A Implementation Details and Experiment Settings

A.1 Implementation Details

Details of ERLBioSeq. Our implementation of ERLBioSeq involves two algorithms: the EA Adalead algorithm (Sinai et al. 2020) and the RL algorithm Dynappo (Angermueller et al. 2019), our implementation refers to the implementation of these two algorithms. In alignment with previous studies (Jain et al. 2022), different parameters are utilized to address diverse problems. In DNA and RNA design, the RL learning rate is established at 0.001; for protein 3msi design, it is set to 0.03. The Adam optimizer is applied uniformly across all design tasks. In the evolutionary context, the parameter κ, employed to control screening sequences with high fitness scores, is configured to 0.05.

Details of LGFEN. In the fitness prediction model LGFEN, the architecture combines CNN with BI-LSTM layer to predict sequence fitness. The 1D-CNN employs 32 filters with different kernel sizes. The concatenated output of these convolutional layers is then fed into Bi-LSTM, capturing features in both forward and backward directions; the LSTM units are configured to 32. Subsequently, the output of the BI-LSTM is flattened and then input into a fully connected layer with ReLU activation. The hidden size parameter of the fully connected layer is set to 100. For predictions, a single-node output layer is used. The model is trained using the MSE loss function and the Adam optimizer with a learning rate of 0.001.

Details of the MutationAndCrossover. The $MutationAndCrossover$ process consists of two elements: mutation and crossover. Within this algorithm, Initially, we filter sequences based on fitness scores to obtain the set R, and subsequently create pairs. At each position of every sequence pair, there is a probability p_1 for an exchange, followed by sequence mutation. For each sequence, a random mutation occurs at each position with a probability p_2. This cycle continues until new $v \cdot B$ sequences are generated and placed in the mutant set M. During the experiment, we set $p_1 = 0.9$ and $p_2 = 0.1$.

A.2 Experiment Settings

In DNA and RNA design tasks, we performed 10 rounds, with 50 sequences proposed in each round. Due to the notably enlarged search space within the protein design task, we carried out a total of 20 search rounds, suggesting 200 sequences per round. To ensure experimental reliability, each result is derived from 10 independent runs. In evolutionary algorithms, the $v \cdot B$ is 500 per round for each task. Neither our ERLBioSeq nor the baselines employ any training data other than that generated during the design rounds.

For each design task, we normalized the binding affinity values of DNA and transcription factors in the DNA design task. In the RNA design problem, we followed the normalization setting from Sinai et al. (2020). However, for the protein design task, we diverged from the experimental protocol of Sinai et al. (2020). We abstained from using sigmoid to modify the output value of PyRosetta. We made this choice due to the concern that such manipulation could excessively simplify the design challenge. Since sigmoid maps value greater than 5 to values near 1, it makes the results of different algorithms very similar. If this approach were adopted, the performance gap between the algorithms might seem minimal, but the actual disparity is significant. Consequently, we decided to discard the sigmoid value mapping and retained the original output of PyRosetta.

B Additional Results

B.1 Additional Results of RNA

To comprehensively evaluate the RNA design problem, we compared the mean value of the objective function $f(x)$. Tables 4 and 5 present the top 10, 50, and 100 results for RNA lengths of 14 and 50, respectively. The findings demonstrate the state-of-the-art performance of our method. Furthermore, we conducted a detailed analysis of the distribution of the top 50 RNA sequences with lengths of 14 and 50, visualizing them through violin plots in Figure 4. It is evident that our algorithm outperforms the baseline in terms of sequence distribution.

B.2 Additional Results of Protein 3MSI

We conducted a thorough assessment of the protein 3MSI design as well. We compared the fitness averages of the top 10, 50, and 100 sequences. The outcomes are depicted in Table 6, revealing the superior performance of our ERLBioSeq algorithm compared to existing methods. In Figure 5, a violin plot comparison of the top 50 protein 3msi sequences designed by different algorithms illustrates our method's superiority over the baseline approach.

	ERLBioSeq	Adalead	Bo	Cbas	Cmaes	Dynappo	GflowNet
Top10 mean	**0.976±0.043**	0.891±0.069	0.770±0.095	0.633±0.046	0.744±0.091	0.806±0.061	0.843±0.063
Top50 mean	**0.927±0.045**	0.841±0.072	0.661±0.097	0.558±0.050	0.646±0.096	0.738±0.062	0.760±0.071
Top100 mean	**0.899±0.049**	0.809±0.078	0.587±0.105	0.522±0.052	0.590±0.101	0.697±0.067	0.706±0.083

Table 4: Mean fitness scores for the top 10, 50, and 100 RNA sequences with a length of 14.

	ERLBioSeq	Adalead	Bo	Cbas	Cmaes	Dynappo	GflowNet
Top10 mean	**0.805±0.043**	0.703±0.045	0.481±0.047	0.363±0.026	0.496±0.033	0.658±0.041	0.660±0.018
Top50 mean	**0.768±0.048**	0.669±0.048	0.435±0.049	0.336±0.026	0.453±0.033	0.622±0.040	0.627±0.022
Top100 mean	**0.737±0.055**	0.645±0.052	0.404±0.054	0.321±0.027	0.428±0.037	0.601±0.044	0.610±0.026

Table 5: Mean fitness scores for the top 10, 50, and 100 RNA sequences with a length of 50.

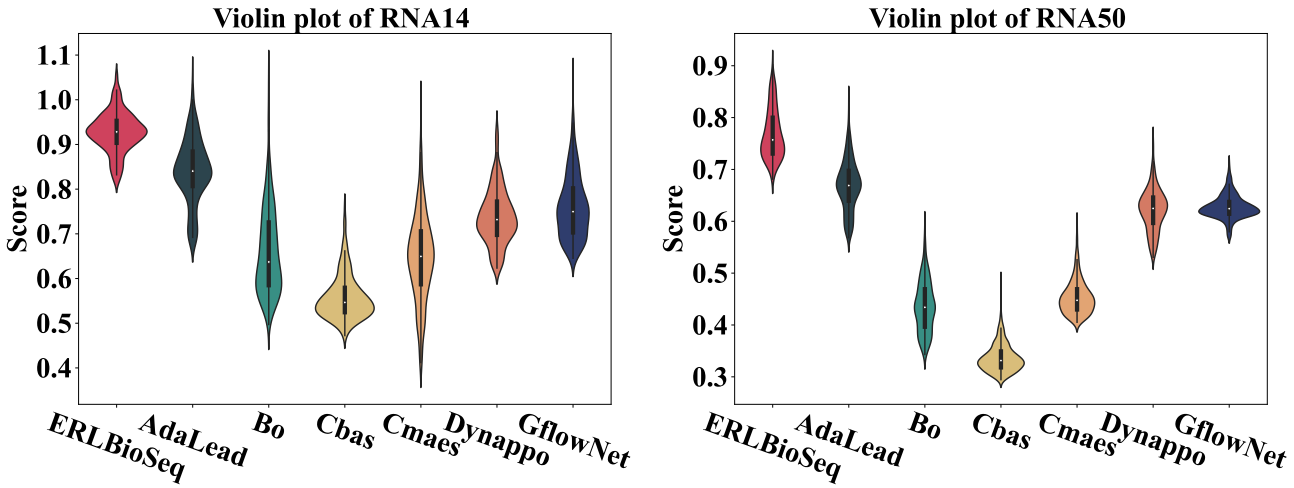

Figure 4: RNA violin plot distribution comparison. The left side displays the violin plot for RNA sequences with a length of 14, while the right side shows the violin plot for RNA sequences with a length of 50.

B.3 Ablation Experiment of LFFEN

We also performed experiments akin to those of Ren et al. (2022). Our aim was to investigate whether our proposed evolutionary reinforcement learning method could yield improved results when using a traditional CNN instead of LGFEN. We compared the performance of the CNN as a predictor against the baseline. The comparison results are presented in Table 7, demonstrating that our evolutionary reinforcement learning method can achieve superior performance even without utilizing LGFEN.

B.4 The Significance of the Results

In order to test the significance of our results, we employed t-test to assess the significance of differences between our method and others in the top 10 values. All our results with p-values < 0.05, the p-values as illustrated in Table 8, results show a significant difference.

	ERLBioSeq	Adalead	Bo	Cbas	Cmaes	Dynappo	GflowNet
Top10 mean	**14.621±2.218**	11.378±1.344	1.923±0.671	-0.358±0.363	1.060±0.581	3.259±1.549	0.067±0.305
Top50 mean	**14.239±2.206**	10.995±1.280	1.505±0.650	-0.679±0.342	0.326±0.595	2.191±1.579	-0.498±0.350
Top100 mean	**14.020±2.204**	10.778±1.251	1.179±0.657	-0.850±0.342	-0.085±0.624	1.657±1.583	-0.771±0.375

Table 6: Mean fitness scores for the top 10, 50, and 100 biological sequences of protein 3msi with a length of 66.

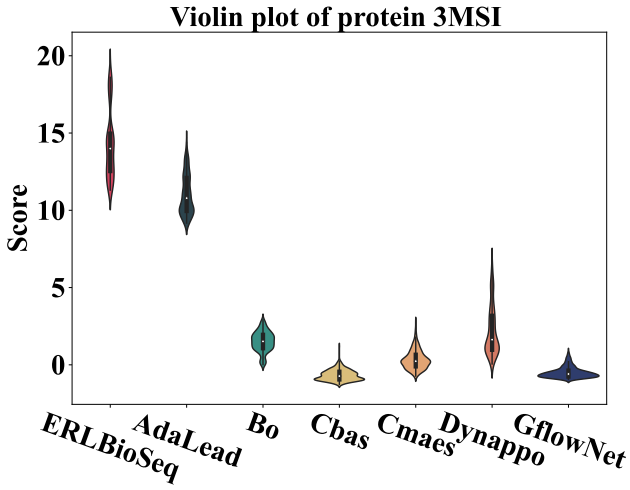

Figure 5: This violin plot compares the distribution of the top 50 sequences from the protein 3msi design task.

		ERLBioSeq + CNN	Adalead	Bo	Cbas	Cmaes	Dynappo	GflowNet
Top10 mean	DNA	**0.987±**	0.984±	0.958±	0.974±	0.967±	0.937±	0.870±
		0.007	0.007	0.021	0.013	0.022	0.026	0.071
	RNA14	**0.942±**	0.891±	0.770±	0.633±	0.744±	0.806±	0.843±
		0.054	0.069	0.095	0.046	0.091	0.061	0.063
	RNA50	**0.713±**	0.703±	0.481±	0.363±	0.496±	0.658±	0.660±
		0.040	0.045	0.047	0.026	0.033	0.041	0.018
	Protein	**14.064±**	11.378±	1.923±	-0.358±	1.06±	3.259±	0.067±
		2.429	1.344	0.671	0.363	0.581	1.549	0.305
Top50 mean	DNA	**0.965±**	0.963±	0.878±	0.928±	0.882±	0.863±	0.735±
		0.017	0.016	0.060	0.034	0.066	0.052	0.113
	RNA14	**0.893±**	0.841±	0.661±	0.558±	0.646±	0.738±	0.760±
		0.057	0.072	0.097	0.050	0.096	0.062	0.071
	RNA50	**0.677±**	0.669±	0.435±	0.336±	0.453±	0.622±	0.627±
		0.043	0.048	0.049	0.026	0.033	0.040	0.022
	Protein	**13.679±**	10.995±	1.505±	-0.679±	0.326±	2.191±	-0.498±
		2.418	1.280	0.650	0.342	0.595	1.579	0.350
Top100 mean	DNA	**0.943±**	0.941±	0.801±	0.886±	0.786±	0.800±	0.629±
		0.028	0.026	0.094	0.053	0.116	0.079	0.159
	RNA14	**0.868±**	0.809±	0.587±	0.522±	0.590±	0.697±	0.706±
		0.058	0.078	0.105	0.052	0.101	0.067	0.083
	RNA50	**0.650±**	0.645±	0.404±	0.321±	0.428±	0.601±	0.610±
		0.048	0.052	0.054	0.027	0.037	0.044	0.026
	Protein	**13.449±**	10.778±	1.179±	-0.850±	-0.085±	1.657±	-0.771±
		2.410	1.251	0.657	0.342	0.624	1.583	0.375

Table 7: Mean comparison of top-k in ERLBioSeq without LGFEN. Use CNN instead of LGFEN to compare the mean fitness of the TopK sequences against the baseline.

	Adalead	Bo	Cbas	Cmaes	Dynappo	GflowNet
DNA	3.181e-07	1.978e-31	1.234e-20	4.612e-18	3.508e-46	1.248e-39
RNA14	1.731e-20	2.455e-48	3.305e-121	4.092e-58	5.559e-57	1.435e-41
RNA50	1.822e-38	8.542e-115	6.259e-160	6.565e-124	4.459e-62	4.389e-77
Protein	1.250e-26	3.395e-121	3.680e-137	2.224e-127	3.661e-100	4.150e-135

Table 8: The significance of the results of our method compared to the alternative method.