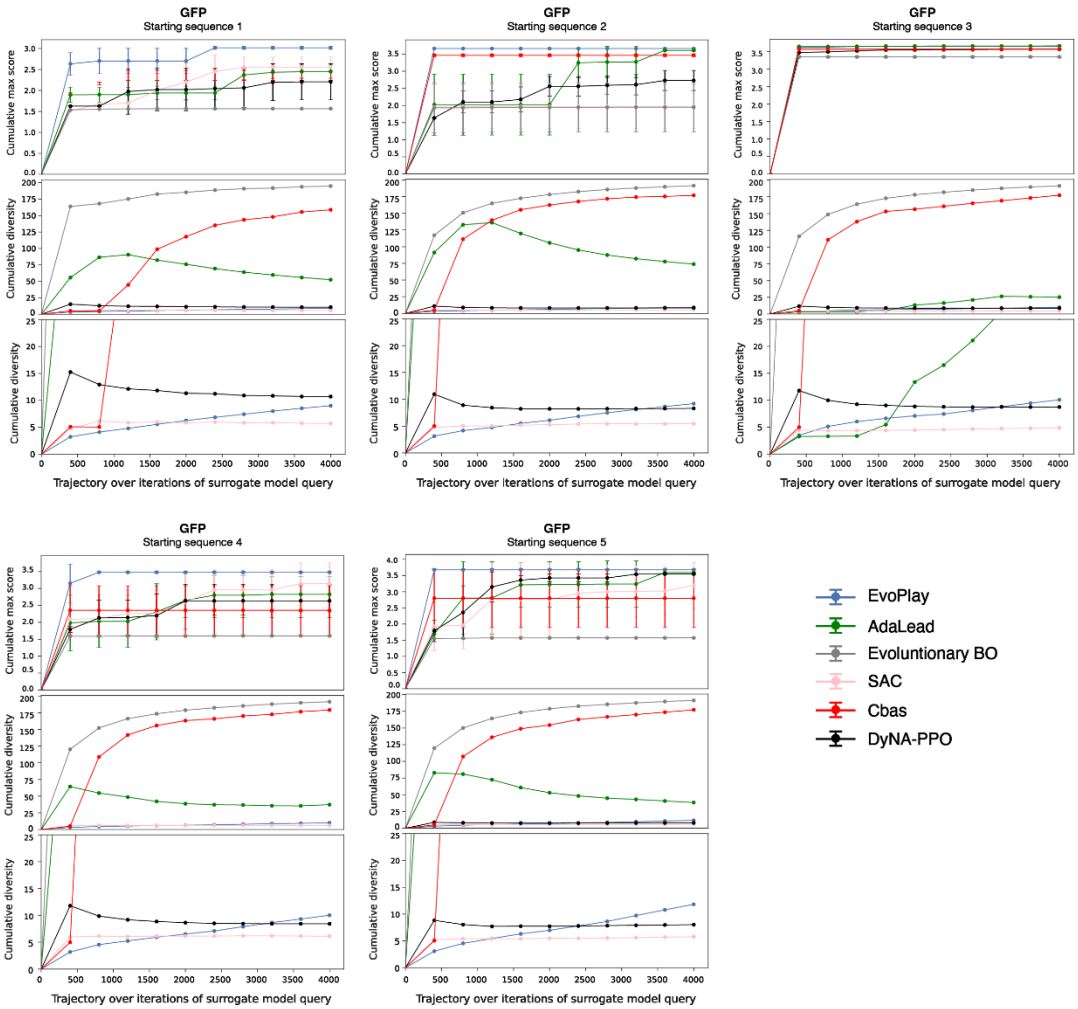
All MD simulations for the target systems are performed using the Gromacs 2020.1 package[12], with the AMBER99SB-ILDN force field[13]. The CTZ ligand topology was derived from Sobtop version 1.0 (dev3.1)[14] and all-atom charges of CTZ were calculated using the Restrained Electrostatic

Potential (RESP) protocol with ORCA 5.0[15], based on the B3LYP-D3/def2-TZVP basic set. The initial protein structures were solvated in a dodecahedron box of TIP3P water molecules[16], with a minimal distance of 12 Å from any edge of the box, and the periodic boundary conditions were applied for PAB1, GFP, protein-peptide and GLuc. The simulation systems were neutralized by adding 0.15M NaCl to simulate physiological conditions. Long-range electrostatics interactions were modelled by the Particle Mesh Ewald (PME) algorithm[17] and the direct space cut-off distance was set to 12 Å. The van der Waals interactions were calculated by a force-switched cut-off between 10 and 12 Å. The linear constraint solver (LINCS) algorithm was applied to constrain h-bonds with an integration time step of 2.0 fs. The system was minimized for 50,000 steps using the steepest descent algorithm. Then, equilibration under the NVT ensemble at 300 K was conducted for 500 ps. NPT equilibration was applied at 1 bar pressure and 300 K temperature for 500 ps, ensuring a balance of solvent molecules around protein residues. The temperature was coupled using the velocity rescaling algorithm[18] with a time constant of 0.1 ps, and pressure was controlled using the Parrinello-Rahman barostat algorithm[19] with a time constant of 1 ps. Finally, the well-equilibrated system was subjected to MD simulations for 200 ns at 300 K without any constraint. All single and multiple-site mutants, as well as the wild type, were analyzed using inbuilt tools (GROMACS 2020.1[12]).

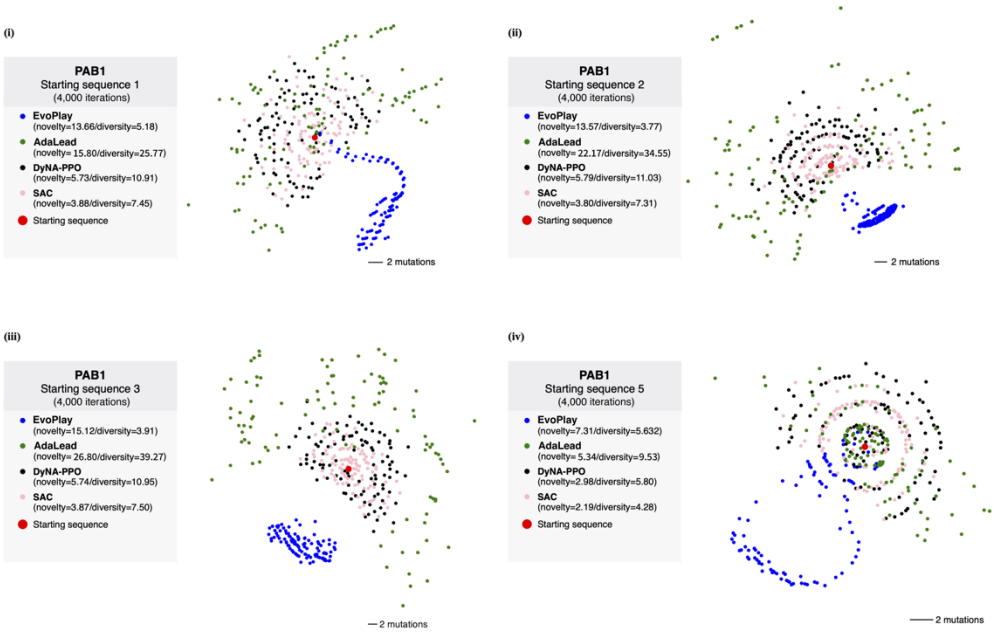
Supplementary Figures



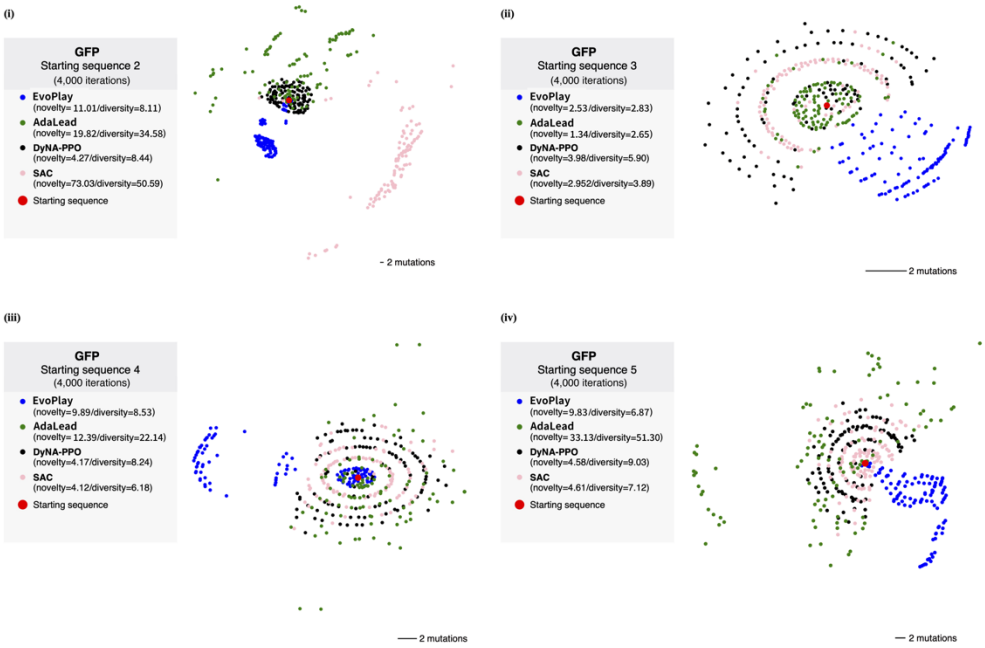
Supplementary Figure 1 Results of cumulative maximum scores and cumulative diversity for 5 starting sequences designed by EvoPlay and five other methods (AdaLead, evolutionary BO, Cbas, SAC and DyNA-PPO) on the PAB1 dataset, 5 repeats for each. The first two subplots depict the cumulative maximum scores and diversity, respectively, sharing the same X-axis. The third subplot provides a magnified view of the second subplot, with the Y-axis ranging from 0 to 25. For the cumulative max score, the error bars indicate the standard deviations (n=5 repeats), and the central line is the mean value. Our results show that EvoPlay can discover high-fitness sequences scored by the TAPE oracle.



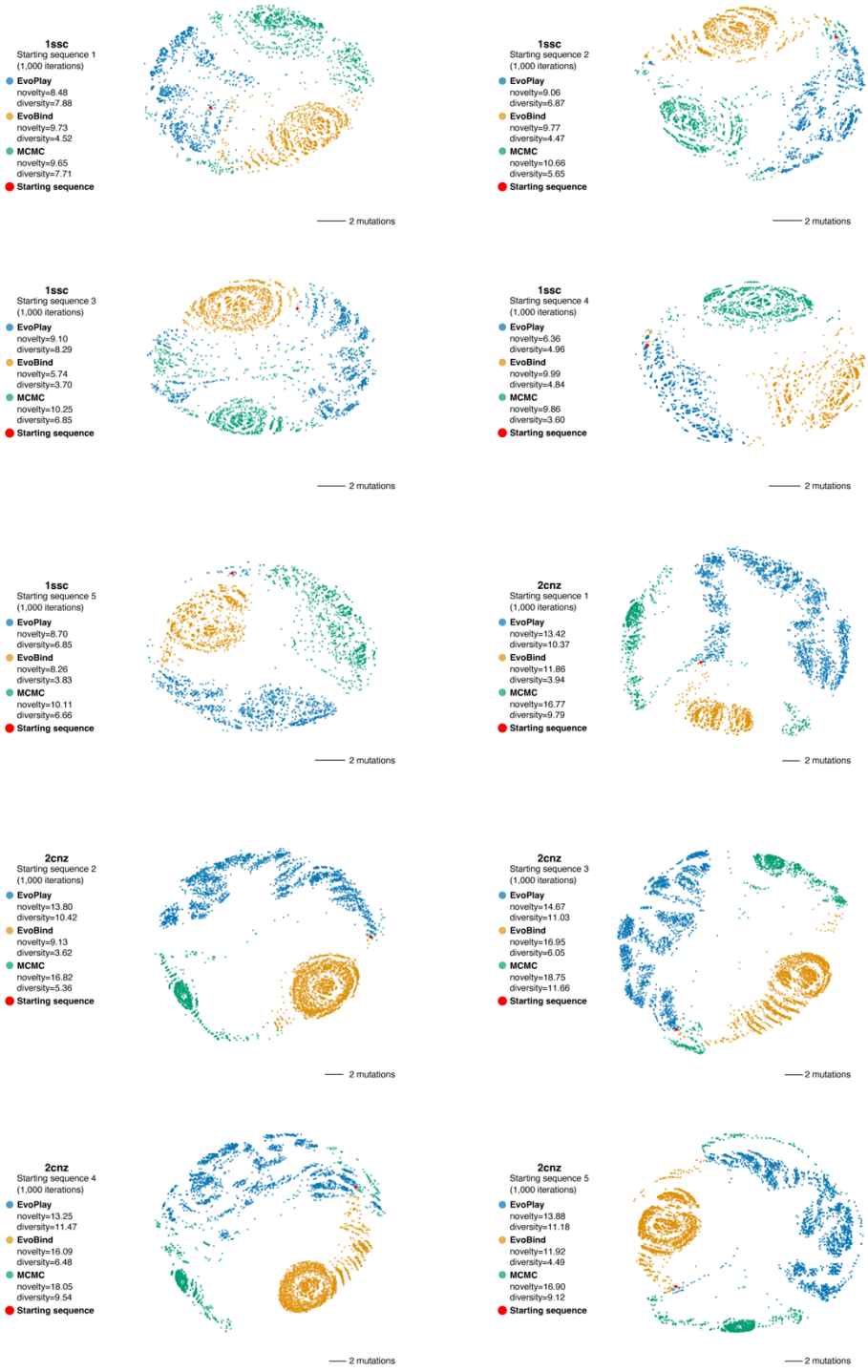
Supplementary Figure 2 Results of cumulative maximum scores and cumulative diversity for 5 starting sequences designed by EvoPlay and five other methods (AdaLead, evolutionary BO, Cbas, SAC and DyNA-PPO) on the GFP dataset, 5 repeats for each. The first two subplots depict the cumulative maximum scores and diversity, respectively, sharing the same X-axis. The third subplot provides a magnified view of the second subplot, with the Y-axis ranging from 0 to 25. For the cumulative max score, the error bars indicate the standard deviations (n=5 repeats), and the central line is the mean value. Our results show that EvoPlay can discover high-fitness sequences scored by the TAPE oracle.



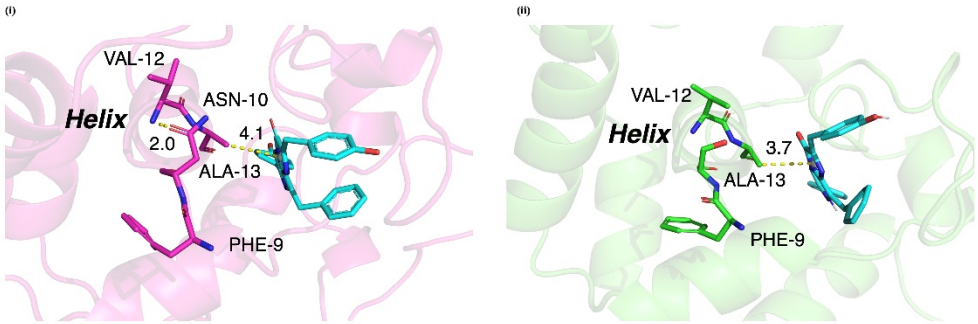
Supplementary Figure 3 Distance-preserving multidimensional scaling plots of another 4 starting sequences (except starting sequence 4 in Fig.2 b(ii)) designed by EvoPlay on PAB1 dataset. Another three benchmark methods (AdaLead, SAC and DyNA-PPO) are compared. The top 100 fitness sequences scored by TAPE oracle of one repeat are plotted, with the value of novelty and diversity indicated. Sequences generated by EvoPlay exhibit higher novelty.



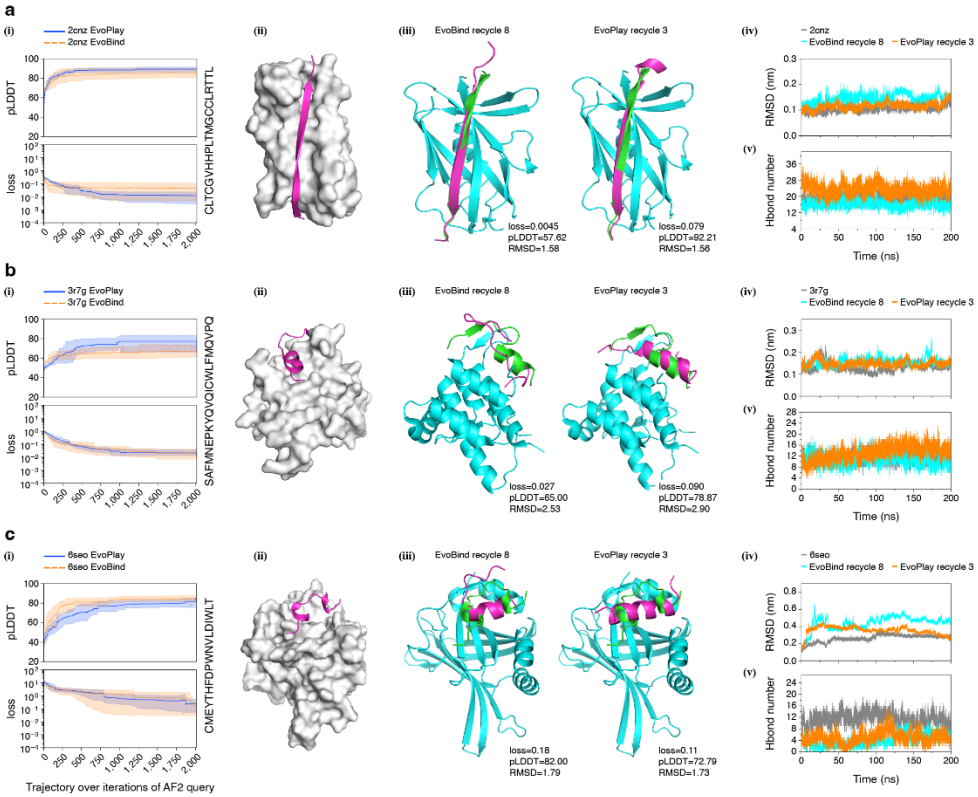
Supplementary Figure 4 Distance-preserving multidimensional scaling plots of another 4 starting sequences (except starting sequence 1 in Fig. 2 b(iv)) designed by EvoPlay on GFP dataset. Another three benchmark methods (AdaLead, SAC and DyNA-PPO) are compared. The top 100 fitness sequences scored by TAPE oracle of one repeat are plotted, with the value of novelty and diversity indicated.



Supplementary Figure 5 Distribution scatter plots of peptide sequences generated by EvoPlay, MCMC, and EvoBind (with 1ssc and 2cnz as the receptor) using 5 randomly selected starting seeds. EvoPlay generates sequences with higher diversity compared to MCMC and EvoBind.

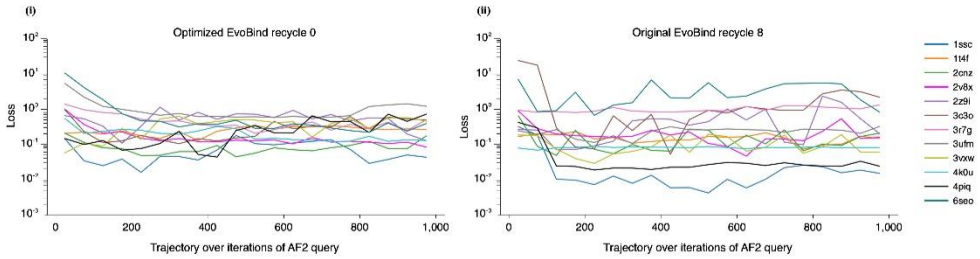


Supplementary Figure 6 A comparison of the binding modes between GLuc-WT and GLuc-MT2 with CTZ. (i), (ii) The binding mode of GLuc-WT-CTZ and GLuc-MT2-CTZ, respectively, was extracted from the last frame of the MD simulation. Key residues are shown in stick mode, colored magenta and green, while CTZ is rendered in cyan. The distance between Ala-13 and CTZ is indicated by a yellow dash. In GLuc-WT, the R group of Val-12 forms a strong hydrophobic interaction with surrounding residues (Leu-144, Leu-35, Met-43, Pro-31). However, this interaction is weakened when Val-12 is mutated to Ala-12, promoting flexibility of the helix structure and enhancing pocket adaptability. The Asn-10 residue in GLuc-WT forms a hydrogen bond interaction with the peptide-bond proton of Val-12, stabilizing the helix. When Asn-10 is mutated to Ser-10, this hydrogen bond is lost, resulting in a more flexible structure. As a result, Ala-13 can more easily participate in the hydrophobic interaction with CTZ.

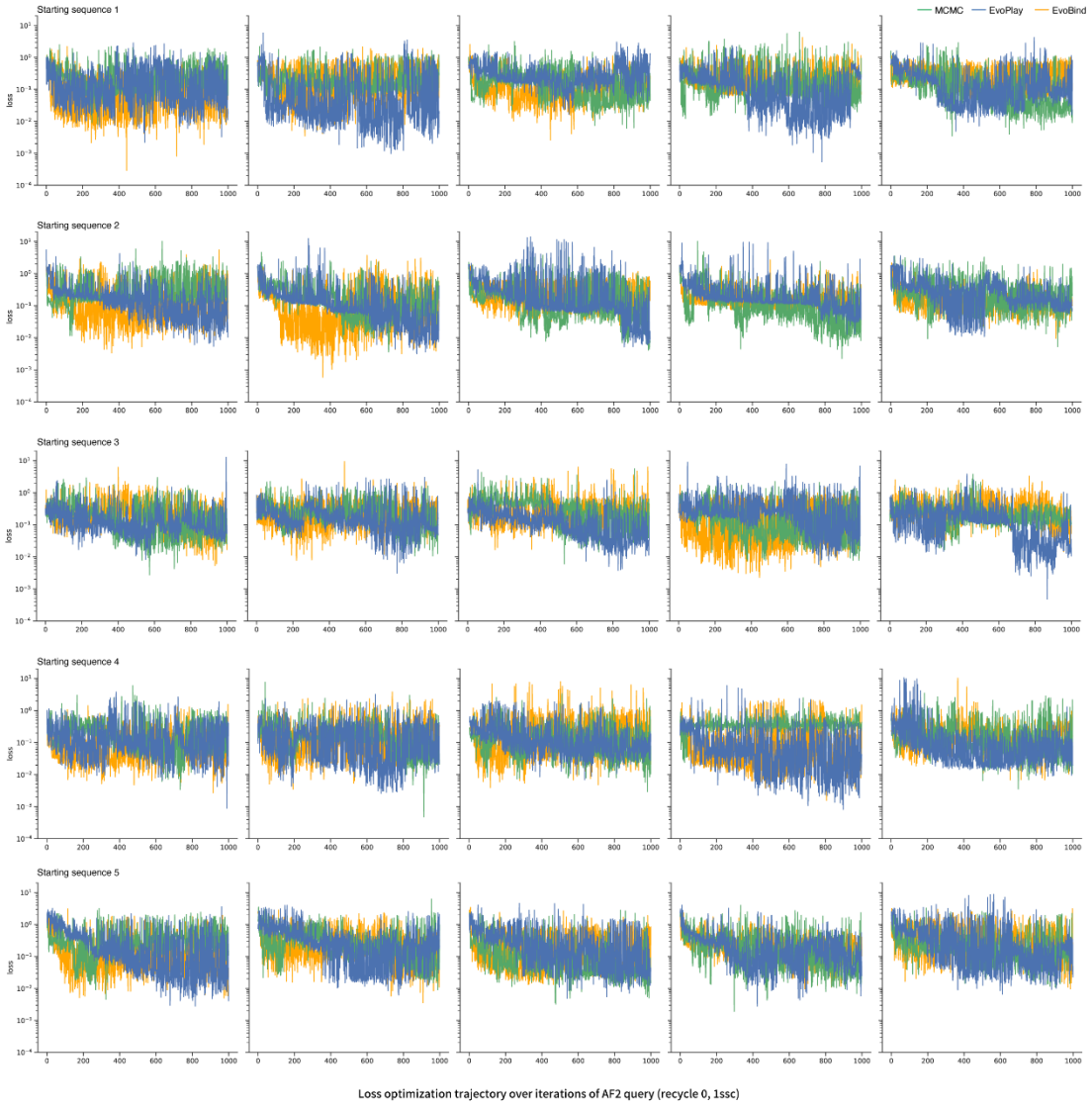


Supplementary Figure 7 EvoPlay produces high-quality peptides of higher pLDDT with less volatile loss. a(i), **b(i), c(i),** Loss and pLDDT trajectories of designed peptides by EvoPlay and EvoBind targeting 2cnz, 3r7g and 6seo (AF2, recycle3), the error bars show the 95% confidence

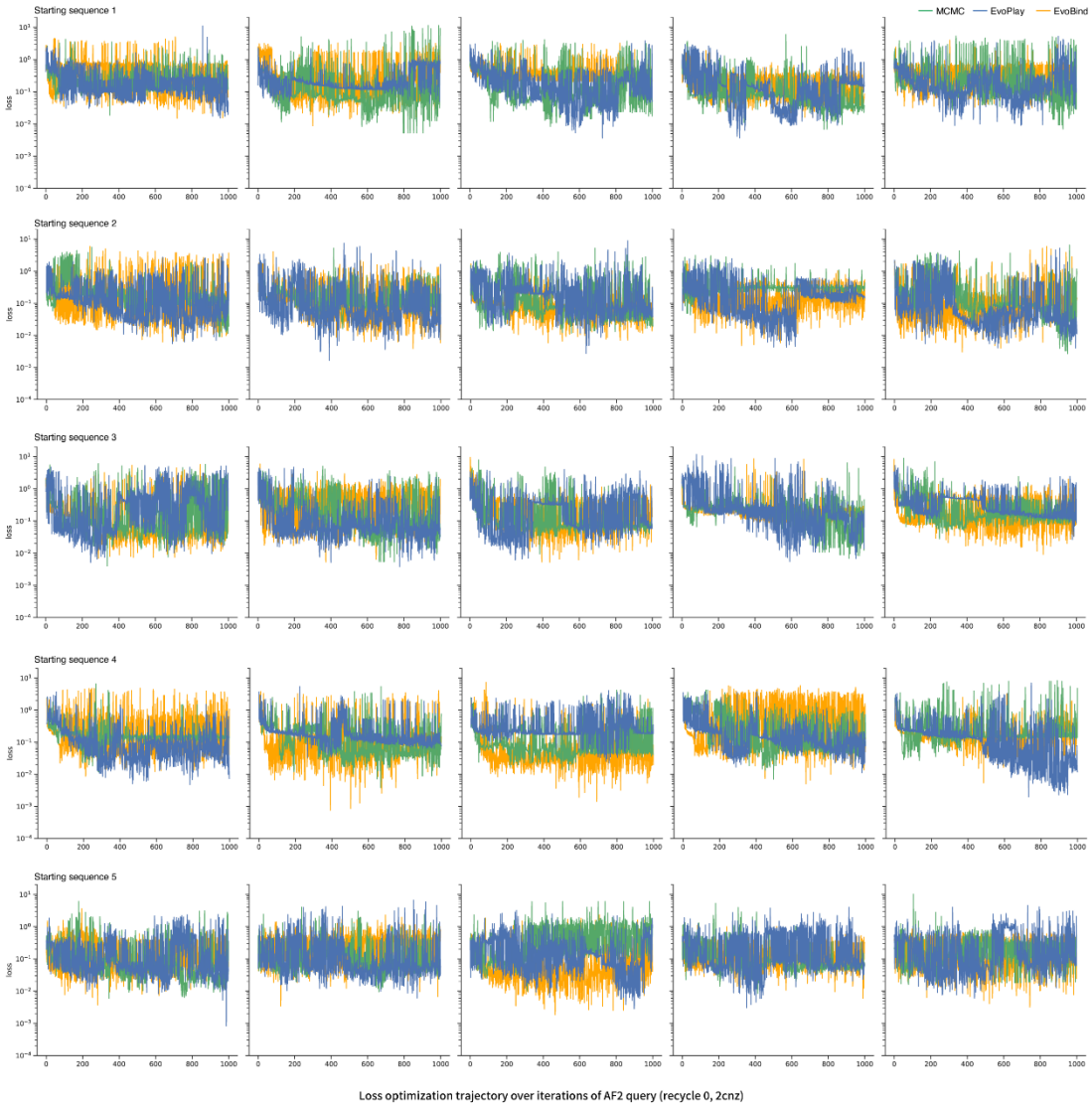
interval. **a(ii)**, **b(ii), c(ii),** The surface plot of 2cnz, 3r7g and 6seo. **a(iii)**, **b(iii), c(iii),** The superposition plot of the designed and native protein-peptide complex (EvoBind, recycle=8; EvoPlay-recycle=0) for 2cnz, 3r7g and 6seo. The color of the native protein and peptide is cyan and green, and the designed peptide is magenta. **a(iv)**, **b(iv), c(iv),** the RMSD changes of peptide-protein complex during 200ns molecular dynamics. **a(v)**, **b(v), c(v),** Hydrogen bond number changes between peptide and protein during 200ns molecular dynamics. For 2cnz, the average complex RMSD of EvoPlay-design, EvoBind-design and native are 0.12±0.013 nm, 0.15±0.018 nm, 0.11±0.013 nm as consistent conformation evaluation. The number of hydrogen bonds formed between protein and peptide of EvoPlay-design (23.49±2.82) is more than that of native (19.01± 1.77) and that of EvoBind (16.32±1.78), showing consistent trends with lower binding free energy of EvoPlay-designs (EvoPlay (-157.94±3.82 kcal/mol), native (-134.73±20.28 kcal/mol) and EvoBind (-118.26±8.47 kcal/mol), Supplementary Table 9). For 3r7g, the average complex RMSD of EvoPlay-design, EvoBind-design and native are 0.13±0.019 nm, 0.14±0.016nm, 0.15±0.017 nm as consistent conformation evaluation. The number of hydrogen bonds formed between protein and peptide of EvoPlay-design (13.32±2.71) is more than that of native (11.26±1.68) and that of EvoBind (10.64±2.10), showing consistent trends with lower binding free energy of EvoPlay-designs (EvoPlay (-96.48±35.81 kcal/mol), native (-81.74±24.87 kcal/mol) and EvoBind (-50.81±14.44 kcal/mol). EvoPlay-designs are fundamentally rational and reasonable. For 6seo, the average complex RMSD of EvoPlay-design, EvoBind-design and native are 0.25±0.043 nm, 0.35±0.079nm, 0.44±0.048 nm. The number of hydrogen bonds formed between protein and peptide of EvoPlay (5.01±2.07) is fewer than that of native (11.46±2.18) but more than that of EvoBind (3.82±2.03), showing consistent trends with lower binding free energy of EvoPlay designs (-44.37±7.08 kcal/mol) than that of EvoBind (-33.93±4.92 kcal/mol) but slightly higher than that of native (-57.15±6.44 kcal/mol). Overall, EvoPlay and EvoBind both perform poorer than native conformation when targeting 6seo. We argue the reason might come from localization of binding sites, as shown in **a(ii)**, **b(ii), c(ii),** the protein binding site of 6seo is flatter and the peptide binds only at the receptor surface. However, 1ssc, 2cnz and 3r7g proteins show a more well-configured binding pocket for peptide to match. The corresponding loss (centroid distance of the designed peptides) decreases to 0.1 for 6seo and it is difficult of further descending. AF2 training is performed on single-chain proteins. For some proteins with distinct active pockets, AF2 might produce unsatisfied prediction. In sum, EvoPlay is well-suited for peptide-binder design of high pLDDT and lower loss. EvoPlay can still design reasonably high pLDDT peptides (pLDDT > 80), even though there are some deviations between native peptides and AF2 predictions.



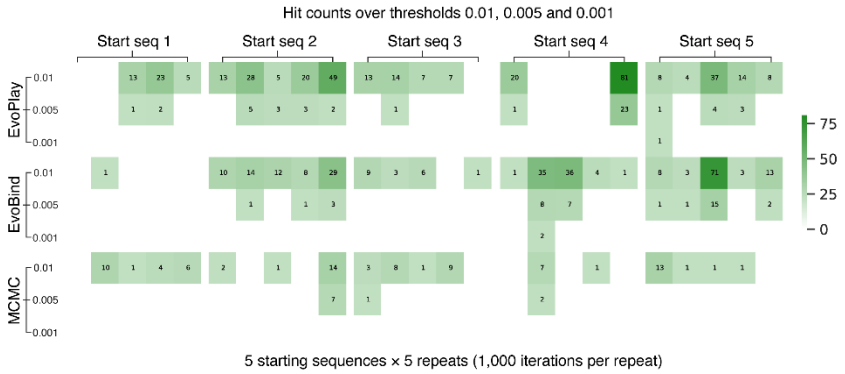
Supplementary Figure 8 Under the same training conditions, we have optimized EvoBind-recycle0 to approximate the performance of the original EvoBind-recycle8. We evaluated this by comparing the training loss and trajectory pattern for 12 peptides. The results showed that EvoBind-recycle0 can achieve similar performance to EvoBind-recycle8, indicating that our optimization has successfully improved the efficiency of the EvoBind algorithm.



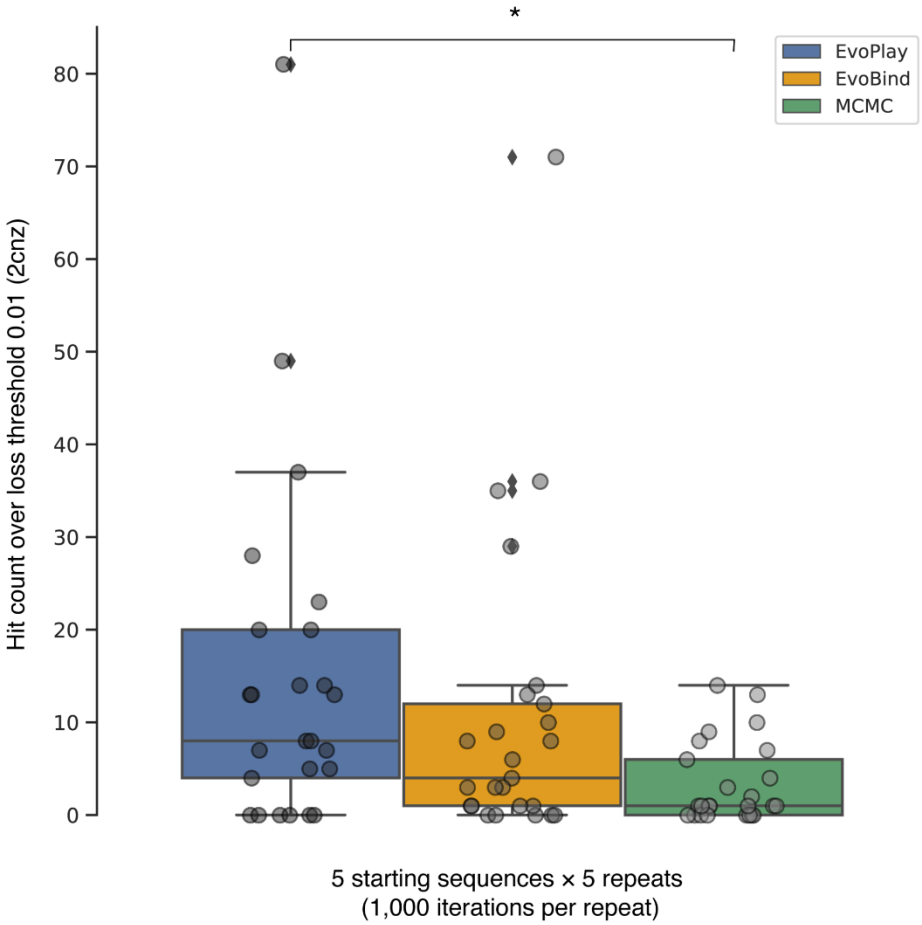
Supplementary Figure 9 EvoPlay, EvoBind, and MCMC methods were each tested in 5 repeated experiments, with 5 starting sequences of the 1ssc receptor. The y-axis represents the loss, and the x-axis represents the trajectory over iterations of the AlphaFold2 query (recycle 0).



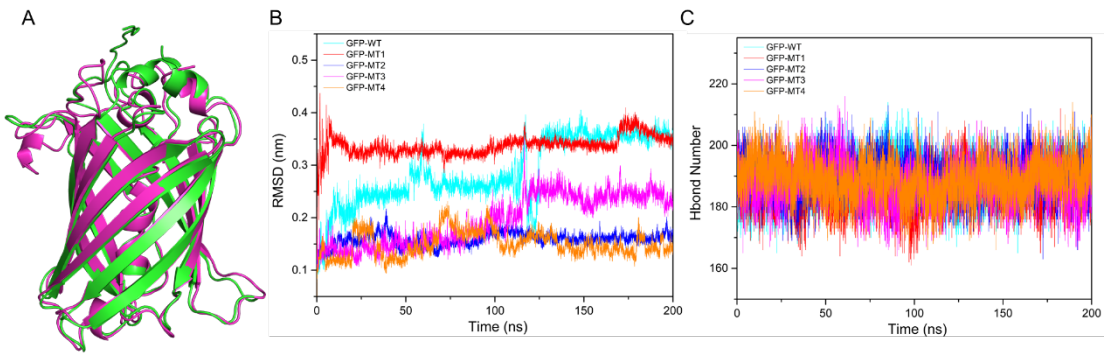
Supplementary Figure 10 EvoPlay, EvoBind, and MCMC methods were each tested in 5 repeated experiments, with 5 starting sequences of the 2cnz receptor. The y-axis represents the loss and the x-axis represents the trajectory over iterations of the AlphaFold2 query (recycle 3).



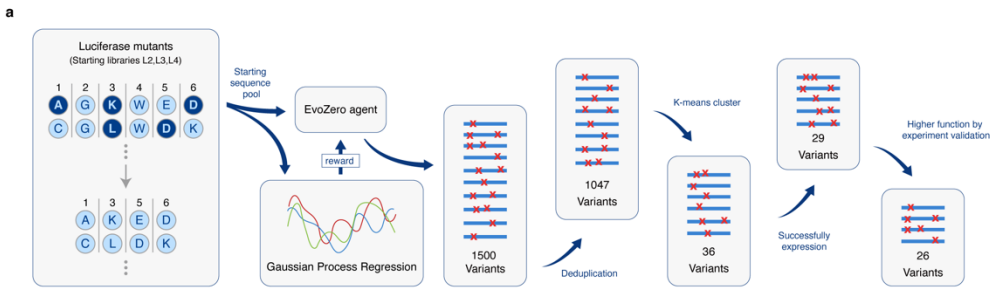
Supplementary Figure 11 The heatmap of EvoPlay, EvoBind and MCMC hit rates on 2cnz. The heatmap displays hit counts for reaching three thresholds (0.01, 0.005 and 0.001) by EvoPlay, EvoBind, and MCMC.



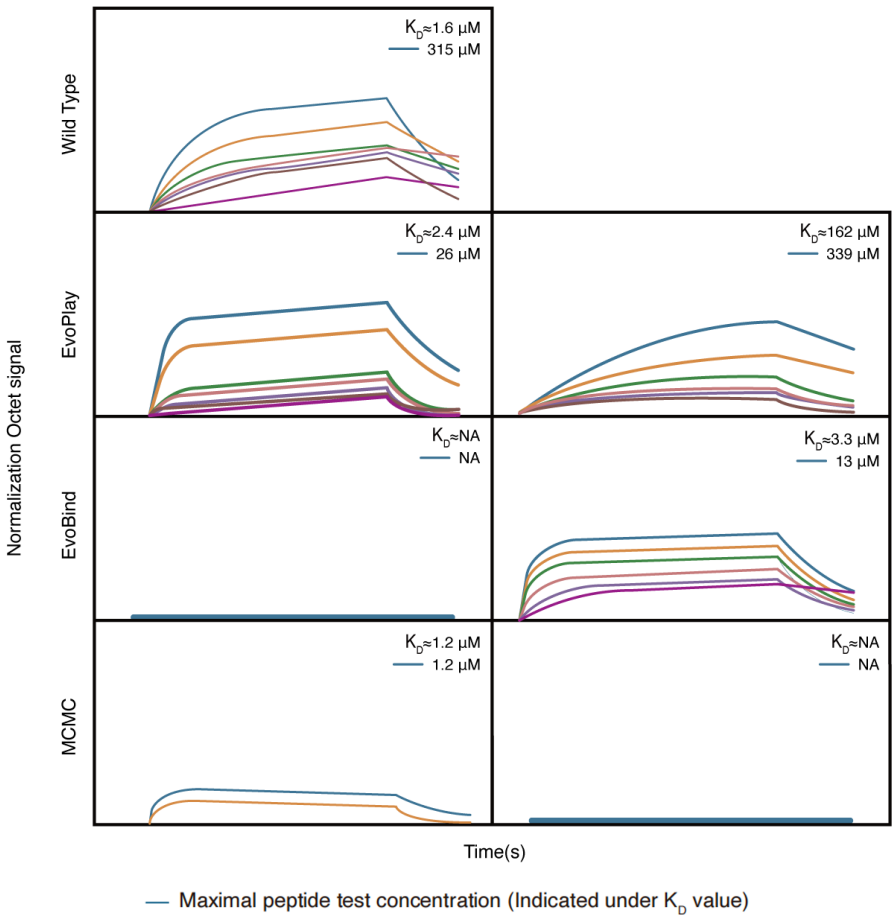
Supplementary Figure 12 The box plot of EvoPlay, EvoBind and MCMC hit rate on 2cnz with a loss threshold of 0.01. The box plot displays the significant difference in hit rates among benchmarking algorithms (n=25 independent repeated experiments for each algorithm). The box plots show the median, the first quartile and the third quartile values. The whiskers extend to points that lie within 1.5-times the interquartile ranges of the lower and upper quartiles and then observations that fall outside of this range are displayed independently. Two-sided Kruskal-Wallis Test is used for statistics analysis, P-Value=0.017, *P < 0.05.



Supplementary Figure 13 GFP design evaluated by MD simulations. (A) Superimposition of native GFP (green) and EvoPlay-designed mutant (magenta) after 200ns molecular dynamics (MD) simulation. **(B)** the RMSD trends of the GFP wild-type (GFP-WT) and designed mutants (GFP-MT1, GFP-MT2, GFP-MT3, GFP-MT4) during 200ns MD simulations. **(C)** The number of hydrogen bonds formed by the mutants and the wild type during the MD simulations.



Supplementary Figure 14 The mutants screening process of GLuc designed by EvoPlay. In GLuc design task, we firstly collect 1,500 designed sequences from 10 independent runs (repeats), and further obtain 1,047 after deduplication. We next implement K-means clustering on all candidates and select the top mutants within each cluster using a surrogate model. Finally, 36 variants are sent for plasmid construction. 29 out of these 36 variants are successfully purified and 26 of them are experimentally validated.



Supplementary Figure 15 Bio-layer interferometry characterization of the binding affinity of designed peptides to the corresponding RNase1 target. Bio-layer interferometry characterization of the binding affinity of designed peptides to the corresponding RNase1 targets. Two-fold serial dilutions are tested for each binder and the highest concentration is labelled. The target RNase1 with His tag (C terminal) are loaded onto Ni-NTA biosensors, and incubated with peptide binders in the solution to measure association and dissociation constants. The Octet signal is normalized by the maximum signal of the cognate binder–target pair. "K_D≈NA" suggests that the binding affinity is not detectable. Result of the wild type is shown in the first row.

Supplementary Tables

Supplementary Table 1 Comparison of EvoPlay against other methods on PAB1 dataset. We choose 5 sequences from the low-fitness pool as starting seeds and repeat the process 5 times for each. The "max" value represents the average cumulative maximum value of the 5 repeats, and the "mean" value is calculated by averaging the top 10 generated sequences (evaluated by the TAPE oracle) from the 5 repeats. The highest score is indicated in bold.

Start sequence	Start fitness	EvoPlay		AdaLead		SAC		Cbas		Evolutionary BO		DyNA-PPO	
		max	mean	max	mean	max	mean	max	mean	max	mean	max	mean
Sequence 1	0.020	**0.73**	**0.59**	0.63	0.49	0.51	0.33	0.53	0.33	0.32	0.28	0.50	0.33
Sequence 2	0.040	**0.55**	**0.48**	0.40	0.28	0.29	0.19	0.34	0.28	0.36	0.31	0.29	0.17
Sequence 3	0.011	**0.54**	**0.45**	0.50	0.41	0.39	0.26	0.41	0.30	0.39	0.34	0.39	0.24
Sequence 4	0.044	**0.59**	**0.55**	0.50	0.45	0.30	0.15	0.34	0.24	0.40	0.36	0.32	0.18
Sequence 5	0.29	**1.09**	0.82	1.05	**0.86**	0.45	0.36	1.00	0.58	0.37	0.34	0.91	0.60

Supplementary Table 2 Comparison of EvoPlay against other methods on GFP dataset. We choose 5 sequences from the low-fitness pool as starting seeds and repeat the process 5 times for each. The "max" value represents the average cumulative maximum value of the 5 repeats, and the "mean" value is calculated by averaging the top 10 generated sequences (evaluated by the TAPE oracle) from the 5 repeats. The highest score is indicated in bold.

Start sequence	Start fitness	EvoPlay		AdaLead		SAC		Cbas		Evolutionary BO		DyNA-PPO	
		max	mean	max	mean	max	mean	max	mean	max	mean	max	mean
Sequence 1	1.56	**3.02**	**2.14**	2.45	1.87	2.64	1.73	2.36	1.80	1.57	1.56	2.21	1.69
Sequence 2	1.43	**3.66**	2.76	3.61	**3.33**	1.97	1.79	3.54	2.39	1.95	1.76	2.73	1.93
Sequence 3	1.30	**3.65**	3.58	3.65	**3.59**	3.61	3.55	3.57	3.51	1.58	1.56	3.61	3.51
Sequence 4	1.60	**3.48**	**2.57**	2.65	1.92	3.17	2.00	2.36	1.72	1.59	1.56	2.64	1.83
Sequence 5	1.30	**3.68**	2.21	3.62	**3.05**	3.22	2.16	2.80	1.68	1.57	1.56	3.55	2.07

Supplementary Table 3 The diversity and novelty of sequences generated by reinforcement learning models evaluated on the PAB1 dataset. We select the top 100 sequences and all generated sequences from one repeat to calculate diversity and novelty, and we report the average value of five repeats. The highest score is marked in bold.

	Start sequence	EvoPlay		DyNA-PPO		SAC	
		novelty	diversity	novelty	diversity	novelty	diversity
Top 100 sequences	Sequence 1	**13.66**	5.18	5.73	10.91	3.88	7.45
	Sequence 2	**13.57**	3.77	5.79	11.03	3.80	7.31
	Sequence 3	**15.12**	3.91	5.74	10.95	3.87	7.50
	Sequence 4	**12.38**	6.55	5.24	10.02	3.74	7.22
	Sequence 5	**7.31**	5.632	2.98	5.80	2.19	4.28
All sequences	Sequence 1	**10.00**	8.25	4.74	9.16	2.92	5.72
	Sequence 2	**9.89**	8.21	4.63	8.94	3.01	5.89
	Sequence 3	**9.90**	8.18	4.71	9.10	3.06	5.98
	Sequence 4	**9.46**	7.75	4.58	8.86	3.04	5.94
	Sequence 5	**9.38**	7.92	4.17	8.10	2.61	5.12

Supplementary Table 4 The diversity and novelty of sequences generated by reinforcement learning models evaluated on the GFP dataset. We select the top 100 sequences and all generated sequences from one repeat to calculate diversity and novelty, and we report the average value of five repeats. The highest score is marked in bold.

	Start sequence	EvoPlay		DyNA-PPO		SAC	
		novelty	diversity	novelty	diversity	novelty	diversity
Top 100 sequences	Sequence 1	**14.52**	8.12	5.64	9.17	3.49	6.89
	Sequence 2	11.01	8.11	4.27	8.44	**73.03**	50.59
	Sequence 3	2.53	2.83	**3.98**	5.90	2.952	3.89
	Sequence 4	**9.90**	8.53	4.17	8.24	4.12	6.18
	Sequence 5	**15.91**	9.16	4.58	9.03	4.61	7.12
All sequences	Sequence 1	**11.91**	9.01	5.42	10.70	2.87	5.71
	Sequence 2	**12.63**	9.19	4.19	8.30	2.74	5.44
	Sequence 3	**14.17**	10.08	4.41	8.74	2.44	4.86
	Sequence 4	**13.22**	10.03	5.24	8.43	3.08	6.11
	Sequence 5	**14.20**	10.81	5.04	8.02	2.91	5.78

Supplementary Table 5 The binding free energies between RNA and PAB1 of the wild-type (WT) and EvoPlay designed mutants, predicted by MM/PBSA (in units of kcal/mol).

Model	*Type*	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{SA}	$\Delta G_{binding}$
WT	*WT-RNA*	-52.42±10.44	12.42±1.18	-9.55±0.7	-6.02±0.05	-55.57±10.53
EvoPlay	*MT1-RNA*	-58.45±9.37	6.45±2.15	-2.56±2.94	-5.92±0.00	-60.49±10.06
	MT2-RNA	-61.67±10.69	0.43±1.08	2.22±0.6	-6.56±0.03	-65.57±10.76
	MT3-RNA	-60.67±7.78	15.55±0.21	-11.48±1.03	-6.83±0.01	-63.43±7.85
	MT4-RNA	-57.18±10.94	2.51±0.5	0.74±0.58	-5.95±0.11	-59.89±10.97
AdaLead	*MT5-RNA*	-47.95±11.44	5.24±1.03	0.48±1.34	-5.74±0.08	-47.98±11.56
	MT6-RNA	-48.89±10.22	9.97±1.34	-6.03±0.53	-5.55±0.06	-50.51±10.32
	MT7-RNA	-53.83±9.6	8.55±1.08	-3.67±0.83	-6.13±0.01	-55.09±9.70
	MT8-RNA	-58.36±9.95	14.88±0.77	-9.97±0.67	-6.36±0.14	-59.80±10.00

ΔE_{vdw}: van der Waals energy.

ΔE_{ele}: electrostatic energy.

ΔG_{GB}: polar contribution to solvation.

ΔG_{SA}: no-polar contribution to solvation.

ΔGbinding: binding free energy.

Supplementary Table 6 The diversity and novelty of peptide sequences against 1ssc and 2cnz generated by EvoPlay and EvoBind starting from 5 random sequences. We choose 1,000 designs of one repeat to calculate the diversity and novelty. We report the average value of all sequences across the 5 repeats.

Peptide type	**Start sequence**	**EvoPlay**		**EvoBind**	
		novelty	diversity	novelty	diversity
1ssc	Sequence 1	**8.76**	**7.37**	8.62	4.55
	Sequence 2	8.29	**6.81**	**9.04**	4.31
	Sequence 3	**8.29**	**7.61**	6.31	3.98
	Sequence 4	8.26	**6.37**	**8.45**	3.96
	Sequence 5	**8.85**	**6.82**	8.69	4.06
2cnz	Sequence 1	**13.31**	**10.84**	11.36	5.07
	Sequence 2	**13.92**	**10.72**	11.14	4.59
	Sequence 3	13.91	**10.99**	**14.78**	6.19
	Sequence 4	13.85	**10.35**	**14.24**	5.04
	Sequence 5	**13.14**	**10.63**	8.14	3.92

Supplementary Table 7 Average time consumed to achieve the same performance on 1ssc (EvoBind recycle8 versus EvoPlay recycle0).

Type	Average time
EvoBind	10 hours 37min
EvoPlay	**3 hours 11min**

Supplementary Table 8 EvoPlay designed peptides compared with AlphaFold2 predicted native peptides. For 1ssc, EvoPlay uses Recycle 0 to design high-quality peptides similar with using Recycle8. For 2cnz, EvoPlay can design high-quality 2cnz at Recycle 3. 2cnz and 1ssc belong to the same class of proteins (uncleaved preproteins or similar synthetic fusion proteins), and thereby with similar design difficulty[20]. For 3r7g, regardless of the recycle numbers, AlphaFold2 can't perform well in predicting the structure of native complex (high pLDDT for receptor, but low for peptide and overall low pLDDT). However, EvoPlay can still produce relatively complex with high pLDDT (79.41) and low loss (0.0932). For 6seo with relatively long centroid distance, EvoPlay-designed peptides have divergent binding pattern with native one.

PDBID	Sequence	Recycle	pLDDT	Loss	IF_dist_peptide	Delta_COM
1ssc-EvoPlay	**EIVCKKIISVY**	**0**	**91.95**	**0.0036**	**4.01**	**0.07**
1ssc-Native	PYVPVHFDASV	0	43.72	0.1057	4.16	0.93
1ssc-Native	PYVPVHFDASV	3	78.51	0.0431	4.16	0.93
1ssc-Native	PYVPVHFDASV	5	93.49	0.0094	3.96	0.73
1ssc-Native	PYVPVHFDASV	8	95.13	0.0089	3.86	0.2
1ssc-Native	PYVPVHFDASV	10	95.62	0.0087	3.86	0.19
2cnz-EvoPlay	**ERQVADGCHGHFLVKVHFNM**	**3**	**91.24**	**0.0451**	**4.03**	**0.90**
2cnz-Native	GSFLPNSEQQKSADIVFSSP	0	84.19	0.0820	3.92	1.55
2cnz-Native	GSFLPNSEQQKSADIVFSSP	3	94.39	0.0739	3.99	1.57
2cnz-Native	GSFLPNSEQQKSADIVFSSP	5	95.45	0.0736	3.99	1.58
2cnz-Native	GSFLPNSEQQKSADIVFSSP	8	95.64	0.0734	3.99	1.58
2cnz-Native	GSFLPNSEQQKSADIVFSSP	10	95.63	0.0734	3.99	1.58
3r7g-EvoPlay	**GHVVAAAQQPNWFAWNQKAHSV**	**3**	**79.41**	**0.0932**	**5.04**	**1.42**
3r7g-Native	KSLYKIKPRHDSGIKAKISMKT	0	34.20	1.7010	3.86	0.19
3r7g-Native	KSLYKIKPRHDSGIKAKISMKT	3	34.01	2.1255	5.04	1.42
3r7g-Native	KSLYKIKPRHDSGIKAKISMKT	5	34.71	2.4577	9.68	5.87
3r7g-Native	KSLYKIKPRHDSGIKAKISMKT	8	34.52	1.9919	9.24	8.01
3r7g-Native	KSLYKIKPRHDSGIKAKISMKT	10	34.64	1.7078	10.34	8.85
6seo-EvoPlay	**GHTHRKRLHADGRDIWIS**	**3**	**72.79**	**0.1134**	**5.51**	**1.34**
6seo-Native	AVPLRARNLPPSFFTEPA	0	31.19	28.4123	9.44	7.43
6seo-Native	AVPLRARNLPPSFFTEPA	3	69.63	0.0992	8.67	6.62
6seo-Native	AVPLRARNLPPSFFTEPA	5	81.76	0.6494	5.51	1.34
6seo-Native	AVPLRARNLPPSFFTEPA	8	85.50	0.6402	24.33	33.93
6seo-Native	AVPLRARNLPPSFFTEPA	10	85.89	0.6540	4.92	1.39

Supplementary Table 9 Binding free energies between peptide and receptor protein predicted by MM/PBSA (kcal/mol) for 1ssc, 2cnz, 3r7g, 6seo.

PDB	Type	Native	EvoPlay-recycle0	EvoBind-recycle8
1ssc	ΔE_{vdw}	-89.58±5.17	-97.08±0.84	-94.1±1.16
	ΔE_{ele}	-310.26±28.06	-306.6±4.15	-382.35±3.1
	ΔG_{GB}	336.05±26.78	326.29±3.44	401.98±4.29
	ΔG_{SA}	-11.07±0.61	-14.74±0.05	-12.96±0.07
	$\Delta G_{binding}$	-74.86±5.79	-96.13±5.45	-87.43±5.42
2cnz	ΔE_{vdw}	-142.44±1.05	-149.93±0.23	-134.82±1.72
	ΔE_{ele}	-431.04±19.21	-577.53±3.05	-278.7±4.44
	ΔG_{GB}	457.52±6.42	589.89±2.28	311.75±7.00
	ΔG_{SA}	-18.76±0.20	-20.37±0.19	-16.48±0.14
	$\Delta G_{binding}$	-134.73±20.28	**-157.94±3.82**	-118.26±8.47
3r7g	ΔE_{vdw}	-80.37±2.96	-89.07±1.51	-55.61±3.08
	ΔE_{ele}	-1472.63±18.9	-991.25±28.11	-775.51±11.93
	ΔG_{GB}	1485.04±15.89	997.56±22.14	789.62±7.53
	ΔG_{SA}	-13.77±0.17	-13.73±0.1	-9.31±0.22
	$\Delta G_{binding}$	-81.74±24.87	**-96.48±35.81**	-50.81±14.44
6seo	ΔE_{vdw}	-82.49±6.64	-70.63±4.03	-41.7±4.73
	ΔE_{ele}	-475.13±4.82	-147.37±5.95	268.82±0.5
	ΔG_{GB}	512.67±2.72	183.54±1.37	283.83±1.16
	ΔG_{SA}	-12.21±0.68	-9.91±0.31	-7.24±0.52
	$\Delta G_{binding}$	-57.15±6.44	**-44.37±7.08**	-33.93±4.92

ΔE_{vdw}: van der Waals energy.

ΔE_{ele}: electrostatic energy.

ΔG_{GB}: electrostatic contribution to solvation.

ΔG_{SA}: no-polar contribution to solvation.

ΔGbinding: binding free energy.

Supplementary Table 10 Binding free energies between GLuc and CTZ predicted by MM/PBSA (kJ/mol).

Type	GLuc-WT-CTZ	GLuc-MT1-CTZ	GLuc-MT2-CTZ
ΔE_{vdw}	-180.87+/-14.47	-221.448+/-11.95	-219.63+/-12.14
ΔE_{ele}	-105.46+/-17.33	-75.83+/-14.85	-76.52+/-12.78
ΔG_{GB}	213.46+/-17.55	201.32+/-14.74	194.14+/-8.83
ΔG_{SA}	-18.57+/-1.03	-20.01+/-0.92	-20.75+/-0.93
$\Delta G_{binding}$	**-91.44+/-11.57**	**-115.95+/-15.03**	**-122.76+/-23.40**

ΔE_{vdw}: van der Waals energy.

ΔE_{ele}: electrostatic energy.

ΔG_{GB}: electrostatic contribution to solvation.

ΔG_{SA}: no-polar contribution to solvation.

ΔGbinding: binding free energy.

Supplementary Table 11 The architecture of EvoPlay.

Input:	the current state
1D-CNN	Number of filters:32, kernel size:3 ReLU activation
1D-CNN	Number of filters:64, kernel size:3 ReLU activation
1D-CNN	Number of filters:128, kernel size:3 ReLU activation
Action policy layers:	State value layers:
1D-CNN: Number of filters:80, kernel size:3, ReLU activation	1D-CNN: Number of filters:20, kernel size:3, ReLU activation
Dense: Output size: L_{seq} *20, Softmax activation	Dense: Output size: 128, ReLU activation Dense: Output size: 1, tanh activation
Output: action policy layer output, state value layer output	

Supplementary Table 12 Ablation results of surrogate models on GFP dataset. Three metrics (Mean Squared Error, R-square, Pearson correlation coefficient (PCC)) are calculated (5-fold cross-validation). Errors represent the standard deviation of 5 folds.

Model	MSE	R_square	PCC	Each_fold_time(s)
UniRep	0.0087±0.0005	0.9375±0.0036	0.9683±0.0018	5360±116
CNN	0.0082±0.0008	0.9412±0.0065	0.9709±0.0028	60.78±5.45
MuFacNet	0.0108±0.0012	0.9150±0.0122	0.9606±0.0042	486.37±7.45
GP	0.0817±0.0011	0.1093±0.0203	0.7130±0.0071	7587±149
Hybrid	0.064±0.0114	0.4680±0.0881	0.744±0.0089	62.02±1.97

5 surrogate models including CNN, GP, UniRep, MuFacNet, Hybrid are evaluated on GFP dataset. CNN model has the highest Pearson correlation coefficient (PCC=0.97) with the best efficiency. UniRep model performs relatively well but is time-consuming. We adopt CNN as the surrogate model for our agent.

Surrogate model:

UniRep: A pre-trained LSTM-based protein language model**.**

CNN: A convolutional neural network, architecture can be found in the Supplementary Table 13.

MuFacNet: A Mutation Factorization Network from PEX, factorizing the composition of mutational effects to the interactions among mutations.

GP: A simple Gaussian process regression model from sklearn package.

Hybrid: A hybrid model by combining a linear model trained on experimentally labelled sequences with a probability feature derived from evolutionary related sequences.

Supplementary Table 13 The architecture of CNN surrogate model.

Input:	Input with $L_{seq} \times 20$
1D-CNN	Number of filters:32,kernel size:5
	ReLU activation
1D-CNN	Number of filters:32,kernel size:5
	ReLU activation
	Max-pooling
1D-CNN	Number of filters:32,kernel size:5
	ReLU activation
	Max-pooling
Dense	Output size: 100, ReLU activation
Dense	Output size: 100, ReLU activation
	Dropout(p=0.25)
Dense	Output size: 1
Fitness Prediction:	A scalar value

Supplementary Table 14 The number of designed peptides targeting each protein. We count the number of Simulation, Move and End, and the median of 25 experiments for each protein is listed.

PDBID	Simulated count	Move count	End count
1ssc (simulation 22)	847	579	1,003
2cnz (simulation 22)	867	623	1,000
3r7g (simulation 11)	430	616	1,000
6seo (simulation 22)	886	425	1,001

Supplementary Table 15 Different in silico evaluation metrics for selected peptides designed by EvoPlay，EvoBind and MCMC sent for BLI assays.

Model	Peptide	pLDDT	TM-score	RMSD
Wild Type	PYVPVHFDASV	-	-	-
EvoPlay	YIQCKEIISVE	90.31	0.96	1.07
	GAPCSHICNTD	89.10	0.96	1.03
	AFFCTKMIHTT	85.58	0.96	1.02
	HCAPHQVVDYK	90.14	0.96	1.00
	PRSCLAWMGAR	90.35	0.95	1.15
EvoBind	PITLSALYIDD	88.04	0.95	1.22
	NYEVHKMISIS	86.85	0.94	1.24
	DPHCHNCHPSS	87.71	0.94	1.3
	SHHCHQLHKLP	85.43	0.88	2.87
	GNCVRWAGTAN	89.79	0.95	1.12
MCMC	KKHCEQLQSTN	87.72	0.95	1.16
	LPTCFGGVRPA	91.30	0.95	1.19
	ECRATKVMGYV	90.46	0.96	1.02
	SQQAVPYIEHD	84.51	0.96	1.09
	HQVVELLAMYN	89.19	0.92	1.50

Supplementary Table 16 The average count of sequences with a fitness score above 0.5 generated by EvoPlay and other methods on the PAB1 dataset. We report the average count of sequences generated by EvoPlay and other models (AdaLead, SAC and DyNA-PPO) with a fitness score above 0.5, based on five repeats of five starting sequences. EvoPlay can produce more high-quality sequences.

Start sequence	Threshold	EvoPlay	AdaLead	SAC	DyNA-PPO
Sequence 1	0.5	8.2	3.2	1.2	0.4
Sequence 2	0.5	72.4	0.2	0.0	0.0
Sequence 3	0.5	20.6	1.0	0.2	0.0
Sequence 4	0.5	70.6	1.6	0.0	0.0
Sequence 5	0.5	18.2	34.0	7.8	6.4

Supplementary Table 17 The average count of sequences with a fitness score above 3.5 generated by EvoPlay and other methods on the GFP dataset. We report the average count of sequences generated by EvoPlay and other models (AdaLead, SAC and DyNA-PPO) with a fitness score above 3.5, based on five repeats of five starting sequences. EvoPlay can produce more high-quality sequences.

Start sequence	Threshold	EvoPlay	AdaLead	SAC	DyNA-PPO
Sequence 1	3.5	0.0	0.0	0.0	0.0
Sequence 2	3.5	3.2	8.0	0.0	1.0
Sequence 3	3.5	55.4	29.8	13.0	3.4
Sequence 4	3.5	8.8	0.6	0.4	0.0
Sequence 5	3.5	102.2	3.6	1.0	0.8

Supplementary Table 18 The high entropy position predicted by the language model with partial overlapping with PSSM matrix and DisEMBL. Among the 35 high entropy positions predicted by the language model, 12 were found to coincide with the high entropy positions predicted by the PSSM matrix, while 23 positions overlapped with the disordered regions predicted by the disorder prediction tool DisEMBL.

Methods	High entropy position
Language model	13, 23, 37, 44, 45, 46, 55, 64, 67, 68, 71, 75, 76, 78, 79, 86, 92, 137, 146, 151, 155, 167, 175, 178, 179, 181, 182, 190, 197, 202, 215, 216, 229, 231, 232
The overlapped positions with PSSM matrix	79, 202, 78, 55, 167, 92, 67, 197, 215, 64, 179, 229
The overlapped positions with DisEMBL	23, 37, 46, 55, 67, 68, 71, 75, 76, 78, 79, 86, 137, 146, 155, 181, 182, 190, 197, 202, 229, 231, 232

Supplementary Table 19 The mutants used in MD simulations of Figure 2d for PAB1. MD simulations on randomly selected mutants from the top 1% designs were performed in Figure 2d.

Algorithms	Mutants	Mutant sites
EvoPlay	MT1	N132G, P135L, E161F, L186C
	MT2	N132C, P135L, D151F, E161F, G167C, L186C, G188T
	MT3	N132G, P135L; MT4: N132G, P135L, E161F
	MT4	N132G, P135L, E161F
Adalead	MT5	N132M, N139V, G167F, A182C, Q194F
	MT6	N132I, N139S, L142Q, F146M, E175P, A179L, E181D
	MT7	N132A, P135W, N139Y, K140Q, D144I, G150S, L153S, G167M, E174K, G177V, K180R, I183D
	MT8	N132F, L133D, H134N, N139D, A141R, D144C, F146A, S147M, F149I, L153H, T159Y, L190D, Y197W

References

1. Sinai, S. et al. AdaLead: A simple and robust adaptive greedy search algorithm for sequence design. Preprint at https://arxiv.org/abs/2010.02141 (2020).
2. Brookes, D., Park, H. & Listgarten, J. Conditioning by adaptive sampling for robust design. In Proc. 36th International Conference on Machine Learning 773-782 (PMLR, 2019).
3. Angermueller, C. et al. Model-based reinforcement learning for biological sequence design. International conference on learning representations (ICLR, 2020).
4. Haarnoja, T., Zhou, A., Abbeel, P. & Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. Proceedings of the 35th International Conference on Machine Learning 1861-1870 (PMLR, 2018).
5. Silver, D. et al. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 362, 1140-1144 (2018).
6. Tareen, A. & Kinney, J. B. Logomaker: beautiful sequence logos in Python. Bioinformatics 36, 2272-2274 (2020).
7. Bryant, P. & Elofsson, A. EvoBind: in silico directed evolution of peptide binders with AlphaFold. bioRxiv (2022).
8. Anishchenko, I. et al. De novo protein design by deep network hallucination. Nature 600, 547-552 (2021).
9. Qiu, Y., Hu, J. & Wei, G.-W. Cluster learning-assisted directed evolution. Nat. Comput. Sci. 1, 809-818 (2021).
10. Wittmann, B. J., Yue, Y. & Arnold, F. H. Informed training set design enables efficient machine learning-assisted directed protein evolution. Cell Syst. 12, 1026-1045. e1027 (2021).
11. Bergstra, James, Daniel Yamins, and David Cox. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. Proceedings of the 30th International Conference on Machine Learning 115-123 (PMLR, 2013).
12. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. Journal of computational chemistry 26, 1701-1718 (2005).
13. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins: Structure, Function, and Bioinformatics 78, 1950-1958 (2010).
14. Lu, T. Sobtop, Version 1.0 (dev3.1), http://sobereva.com/soft/Sobtop (accessed on 2022-Aug-9). (2022).
15. Neese, F. Software update: The ORCA program system—Version 5.0. Wiley Interdisciplinary Reviews: Computational Molecular Science 12, e1606 (2022).
16. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926-935 (1983).
17. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089-10092 (1993).
18. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. J. Chem. Phys. 126, 014101 (2007).
19. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. J. Appl. Phys. 52, 7182-7190 (1981).
20. Tsaban, T. et al. Harnessing protein folding neural networks for peptide–protein docking. Nat.

Commun. 13, 1-12 (2022).