Supplemental information

ACP-TransLSTM: A Novel Deep Learning Framework for Anticancer Peptide Prediction Using Multi-Source Feature Integration

Jinxin Liu[1], Zhenming Wu[1] and Jin Zhao[1*]

[1] School of Computer Science and Technology, Qingdao University, Ningxia Road, 266071, Qingdao, Shandong, China
zhaojin@qdu.edu.cn

Contents

1. Supplemental Notes

Supplemental Note 1.1: Hyperparameter Setting

In the CNN + Bi-LSTM model, we first tested the CNNs individually and in order to find the best hidden layer settings, we chose different numbers of filter layers. We selected six different filter sizes: 8, 16, 32, 64, 128 and 256. As illustrated in Table S16, the model with 256 filters achieved the best performance and the highest AUC value. The choice of 256 filters is a trade-off between computational complexity and the ability to capture diverse local patterns in peptide sequences. A larger number of filters can potentially capture more complex patterns but may lead to overfitting and increased computational cost. When we decreased the number of CNN filters from 256 to 128, the average ACC across the six datasets decreased from 0.923 to 0.918, and the AUC decreased from 0.984 to 0.976. This indicates that reducing the number of filters weakens the model's ability to capture local sequence patterns, leading to a decline in performance. Therefore, we chose to use 256 filters when building the model. After determining the parameters of the CNN model, we compared the performance of different hidden cell counts (8, 16, 32, 64, 128, 256) in order to find the optimal hidden cell settings. The model with 64 hidden cells achieving the best performance and the highest AUC value. This number of units is determined through a series of preliminary experiments. A smaller number of units may not be able to fully capture the long-range dependencies in peptide sequences, resulting in suboptimal performance. Conversely, an excessive number of units can cause overfitting. Adjusting the number of Bi-LSTM units also affected the model's performance. When the number of units was reduced to 32, the ACC across the six datasets dropped from 0.916 to 0.902, and the MCC decreased from 0.88 to 0.856. This shows that fewer units limit the model's ability to capture long-range dependencies. On the other hand, increasing the number of units to 128 led to overfitting.

To accurately determine the optimal critical probability for differentiating between ACPs and non-ACPs, we carried out a series of experiments on six datasets. We selected several different probability values, namely 0.3, 0.4, 0.5, 0.6 and 0.7 for testing. The experimental results show that when the critical probability is set to 0.5, the model exhibits the most excellent performance.

Supplemental Note 1.2: Dataset Information

ACPred-Fuse: Positive samples were sourced from the CancerPPD ACP database, while negative samples were obtained from Swiss-Prot. The negative samples were selected to represent non-anticancer peptides, mainly those with no reported anticancer function. CD-HIT was used to remove sequences with similarity above 0.8. The dataset consists of 250 positive and 250 negative samples for training and 82 positive and 2,628 negative samples for testing. In the test set of the ACPred-Fuse dataset, there exists a substantial disparity in the ratio of positive to negative samples.

Specifically, there are merely 82 positive samples, in contrast to 2,628 negative samples. This imbalance can be attributed to the origin of the samples. The negative samples were obtained from a large-scale general protein database, Swis-Prot. Given that Swiss-Prot contains mostly non-anticancer peptides, the number of negative samples is significantly large. Conversely, the positive samples were meticulously chosen from a specialized ACP database, which accounts for their relatively small quantity. samples.
ACPred-FL: Comprises 250 positive and 250 negative training samples, with a test set of 82 positive and 82 negative samples.
ACP240: Includes 129 experimentally validated anticancer peptides as positive samples and 111 antimicrobial peptides without known anticancer activity as negative samples.
ACP740: CD-HIT filtered out sequences with similarity above 0.9. The final dataset contains 376 positive and 364 negative samples. ACP240 and ACP740 have no overlapping sequences, and both underwent 5-fold cross-validation.
ACPmain: Consists of 861 positive ACPs and 861 negative samples compiled from ACP-DL, ACPP, ACPred-FL, AntiCP, and iACP.
ACP530: Contains 265 positive ACPs and 265 negative samples randomly selected from Swiss-Prot.

Supplemental Note 1.3: Evaluation Metrics

We compared our model with state-of-the-art methods, including ACP-DA, ACP-DL, ACP-MHCNN, iACP, CL-ACP, ACP-check, ACPScanner and TriNet. The evaluation was conducted using multiple performance metrics, including accuracy (ACC), Matthews correlation coefficient (MCC), sensitivity (SE), specificity (SP), area under the ROC curve (AUC) and F1-score. The mathematical formulations for these evaluation metrics are defined as follows:

$$AAC = \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

$$MCC = \frac{(TP\times TN)-(FN\times FP)}{\sqrt{(TP+FN)\times(TN+FP)\times(TP+FP)\times(TN+FN)}} \tag{2}$$

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{3}$$

$$SP = \frac{TN}{FP+TN} \tag{4}$$

$$F1-score = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{5}$$

where TP (true positive) represents correctly predicted positive samples, TN (true negative) denotes correctly predicted negative samples, FP (false positive) refers to negative samples incorrectly classified as positive and FN (false negative) represents positive samples incorrectly classified as negative. Additionally, the precision ($\frac{TP}{TP+FP}$)

measures the proportion of correctly predicted positive samples among all predicted positives and recall ($\frac{TP}{\mathrm{TP+FN}}$) represents the proportion of correctly identified positive samples out of all actual positives.

2. Supplemental Figures

Supplemental Figure 2.1: Amino acid composition information

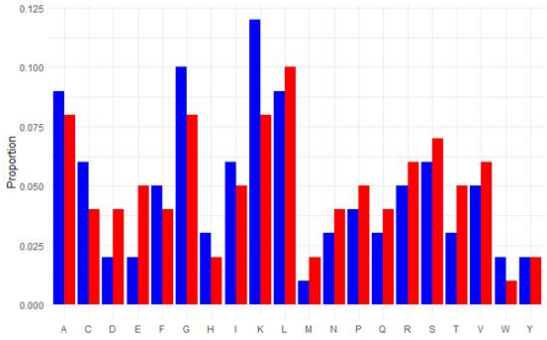

Fig. S1. Amino acid composition of ACPs and non-ACPs on ACP240 dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

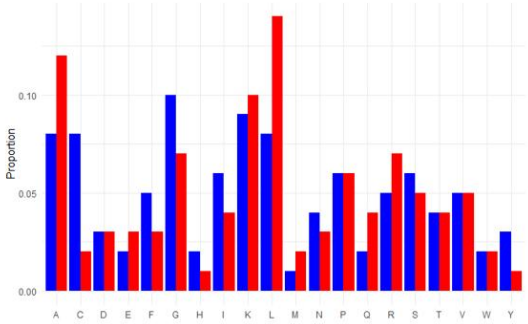

Fig. S2. Amino acid composition of ACPs and non-ACPs on ACP740 dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

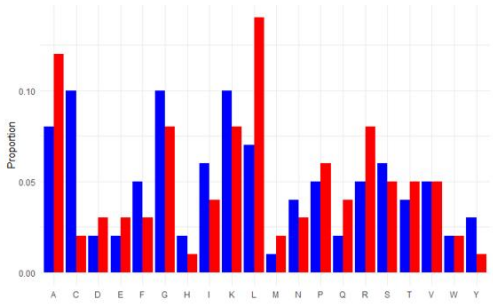

Fig. S3. Amino acid composition of ACPs and non-ACPs on ACP530 dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

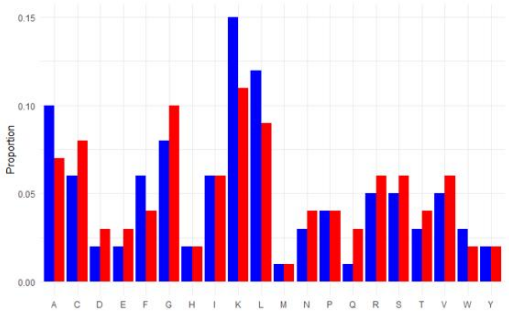

Fig. S4. Amino acid composition of ACPs and non-ACPs on ACPmain dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

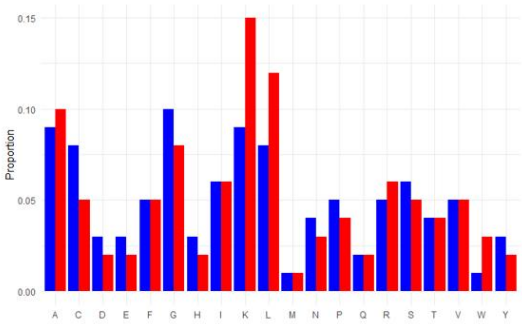

Fig. S5. Amino acid composition of ACPs and non-ACPs on ACPred-FL dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

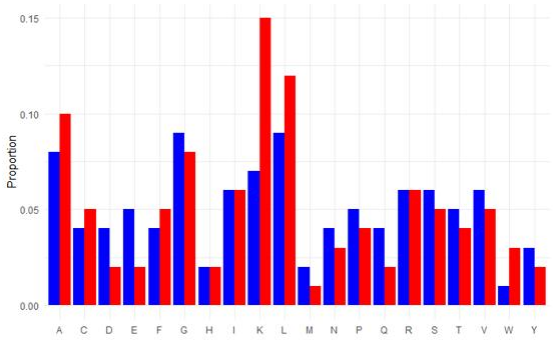

Fig. S6. Amino acid composition of ACPs and non-ACPs on ACPred-Fuse dataset. The type represented by blue is ACPs, and the type represented by red is non-ACPs.

Supplemental Figure 2.2: Charge distribution

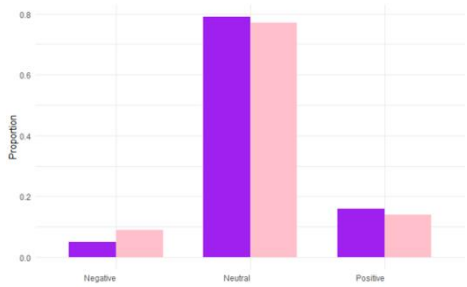

Fig. S7. Charge distribution of ACPs and non-ACPs on ACP240 dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

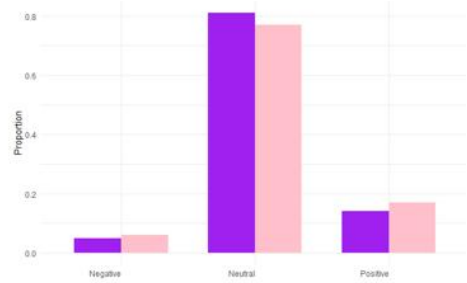

Fig. S8. Charge distribution of ACPs and non-ACPs on ACP740 dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

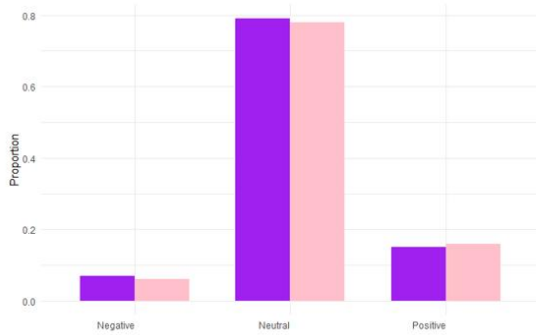

Fig. S9. Charge distribution of ACPs and non-ACPs on ACP530 dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

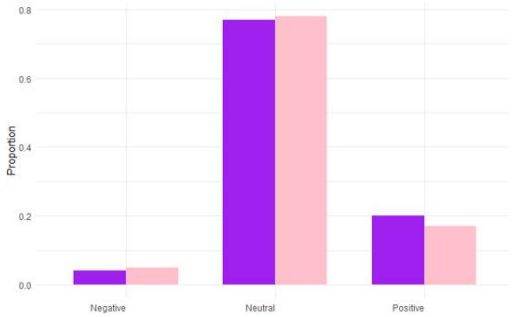

Fig. S10. Charge distribution of ACPs and non-ACPs on ACPmain dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

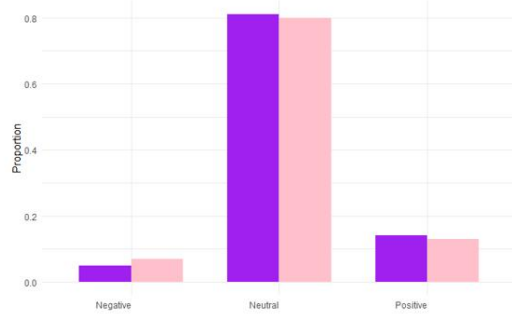

Fig. S11. Charge distribution of ACPs and non-ACPs on ACPred-FL dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

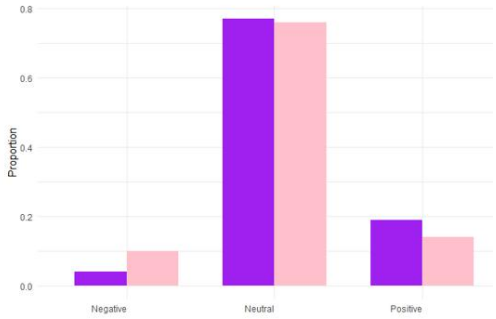

Fig. S12. Charge distribution of ACPs and non-ACPs on ACPred-Fuse dataset. The type represented by purple is ACPs, and the type represented by pink is non-ACPs.

Supplemental Figure 2.3: Length distribution

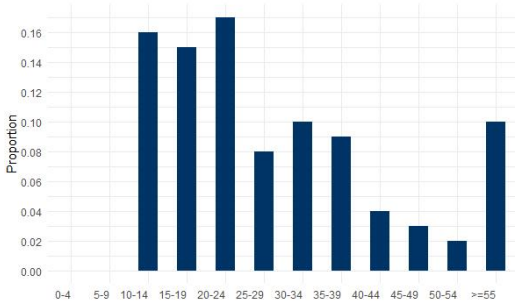

Fig. S13. Length distribution of ACPs and non-ACPs on ACP240 dataset.

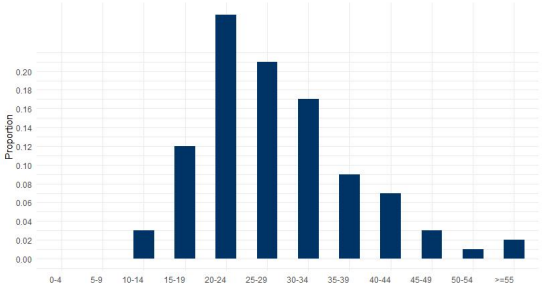

Fig. S14. Length distribution of ACPs and non-ACPs on ACP740 dataset.

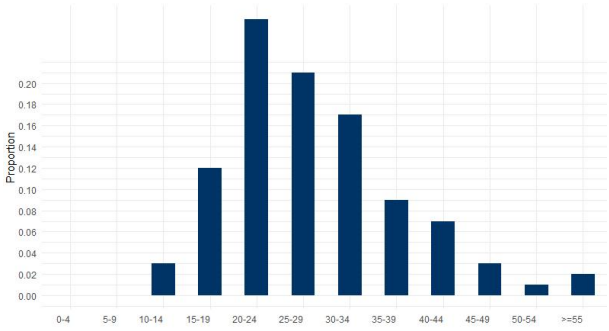

Fig. S15. Length distribution of ACPs and non-ACPs on ACP530 dataset.

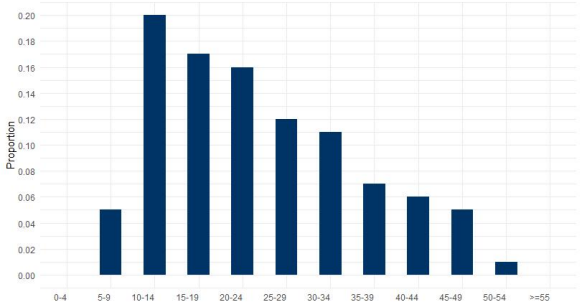

Fig. S16. Length distribution of ACPs and non-ACPs on ACPmain dataset.

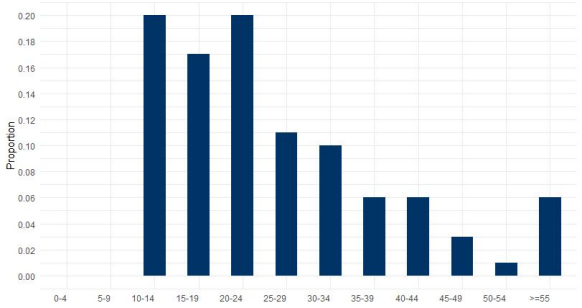

Fig. S17. Length distribution of ACPs and non-ACPs on ACPred-FL dataset.

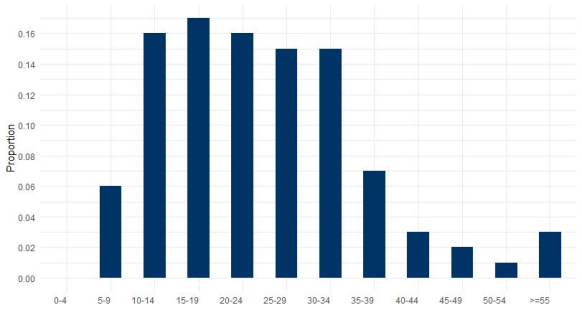

Fig. S18. Length distribution of ACPs and non-ACPs on ACPred-Fuse dataset.

Supplemental Figure 2.4: Feature Fusion and Ablation Experiment Results

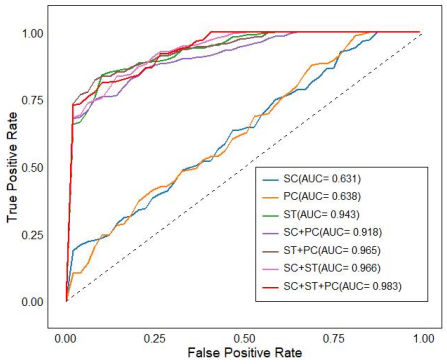

Fig. S19. ROC curve comparison for different feature fusion strategies on the ACPmain dataset. SC: sequence composition features, ST: structural features, PC: physicochemical features.

The ROC curve comparison for the different feature combinations is displayed, further illustrating the benefits of feature fusion in enhancing model performance. There are significant differences in the performance of single features. The AUC value of the SC feature is only 0.631, indicating that its ability to distinguish between ACPs and non-ACPs when used alone is quite limited. The ST feature has an AUC of 0.943, demonstrating a strong discriminatory power. This means that the model using this feature can efficiently separate ACPs from non- ACPs, and the probability-based predictions are highly accurate, with an excellent ability to identify positive and negative samples. Regarding the feature fusion strategies, the AUC of the fusion of SC and ST is 0.966, which is higher than the values when they are used separately. Indicating that feature fusion enhances the model's discriminatory ability, enabling it to more accurately predict the probability of a sample being an ACP. The AUC of the fusion of ST and PC is 0.965. It also improves the model's performance in

probability-based classification, suggesting that the complementary information between these two features helps the model better capture the characteristics of ACPs and non-ACPs. When all three features of SC, ST, and PC are fused, the AUC reaches the highest value of 0.983. This fully demonstrates that comprehensively utilizing these three features can achieve the most accurate probability-based classification. The model's ability to distinguish between ACPs and non-ACPs is extremely excellent, and the prediction results are highly reliable.

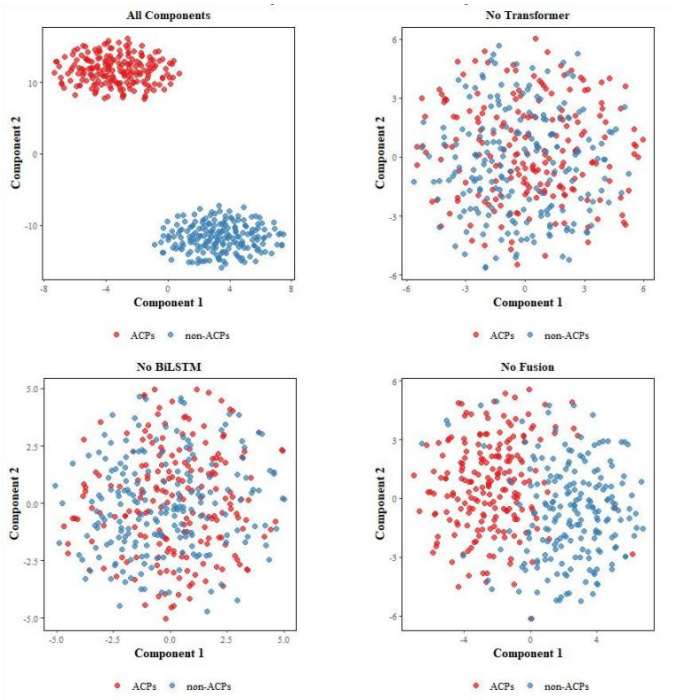

Fig. S20. t-SNE dimensionality reduction was applied to visualize after removing Transformer layers, Bi-LSTM modules and feature fusion mechanisms, and all components respectively.

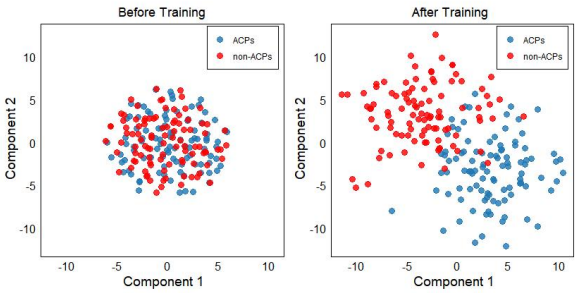

Fig. S21. t-SNE dimensionality reduction was applied to visualize the distribution of samples in the two-dimensional space before and after classification.

To gain deeper insight into the distribution of feature information in ACP-TransLSTM, we employ t-distributed stochastic neighbor embedding (t-SNE) for feature visualization. t-SNE is a dimensionality reduction technique that helps illustrate the clustering patterns of high-dimensional data in a two-dimensional space, making it easier to observe how well the model differentiates anticancer peptides from non-anticancer peptides. Before training, the sequences exhibit a highly mixed distribution, indicating that ACPs and non-ACPs are not easily distinguishable based on raw features alone. However, after training, the feature representations become more structured, with ACPs and non-ACPs forming clearly separable clusters. This demonstrates that the model effectively learns meaningful feature representations through the fusion of amino acid sequence composition, structural and physicochemical features.

3. Supplemental Tables

Supplemental Table 3.1: Classification results of different amino acids

Table S1. Classification of amino acids into five groups based on physicochemical properties

Physicochemical property	Amino acid
Aliphatic group(G1)	G, A, V, L, M, I
Aromatic group(G2)	F, Y, W
Positive charge group (G3)	K, R, H
Negative charge group(G4)	D, E
Uncharged group (G5)	S, T, C, P, N, Q

Table S2. Classification of amino acids into seven groups based on structural characteristics.

Group	Amino acid
G1	A, G, V
G2	I, L, F, P
G3	Y, M, T, S
G4	H, N, Q, W
G5	R, K
G6	D, E
G7	C

Supplemental Table 3.2: Experimental Results of different filters

Table S3. Performance comparison of different filters of CNN on ACP530 dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACP530	8	0.915	0.926	0.905	0.916	0.834	0.979
	16	0.925	0.925	0.921	0.925	0.853	0.98
	32	0.91	0.902	0.918	0.908	0.828	0.981
	64	0.924	**0.936**	0.911	0.925	0.851	**0.984**
	128	0.834	0.785	0.883	0.816	0.7	0.963
	256	**0.931**	0.933	**0.929**	**0.931**	**0.863**	**0.984**

Table S4. Performance comparison of different filters of CNN on ACP240 dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACP240	8	0.912	0.903	0.921	0.911	0.824	0.972
	16	0.915	0.939	0.89	0.917	0.831	0.975
	32	0.928	0.907	0.949	0.927	0.857	0.982
	64	0.924	0.936	0.911	0.925	0.851	**0.984**
	128	0.929	**0.946**	0.912	0.931	0.86	0.981
	256	**0.931**	0.911	**0.951**	**0.932**	**0.863**	0.982

Table S5. Performance comparison of different filters of CNN on ACP740 dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACP740	8	0.919	0.947	0.891	0.921	0.84	0.977
	16	0.918	0.9	0.934	0.917	0.837	0.975
	32	0.92	0.889	**0.952**	0.918	0.843	0.979
	64	0.921	**0.954**	0.888	0.924	0.845	0.979
	128	0.909	0.959	0.858	0.913	0.822	0.975
	256	**0.927**	0.921	0.932	**0.926**	**0.854**	**0.98**

Table S6. Performance comparison of different filters of CNN on ACPmain dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACPmain	8	0.913	0.942	0.884	0.916	0.828	0.974
	16	0.918	0.915	0.922	0.918	0.837	0.976
	32	0.919	0.95	0.888	0.922	0.84	0.978
	64	**0.923**	0.923	**0.923**	0.923	0.847	0.978
	128	0.918	0.939	0.897	0.92	0.838	0.976
	256	**0.923**	**0.972**	0.874	**0.927**	**0.851**	**0.984**

Table S7. Performance comparison of different filters of CNN on ACPred-Fuse dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACPred-Fuse	8	0.914	0.916	0.912	0.915	0.83	0.974
	16	0.923	0.887	**0.959**	0.92	0.849	0.981
	32	**0.925**	0.914	0.935	0.924	**0.851**	0.98
	64	0.922	0.911	0.933	0.92	0.84	0.978
	128	0.92	0.9	0.933	0.918	0.84	0.976
	256	0.924	**0.959**	0.889	**0.926**	**0.851**	**0.983**

Table S8. Performance comparison of different filters of CNN on ACPred-FL dataset.

Dataset	Filter	ACC	SE	SP	F1	MCC	AUC
ACPred-FL	8	0.906	0.887	0.924	0.9	0.813	0.969
	16	0.913	0.824	0.902	0.914	0.827	0.971
	32	**0.928**	0.9	**0.954**	**0.926**	**0.857**	**0.982**
	64	0.911	**0.956**	0.866	0.915	0.826	0.977
	128	0.922	0.91	0.934	0.92	0.844	0.978
	256	0.918	0.953	0.883	0.921	0.839	0.979

Table S9. Comparison of the average performance of different filters of CNN on six datasets.

Filter	ACC	SE	SP	F1	MCC	AUC
8	0.913	0.941	0.883	0.917	0.827	0.973
16	0.917	0.914	0.921	0.919	0.837	0.977
32	0.919	0.949	0.887	0.923	0.844	0.979
64	0.922	0.923	**0.922**	0.91	0.843	0.978
128	0.918	0.949	0.896	0.925	0.839	0.975
256	**0.923**	**0.971**	0.875	**0.927**	**0.851**	**0.984**

Supplemental Table 3.3: Experimental Results of different cells

Table S10. Performance comparison of different numbers of hidden cells on ACP530 dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACP530	8	0.892	0.84	**0.945**	0.883	0.796	0.976
	16	0.902	0.885	0.909	0.902	0.815	0.98
	32	0.923	0.95	0.895	0.923	0.838	0.982
	64	**0.926**	0.94	0.915	**0.925**	**0.851**	**0.983**
	128	0.904	0.91	0.826	0.911	0.817	0.972
	256	0.912	**0.952**	0.883	0.906	0.83	0.978

Table S11. Performance comparison of different numbers of hidden cells on ACP240 dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACP240	8	0.906	0.915	0.896	0.901	0.862	0.942
	16	0.91	**0.919**	0.902	0.911	0.871	0.944
	32	0.901	0.897	**0.925**	0.906	0.873	0.94
	64	**0.921**	0.912	**0.925**	**0.924**	**0.886**	**0.956**
	128	0.911	0.909	0.914	0.913	0.873	0.945
	256	0.913	0.908	0.919	0.913	0.876	0.946

Table S12. Performance comparison of different numbers of hidden cells on ACP740 dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACP740	8	0.9	0.911	0.879	0.883	0.85	0.94
	16	0.902	0.895	**0.928**	0.911	0.872	0.945
	32	0.909	**0.923**	0.895	0.909	0.868	0.948
	64	**0.916**	0.918	0.904	**0.913**	**0.883**	**0.953**
	128	0.901	0.94	0.868	0.902	0.855	0.939
	256	0.907	0.92	0.875	0.908	0.866	0.947

Table S13. Performance comparison of different numbers of hidden cells on ACPmain dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACPmain	8	0.903	0.901	0.906	0.904	0.857	0.941
	16	0.914	0.92	0.908	**0.915**	0.879	0.944
	32	0.902	0.879	0.926	0.902	0.856	0.932
	64	**0.916**	**0.928**	0.889	0.914	**0.88**	**0.945**
	128	0.905	0.871	**0.929**	0.905	0.867	0.94
	256	0.906	0.927	0.873	0.908	0.876	0.941

Table S14. Performance comparison of different numbers of hidden cells on ACPred-Fuse dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACPred-Fuse	8	0.89	0.852	**0.926**	0.887	0.832	0.93
	16	0.898	0.863	0.913	0.89	0.85	0.936
	32	0.912	0.915	0.909	0.912	0.872	0.942
	64	**0.916**	0.906	0.916	**0.915**	**0.893**	**0.946**
	128	0.915	**0.922**	0.905	**0.915**	0.883	0.945
	256	0.912	0.92	0.906	0.912	0.875	0.943

Table S15. Performance comparison of different numbers of hidden cells on ACPred-FL dataset.

Dataset	Cell	ACC	SE	SP	F1	MCC	AUC
ACPred-FL	8	0.892	0.897	0.921	0.904	0.846	0.938
	16	0.913	0.934	0.89	**0.913**	**0.887**	0.943
	32	0.91	0.9	0.92	0.91	0.871	0.941
	64	0.916	0.89	0.921	0.905	0.872	0.943
	128	**0.918**	0.901	**0.927**	**0.913**	0.878	**0.945**
	256	0.905	**0.92**	0.881	0.906	0.862	0.94

Table S16. Comparison of the average performance of different numbers of hidden cells on six datasets.

Cell	ACC	SE	SP	F1	MCC	AUC
8	0.904	0.903	0.916	0.904	0.857	0.939
16	0.915	0.921	0.908	**0.915**	0.879	0.942
32	0.903	0.889	0.924	0.902	0.856	0.922
64	**0.916**	**0.928**	0.887	0.914	**0.88**	**0.945**
128	0.907	0.872	**0.929**	0.905	0.867	0.938
256	0.908	0.926	0.873	0.908	0.876	0.939

Supplemental Table 3.4: Experimental Results of different critical probabilities

Table S17. Performance comparison of different critical probabilities on ACP530 dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACP530	0.3	0.891	0.857	0.93	0.917	0.796	0.967
	0.4	0.902	0.927	0.79	0.91	0.828	0.97
	0.5	**0.925**	**0.93**	0.92	**0.926**	**0.854**	**0.973**
	0.6	0.922	0.903	0.918	0.92	0.848	0.969
	0.7	0.916	0.875	**0.96**	0.913	0.841	0.958

Table S18. Performance comparison of different critical probabilities on ACP240 dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACP240	0.3	0.912	0.925	0.798	0.868	0.771	0.923
	0.4	0.87	0.906	0.853	0.85	0.746	0.889
	0.5	**0.932**	**0.954**	0.896	**0.894**	**0.865**	**0.972**
	0.6	0.928	0.906	0.951	0.89	0.858	0.969
	0.7	0.909	0.841	**0.968**	0.853	0.826	0.943

Table S19. Performance comparison of different critical probabilities on ACP740 dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACP740	0.3	0.881	0.892	0.87	0.873	0.762	0.91
	0.4	0.88	0.884	0.875	0.878	0.76	0.912
	0.5	0.921	**0.943**	0.907	0.898	0.846	0.967
	0.6	**0.923**	0.879	**0.959**	**0.9**	**0.851**	**0.97**
	0.7	0.912	0.863	0.941	0.881	0.829	0.961

Table S20. Performance comparison of different critical probabilities on ACPmain dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACPmain	0.3	0.872	0.857	0.91	0.864	0.756	0.892
	0.4	0.902	0.917	0.883	0.884	0.807	0.907
	0.5	**0.915**	0.859	**0.925**	**0.906**	**0.837**	**0.963**
	0.6	0.903	**0.929**	0.859	0.894	0.828	0.96
	0.7	0.902	0.919	0.865	0.883	0.83	0.958

Table S21. Performance comparison of different critical probabilities on ACPred-Fuse dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACPred-Fuse	0.3	0.894	0.867	0.923	**0.894**	0.816	0.916
	0.4	0.883	**0.936**	0.786	0.876	0.803	0.903
	0.5	**0.911**	0.904	0.909	0.884	**0.834**	**0.958**
	0.6	0.908	0.882	0.92	0.862	0.824	0.956
	0.7	0.909	0.852	**0.941**	0.88	0.826	0.955

Table S22. Performance comparison of different critical probabilities on ACPred-FL dataset.

Dataset	Pro	ACC	SE	SP	F1	MCC	AUC
ACPred-FL	0.3	0.892	0.859	0.912	0.914	0.83	0.911
	0.4	0.92	**0.927**	0.786	**0.931**	0.853	0.946
	0.5	**0.925**	0.913	**0.931**	0.93	**0.859**	**0.951**
	0.6	0.924	0.923	0.918	0.923	0.86	0.95
	0.7	0.916	0.896	0.925	0.913	0.846	0.943

Supplemental Table 3.5: Experimental Results of different component combinations

Table S23. Performance comparison of different component combinations on ACP240 dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACP240	-Ex Transformer layers	0.85	0.73	0.70	0.78	0.75
	-Ex Bi-LSTM modules	0.87	0.84	0.74	0.83	0.81
	-Ex feature fusion mechanisms	0.79	0.74	0.68	0.80	0.76

Table S24. Performance comparison of different component combinations on ACP740 dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACP740	-Ex Transformer layers	0.81	0.76	0.69	0.83	0.75
	-Ex Bi-LSTM modules	0.84	0.86	0.74	0.77	0.85
	-Ex feature fusion mechanisms	0.75	0.73	0.66	0.73	0.79

Table S25. Performance comparison of different component combinations on ACP530 dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACP530	-Ex Transformer layers	0.71	0.76	0.52	0.73	0.70
	-Ex Bi-LSTM modules	0.65	0.70	0.74	0.77	0.85
	-Ex feature fusion mechanisms	0.68	0.77	0.66	0.80	0.58

Table S26. Performance comparison of different component combinations on ACPmain dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACPmain	-Ex Transformer layers	0.73	0.75	0.53	0.76	0.68
	-Ex Bi-LSTM modules	0.66	0.69	0.49	0.67	0.68
	-Ex feature fusion mechanisms	0.64	0.67	0.47	0.71	0.58

Table S27. Performance comparison of different component combinations on ACPred-Fuse dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACPred-Fuse	-Ex Transformer layers	0.72	0.73	0.26	0.65	0.77
	-Ex Bi-LSTM modules	0.68	0.69	0.19	0.68	0.69
	-Ex feature fusion mechanisms	0.62	0.61	0.24	0.58	0.63

Table S28. Performance comparison of different component combinations on ACPred-FL dataset.

Dataset	Component	ACC	AUC	MCC	SE	SP
ACPred-FL	-Ex Transformer layers	0.82	0.87	0.75	0.80	0.82
	-Ex Bi-LSTM modules	0.78	0.84	0.71	0.76	0.74
	-Ex feature fusion mechanisms	0.72	077	0.65	0.70	0.72

Table S29. Comparison of the average performance of different component combinations

Component	ACC	AUC	MCC	SE	SP
-Ex Transformer layers	0.753	0.736	0.575	0.758	0.745
-Ex Bi-LSTM modules	0.768	0.762	0.602	0.747	0.77
-Ex feature fusion mechanisms	0.716	0.715	0.56	0.72	0.677
All Components	**0.912**	**0.93**	**0.727**	**0.882**	**0.897**

To further understand the contribution of each component in the ACP-TransLSTM model, we conducted ablation studies. We systematically removed or replaced the core architecture components, including the Transformer layers, Bi-LSTM modules and feature fusion mechanisms. When the Transformer layers were removed from the ACP-TransLSTM model, the model's performance dropped significantly. Transformer layers play a crucial role in capturing global dependencies and semantic information in peptide sequences, which is essential for accurate ACP prediction. After removing the Bi-LSTM modules, the model also showed a decline in performance. On the ACP740 dataset, the ACC decreased from 0.94 to 0.84 and the MCC decreased from 0.86 to 0.74. The Bi-LSTM modules are effective in handling long-range dependencies in sequential data. When the feature fusion mechanisms were removed and the model only used a single type of feature for prediction, the performance was

severely degraded. On the ACPmain dataset, the ACC decreased from 0.87 to 0.64 and the MCC decreased from 0.67 to 0.47. This highlights the importance of the feature fusion mechanisms in integrating different types of information to improve the model's performance. A comprehensive analysis of these results offers valuable insights into the significance of each component in the model. When all components are integrated in the ACP-TransLSTM model, it achieves an impressive performance with an ACC of 0.912, an AUC of 0.93, an MCC of 0.727, an SE of 0.882 and an SP of 0.897. This indicates that the combination of all elements works to effectively capture the complex characteristics of anticancer peptides. In contrast, when the Transformer layers are removed, the model's performance deteriorates significantly. The ACC drops to 0.753 and the AUC to 0.736.

Supplemental Table 3.6: Experimental Results of different methods

Table S30. Comparison of eight methods on ACP240 dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACP240	ACP-DA	0.80	0.85	0.65	0.78	0.76
	ACP-DL	0.70	0.82	0.60	0.75	0.72
	ACP- MHCNN	0.85	0.88	0.72	0.82	0.79
	iACP	0.78	0.86	0.68	0.80	0.77
	CL-ACP	0.88	0.90	0.75	0.86	0.85
	ACP-check	0.90	0.92	0.78	0.88	0.87
	TriNet	0.87	0.89	0.75	0.89	0.81
	ACPScanner	0.88	0.91	0.74	0.85	0.85
	ACP-TransLSTM	**0.92**	**0.93**	**0.80**	**0.90**	**0.89**

Table S31. Comparison of eight methods on ACP740 dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACP740	ACP-DA	0.81	0.74	0.58	0.80	0.82
	ACP-DL	0.81	0.89	0.62	0.81	0.80
	ACP- MHCNN	0.86	0.90	0.72	0.89	0.83
	iACP	0.81	0.86	0.61	0.87	0.74
	CL-ACP	0.84	0.91	0.68	0.83	0.85
	ACP-check	0.87	0.92	0.75	0.86	0.88
	TriNet	0.90	0.89	0.82	0.87	0.89
	ACPScanner	0.91	0.96	0.80	0.89	0.90
	ACP-TransLSTM	**0.94**	**0.98**	**0.86**	**0.93**	**0.92**

Table S32. Comparison of eight methods on ACP530 dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACP530	ACP-DA	0.85	0.84	0.73	0.76	0.78
	ACP-DL	0.79	0.78	0.69	0.74	0.80
	ACP- MHCNN	0.73	0.71	0.60	0.69	0.71
	iACP	0.82	0.82	0.75	0.78	0.82
	CL-ACP	0.55	0.39	0.43	0.77	0.39
	ACP-check	**0.93**	**0.96**	**0.85**	0.80	**0.96**
	TriNet	0.91	0.90	0.83	0.85	0.82
	ACPScanner	0.90	0.92	0.84	0.86	0.94
	ACP-TransLSTM	0.91	0.95	0.83	**0.88**	0.95

Table S33. Comparison of eight methods on ACPmain dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACPmain	ACP-DA	0.75	0.75	0.53	0.77	0.73
	ACP-DL	0.53	0.46	0.09	0.86	0.21
	ACP- MHCNN	0.73	0.71	0.46	0.79	0.67
	iACP	0.55	0.47	0.11	0.78	0.32
	CL-ACP	0.45	0.39	0.12	0.67	0.23
	ACP-check	0.78	0.85	0.56	0.80	0.77
	TriNet	0.80	0.79	0.62	0.77	0.79
	ACPScanner	0.79	0.77	0.58	0.76	0.78
	ACP-TransLSTM	**0.87**	**0.89**	**0.67**	**0.89**	**0.80**

Table S34. Comparison of eight methods on ACPred-Fuse dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACPred -Fuse	ACP-DA	0.86	0.85	0.29	0.68	0.89
	ACP-DL	0.82	0.83	0.22	0.68	0.83
	ACP- MHCNN	0.88	0.86	0.29	0.70	0.88
	iACP	0.88	0.76	0.23	0.55	0.89
	CL-ACP	0.85	0.85	0.29	0.70	0.86
	ACP-check	**0.91**	**0.90**	**0.37**	0.73	**0.92**
	TriNet	0.88	0.89	0.36	0.80	0.87
	ACPScanner	0.87	0.88	0.34	**0.81**	0.84
	ACP-TransLSTM	0.88	0.86	0.35	0.79	0.90

Table S35. Comparison of eight methods on ACPred-FL dataset.

Dataset	Model	ACC	AUC	MCC	SE	SP
ACPred-FL	ACP-DA	0.77	0.83	0.56	0.68	0.87
	ACP-DL	0.79	0.84	0.59	0.74	0.84
	ACP- MHCNN	0.77	0.82	0.55	0.67	0.87
	iACP	0.77	0.80	0.49	0.68	0.80
	CL-ACP	0.88	0.94	0.78	0.81	0.94
	ACP-check	0.91	0.94	0.82	0.87	**0.95**
	TriNet	0.88	0.90	0.77	0.86	0.85
	ACPScanner	0.89	0.93	0.76	0.89	0.88
	ACP-TransLSTM	**0.92**	**0.97**	**0.85**	**0.90**	0.92