Research Article

ENsiRNA: A Multimodality Method for siRNA-mRNA and Modified siRNA Efficacy Prediction Based on Geometric Graph Neural Network

Wenchong Tan [1], Mingshu Dai [1], Shimin Ye [1], Xin Tang [1], Dawei Jiang [1], Dong Chen [2], and Hongli Du [1,*]

*1 - **School of Biology and Biological Engineering,** South China University of Technology, Guangzhou, Guangdong Province 510000, China*

*2 - **Fangrui Institute of Innovative Drugs,** South China University of Technology, Guangzhou, China*

***Correspondence to Hongli Du:** hldu@scut.edu.cn (H. Du)*

https://doi.org/10.1016/j.jmb.2025.169131

Edited by Zhi John Lu

Abstract

With the rise of small interfering RNA (siRNA) as a therapeutic tool, effective siRNA design is crucial. Current methods often emphasize sequence-related features, overlooking structural information. To address this, we introduce ENsiRNA, a multimodal approach utilizing a geometric graph neural network to predict the efficacy of both standard and modified siRNA. ENsiRNA integrates sequence features from a pretrained RNA language model, structural characteristics, and thermodynamic data or chemical modifications to enhance prediction accuracy. Our results indicate that ENsiRNA outperforms existing methods, achieving over a 13% improvement in Pearson Correlation Coefficient (PCC) for standard siRNA across various tests. For modified siRNA, despite challenges associated with RNA folding methods, ENsiRNA still demonstrates competitive performance in different datasets. This novel method highlights the significance of structural information and multimodal strategies in siRNA prediction, advancing the field of therapeutic design.

Introduction

Small interfering RNA (siRNA) has the potential to treat a broad range of diseases. The mechanism, RNA interference (RNAi), is a critical and spontaneous cellular pathway for posttranscriptional gene regulation in numerous eukaryotic organisms.[1] The initial step of RNAi entails the processing and cleavage of longer double-stranded RNA into siRNAs, a process carried out by an RNase III-like enzyme known as Dicer.[2] After formation, the antisense strand of the siRNA duplex is then incorporated by a multiprotein complex called RISC, an RNA induced silencing complex, which is capable of cleaving target mRNAs.[3]

siRNA therapies face challenges in targeted delivery, endosomal escape, and sustained efficacy, necessitating improved delivery methods, endosomal escape mechanisms, and in vivo stability for effective outcomes.[4] These challenges lead to two critical steps in the pipeline of designing siRNAs. Initially, siRNA sequences are designed using sophisticated algorithms to maximize the efficacy on target mRNA and minimize the risk of off-target effects.[5] Subsequently, these sequences undergo extensive chemical modifications to enhance their bioavailability, safety, and stability within the biological environment.[6] Based on the modification sites in nucleic acids, these structures can be categorized into three classes: phosphonate modifications, ribose modifications, and base modifications.[7] For example, ribose modification 2′-O

methylated has been shown to increase both the binding affinity and nuclease resistance of siRNA. Phosphonate modification phosphorothioate confers resistance to nucleases and enhance hydrophobic protein-binding properties. In contrast, base modifications or replacements are not widely used for siRNA drugs yet.[8–10] All of these modifications, including the difference in the modified number and position, result in different efficacy and cytotoxicity, presenting challenges for therapeutic siRNA design.[11,12] Five siRNA therapeutics are approved for liver diseases, with many candidates in trials, highlighting siRNA's therapeutic potential.[13]

Currently, numerous deep learning and machine learning approaches have been developed for predicting siRNA-mRNA efficacy as well as the efficacy of modified siRNAs. With the accumulation of experimental data, machine learning methods such as Biopredsi, DISR, and i-Score have enabled predictions of siRNA activity.[14,15,5] Convolutional Neural Network (CNN) and Graph Neural Network (GNN) methods, which utilize deep learning networks, have been designed in recent years.[16,17] Although GNN methods have shown better performance than CNN methods on internal datasets by generating predictions from interaction nodes between siRNA and mRNA, they face out-of-distribution (OOD) generalization issues due to the different distributions of siRNA and mRNA in new datasets. Recently, a new Transformer based method showcased the advantages of combining siRNA prediction with RNA pretrained language models.[18,19] For modified siRNA prediction, SMEpred and Dominic D. Martinelli utilized machine learning techniques such as SVM and Random Forest, encoding modified categories using one-hot encoding.[20,21] With the development of deep learning, the newest method, Cm-siRPred, constructed the model by cross-attention 3D-CNN, with the modified embedding in MACCS and RNA properties[22].

Despite these advancements, current methods do not directly utilize siRNA 3D structure, hindering the natural and effective capture of nucleotide interactions and comprehensive information. Furthermore, current deep learning methods are not suitable for geometric graph data, which includes most biological structures, such as molecules and proteins.[23] The structure of siRNA, characterized as a geometric graph, necessitates a geometric graph deep learning method for handling. EGNN, a model with E(3)-equivariance, has been widely applied to molecular, protein properties prediction and docking as well as serving as protein and molecule generative model.[24–26] While previous structural information has been utilized in the biomedical field, existing siRNA efficacy prediction methods lack a direct 3D structural representation that can effectively capture intricate nucleotide interactions through advanced deep learning techniques. Lastly, existing methods typically focus solely on siRNA-mRNA prediction or modified siRNA prediction, lacking a comprehensive approach for the siRNA prediction workflow.

To address these challenges, we propose ENsiRNA: a multimodality method for siRNA-mRNA and modified siRNA efficacy prediction based on geometric GNN. By directly utilizing the structure of double-stranded siRNA generated by ViennaRNA and FARFAR2, ENsiRNA simultaneously encodes the nucleic acid sequences and thermodynamic or modification features as node features, generating predictions from the global node connected to the nodes of siRNA or mRNA.[27,28] As the first multimodality method for siRNA design, ENsiRNA achieves optimal performance and demonstrates real-world applicability.

Material and Methods

Overview of the workflow

The overview workflow and multimodality strategy of our method are illustrated in Figure 1A, which includes siRNA-mRNA efficacy prediction method ENsiRNA and modified siRNA efficacy prediction method ENsiRNA-mod. The design pipeline of therapeutic siRNA encompasses target gene and transcript evaluation, sequence design, filtration for off-target effects and SNPs (Single Nucleotide Polymorphism) as well as modification design, Figure 1B shows where our method fits in 10,29. The detailed model architectures of ENsiRNA and ENsiRNA-mod are depicted in Figure 1C and D, respectively, sharing the multimodality strategy with sequence-based and structure-based information.

Data source and processing

The data for siRNA-mRNA prediction were first collected from the work of Rosa,[17] which originates from the experimental data of Huesken,[14] Harborth,[30] Ui-Tei,[31] Vickers,[32] Khovorova[33] and Katoh.[34] The dataset containing a single mRNA is reserved for external testing, while the rest is used for 5-fold cross-validation. To demonstrate our method's efficacy, we also collected real-world drug design data from patent US20230026968A1, related to an FDA-approved drug, targeting human ALAS1 mRNA (NM_000688.4). Additionally, we evaluate the potential cross-species application of our method in discussion, with the data from a cross-species siRNA study targeting CKIP-1, which is detailed in supplementary data.[35]

Although siRNAs may utilize overhangs to enhance the formation of RISC-duplex complexes, in real-world drug design, 19-mer siRNAs are first confirmed and then extended. These overhangs do not play a critical role in target mRNA

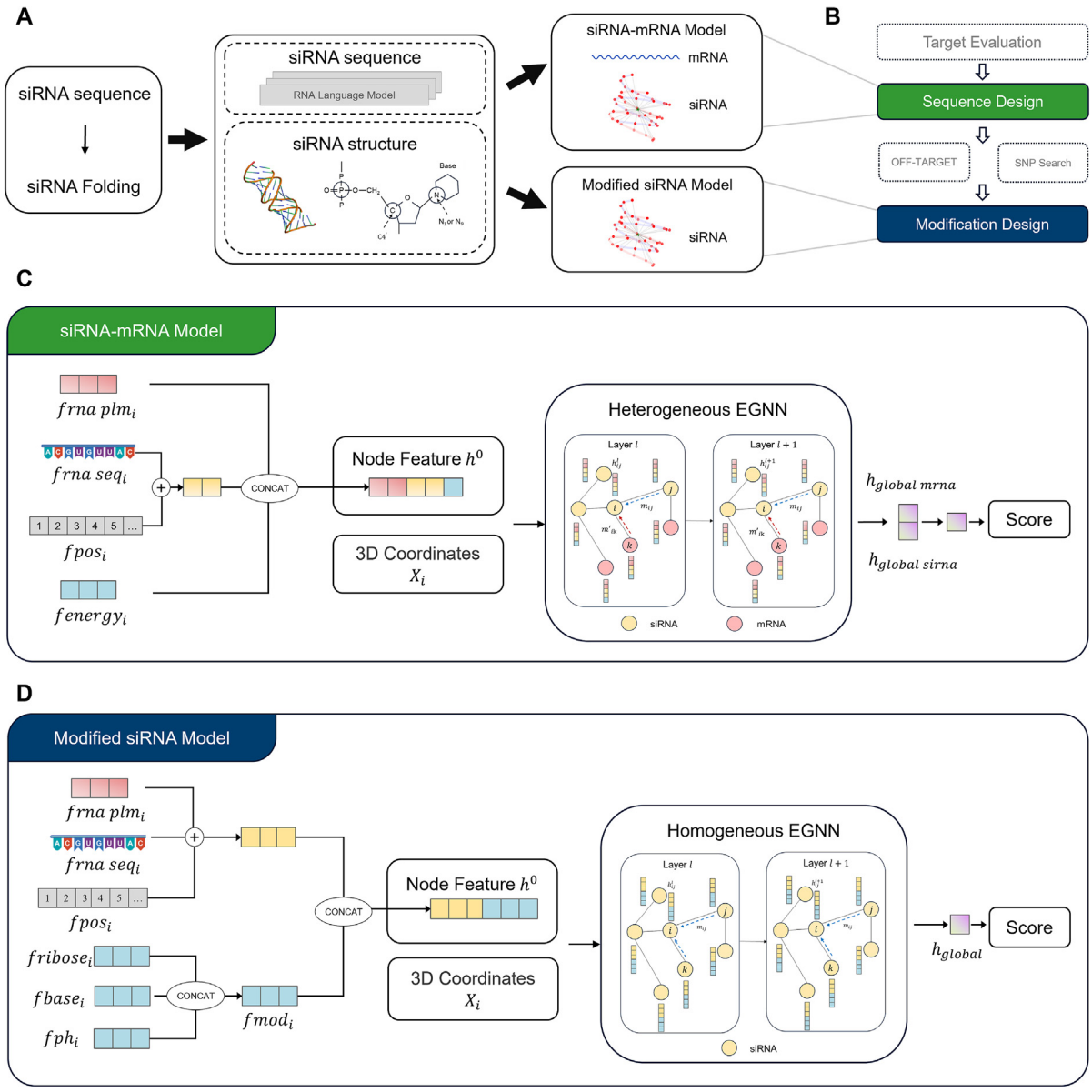

Figure 1. (**A**) Figure depicted the overview workflow of ENsiRNA. (**B**) Figure depicted the principal process of therapeutic siRNA design and where our method fits in. (**C**) Figure depicted the model architecture of ENsiRNA. (**D**) Figure depicted the model architecture of ENsiRNA-mod.

recognition.[12] As a result, all siRNAs in our method and comparisons are set to a length of 19 nucleotides. The mRNA sequences consist of the target region plus 21 nucleotides in both upstream and downstream regions, resulting in a total length of 61 nucleotides. This optimal length setting was determined based on previous studies.[16,17]

For 5-fold cross-validation, data are scaled to [0, 1] to balance experimental sources, with efficacy distribution shown in Figure S1. Data are binned into 10 efficacy score categories and then split into 5 folds. The number of siRNA and mRNA sequences is detailed in Table 1.

After constructing the dataset, double-strand siRNA sequences are transformed into 3D structures. The secondary structure is predicted using RNAplex from ViennaRNA,[28] and 3D structures are built with FARFAR2, a Rosetta-based RNA folding method using full-atom refinement.[27]

For modified siRNA prediction, we firstly collected and processed data from the siRNAmod database and relevant patents (Table 1).[36] After thorough review of SMILES chemical formulas and classification into ribose, base, and phosphonate substructures, we established a dataset of 3,512 high-quality entries (3,272 from siRNAmod, 240 from patents), focusing on sequences under 25 nucleotides with concentrations below 100 nM. We supplemented this with recent modifications from patents US20220331351A1 and

Table 1 The data source and distribution of ENsiRNA and ENsiRNA-mod.

		Num of siRNAs	Num of mRNAs	Data source
Dataset of ENsiRNA	5-fold cross-validation	2838	45	Huesken, Harborth, Ui-Tei, Vickers, Khovorova
	External test set	702	1	Katoh
	Real world drug design dataset	232	1	patent
Dataset of ENsiRNA-mod	5-fold cross-validation & Internal test set	3514	\	siRNAmod (3274), patent (240)
	siRNA drugs dataset	7	\	siRNA drugs
	Real world drug design dataset	15	\	patent

US20220090067A1. Figure S2A displays modification ratios at each siRNA position, while Figure S2B uses a t-SNE graph with RDKit-generated Morgan fingerprints to show substructure property distributions. Finally, the data were divided into efficacy-based bins for cross-validation and testing.

To evaluate our method's performance beyond training data, we assembled an external test set comprising five FDA-approved siRNA drugs and two additional clinical candidates, which is detailed in Table S1. We also tested our method on a dataset from patent US20230026968A1, similar to Givosiran but with different modification positions, aligning with real-world drug design.

Architecture of siRNA-mRNA model

To effectively address the multimodal features of siRNA, particularly in terms of structure, we have selected the EGNN as our primary model. We represent the nucleic acid structure using the RNA backbone atoms C4′, P, and glycosidic N.[37] Due to current limitations in accurately predicting long mRNA tertiary structures, which are typically represented as non-geometric data, we designed a heterogeneous EGNN network that processes both siRNA geometric data and mRNA non-geometric data.[38,39]

The overview structure of the model is presented in Figure 1B. Additionally, the overview of ENsiRNA is provided in Eq. (1), where $\mathscr{G}_{mRNA-siRNA}$ denotes the graph of mRNA and double strand of siRNA. The heterogeneous graph of $\mathscr{G}_{mRNA-siRNA}$ comprises mRNA nodes $\mathscr{V}_{mRNA}$ and siRNA nodes $\mathscr{V}_{siRNA}$, with sequence, base-pair, self-loop, and global edges. Sequence edges are based on mRNA or siRNA sequences, base-pair edges on siRNA duplex or siRNA-mRNA pairs, and global edges connect nodes to either mRNA or siRNA. The architecture of the heterogeneous EGNN will be detailed alongside the homogeneous EGNN.

$$h_i = \text{heterogeneous EGNN}(\mathscr{G}_{mRNA-siRNA}),\quad i \in \mathscr{V}_{mRNA}, \mathscr{V}_{siRNA} \tag{1}$$

Input of siRNA-mRNA model. In the siRNA-mRNA model, the input of each node in $\mathscr{V}_{siRNA}$ are associated with the representation (h_i, X_i), where h_i includes sequence and energy features, and $X_i$indicates the 3D coordinates of three backbone atoms of nucleic acid, as structure-related information. $\mathscr{V}_{mRNA}$ nodes are solely represented as h_i, without structure information.

$$h_i^0 = f(rna_i, energy_i) = \text{CONCAT}(frna_i, fenergy_i),\quad i \in \mathscr{V}_{mRNA}, \mathscr{V}_{siRNA} \tag{2}$$

$$frna_i = \text{CONCAT}((fseq_i + fpos_i), frnaplm_i),\quad i \in \mathscr{V}_{mRNA}, \mathscr{V}_{siRNA}$$

$$fenergy_i = fenergy_{raw} - fenergy_i^{mutate} \tag{3}$$

In detail, node feature h_i^0 combines $frna_i$ and $fenergy_i$, representing nucleic acid and thermal properties. Firstly, $frna_i$ is concatenated with $frnaplm_i$and the concatenation of $f\text{seq}_i$ and $f\text{pos}_i$, where $f\text{seq}_i$ and $f\text{pos}_i$ define the learnable embeddings of four types of nucleic acid and their position, while the $frnaplm_i$ represents the embedding generated by the pretrained RNA language model, RNA-FM.[40] RNA-FM, a BERT model with 12 transformer encoders, is pretrained on 23 million non-coding RNA sequences. Its embeddings provide sequence and evolutionary information, enhancing our sequence features and also structure features, especially for mRNA.[41]

In addition to leveraging the structural features, our method differs in energy calculation by computing energies for single chains (mRNA, sense, and antisense siRNA) and duplex chains (mRNA-antisense siRNA, sense-antisense siRNA) to capture siRNA silencing stages. We allocate energy contributions to each nucleic acid depicted in Eq. (3), where $fenergy\prime_i$ is calculated by mutating each nucleic acid to X in the ViennaRNA fold_compound function for single chain and duplexfold for duplex chain. The dimensions of these features are detailed in Table S2.

To enhance generalization, we use three backbone atoms to represent nucleic acid structures. Consequently, the input X_i of $\mathscr{G}_{mRNA-siRNA}$ here is the coordinates of three backbone atoms in nucleic acid, which are $X_i^{C4'}$, X_i^{P}, $X_i^{N9'}$ in A, G (4) and $X_i^{C4'}$, X_i^{P}, $X_i^{N1'}$ in C, U (5).

$$X_i = \left(X_i^{C4'}, X_i^{P}, X_i^{N9'}\right), \quad i \in A, G \tag{4}$$

$$X_i = \left(X_i^{C4'}, X_i^{P}, X_i^{N1'}\right), \quad i \in C, U \tag{5}$$

Prediction and training for siRNA-mRNA model. After inputting h_i, X_i into Heterogeneous EGNN, as detailed in Eq. (6), the output of the model includes global nodes $h_{mRNA\ global}$ and $h_{mRNA\ global}$ Once concatenated, these outputs will be connected to Multi-Layer Perceptron (MLP) $\varnothing_p$ and activated by a sigmoid function to normalize the results to the range [0, 1]. The loss is then computed using the smooth L1 loss, as described in Eq. (7), where y represents the predicted efficacy of model and $\hat{y}$ denotes the truth label.

$$y = \text{sigmod}\left(\varnothing_p\left(\text{CONCAT}\left(h_{mRNA\ \ global}, h_{siRNA\ \ global}\right)\right)\right) \tag{6}$$

$$Loss(y, \hat{y}) = \begin{cases} 0.5 \cdot (y - \hat{y})^2, & if\ \ |y - \hat{y}| < 1 \\ |y - \hat{y}| - 0.5, & if\ \ |y - \hat{y}| \geq 1 \end{cases} \tag{7}$$

Architecture of modified siRNA model

The modified siRNA model architecture resembles the siRNA-mRNA model but differs in graph construction, feature representation, and EGNN structure. The overview of ENsiRNA-mod is provided in Eq. (8), where $\mathcal{G}_{modifiedsiRNA}$ denotes the graph of modified siRNA with nodes $\mathcal{V}_{siRNA}$. The graph of $\mathcal{G}_{modifiedsiRNA}$ here only includes nodes of siRNA with edges comprising sequence edges, base-pair edges, self-loop edges and global edges. While base-pair edges here only contain the edges between sense siRNA and antisense siRNA, the only one global node connects to the nodes of siRNA.

$$h_i = \text{homogeneous EGNN}\left(\mathcal{G}_{modified\ \ siRNA}\right), \quad i \in \mathcal{V}_{siRNA} \tag{8}$$

Input of modified siRNA model. Compared with the siRNA-mRNA model, the model of modified siRNA only includes the nodes in $\mathcal{V}_{siRNA}$, which is represented as (h_i, X_i). Here, h_i denotes node feature including sequence-related and modification-related information, and $X_i$indicates the 3D coordinates of three backbone atoms of nucleic acid, as structure-related information.

The nucleic acid information and the modification information as input feature h_i^0 for each node in G_{siRNA} are represented by Eqs. (9)–(11)

$$h_i^0 = f(rna_i, mod_i) = \text{CONCAT}(frna_i + fmod_i), \quad i \in \mathcal{V}_{siRNA} \tag{9}$$

$$frna_i = fseq_i + fpos_i + frna\ \ plm_i, \quad i \in \mathcal{V}_{siRNA} \tag{10}$$

$$fmod_i = f(ribose_i, base_i, ph_i) = CONCAT(fribose_i, fbase_i, fph_i), \quad i \in \mathcal{V}_{siRNA}\# \tag{11}$$

The modified siRNA model features resemble the siRNA-mRNA model, with specific differences. Generally, input feature h_i^0 is concatenated with $frna_i$ and $fmod$. On one hand, $frna_i$ is added by $fseq_i$ (4 nucleic acid types), $fpos_i$ (sinusoidal positional encoding) and $frna\ plm_i$(RNA pretrained language model embedding). On the other hand, $fmod_i$ encompasses features from three substructures of nucleic acids, which are concatenated with $fribose_i$, $fbase_i$ and fph_i – the feature of ribose, base and phosphonate substructure. Table S3 details dimension information. Each nucleic acid substructure is represented by a 512-dimensional Morgan fingerprint (RDKit), reduced to 128 dimensions by a linear layer. The pretrained RNA language model embedding is similarly reduced from 640 to 128 dimensions.

Prediction of modified siRNA model. After inputting h_i, X_i into the EGNN model, as detailed in Eq. (12), the model output h_{global}, after being connected to MLP $\varnothing_p$, will be activated by a sigmoid function to be normalized to [0, 1] (12). Subsequently, the loss is computed using the smooth L1 loss, where y represents the predicting efficacy of model and y denotes the truth label, same as Eq. (8).

$$y = \text{sigmod}\left(\varnothing_p\left(h_{global}\right)\right) \tag{12}$$

Model structure of heterogeneous EGNN and homogeneous EGNN

To capture the three backbone atoms—P, C4′ and either N1 or N9—we utilize an improved version of dyMEAN, which could take account of different atoms by multi-channel EGNN.[25] Furthermore, to incorporate the non-geometric data of mRNA in siRNA-mRNA prediction, we designed a heterogeneous network based on multi-channel EGNN. Here are the details of the heterogeneous EGNN model for siRNA-mRNA prediction and the homogeneous EGNN for modified siRNA prediction.

In our method, the heterogeneous EGNN model consists of $l = 2$ layers. As shown in Figure 1B, the node feature of each layer will be stacked and output through a linear layer. Notably, since we only need h_i^l for prediction, the l-th layer will only update the hidden vector h_i^l, as demonstrated in the original paper.[26] As a result, the geometric messages are actually invariant.

The l-th layer of heterogeneous EGNN is shown in Eqs. (13)–(17)

$$h_i^{l+1} = \varnothing_h\left(h_i^l, \sum m_{ij}, \sum mI_{ik}\right), \quad i, j \in \mathcal{V}_{siRNA}, k \in \mathcal{V}_{mRNA} \tag{13}$$

$$h_i^{l+1} = \varnothing_h\left(h_i^l, \sum mI_{ij}\right), \quad i \in \mathcal{V}_{mRNA}, \quad j \in \mathcal{V}_{mRNA} \bigcup \mathcal{V}_{siRNA} \tag{14}$$

$$m_{ij} = \varnothing_m\left(h_i^l, h_j^l, h_{ij}^l\right), \quad i,j \in \mathscr{V}_{siRNA} \tag{15}$$

$$m\prime_{ik} = \varnothing_{m'}\left(h_i^l, h_k^l\right), \quad i \in \mathscr{V}_{siRNA} \bigcup \mathscr{V}_{mRNA}, k \in \mathscr{V}_{mRNA} \tag{16}$$

$$h_{ij}^l = \frac{T_R(X_i, X_j)}{\|T_R(X_i, X_j)\|_F + \epsilon}, \quad i,j \in \mathscr{V}_{siRNA} \tag{17}$$

Generally, h_i^{l+1} is the updated hidden vector of the l-th layer, Eq. (13) denotes the message update of the siRNA nodes, while Eq. (14) showcases the message update of the mRNA nodes. Here, the $\varnothing_m$, $\varnothing_{m'}$, $\varnothing_h$ are MLPs consisting of two layers of linear transformations with a SILU activation function. Specifically, h_i^{l+1} for the siRNA node is concatenated from h_i^l, $\sum m_{ij}$ and $\sum m\prime_{ik}$, which represent the self-loop edge attributes, the edge attributes between siRNA and siRNA, and the edge attributes between siRNA and mRNA, respectively. On the other hand, h_i^{l+1} of mRNA node is concatenated from h_i^l and $\sum m\prime_{ik}$, which represents the self-loop edge attributes, the edge attributes between mRNA and siRNA or between mRNA and mRNA. Furthermore, m_{ij} concatenates the hidden vector of h_i^l, h_j^l and h_{ij}^l, which serve as the geometric message between v_i and v_j processed by the geometric relation extractor T_R. Conversely, $m\prime_{ik}$ only concatenates the non-geometric hidden vector h_i^l and h_k^l.

While the homogeneous EGNN is similar, the differences in the l-th layer EGNN are shown in Eqs. (18)–(20).

$$h_i^{l+1} = \varnothing_h\left(h_i^l, \sum m_{ij}\right), \quad i,j \in \mathscr{V}_{siRNA} \tag{18}$$

$$m_{ij} = \varnothing_m\left(h_i^l, h_j^l, h_{ij}^l\right), \quad i,j \in \mathscr{V}_{siRNA} \tag{19}$$

$$h_{ij}^l = \frac{T_R(X_i, X_j)}{\|T_R(X_i, X_j)\|_F + \epsilon}, \quad i,j \in \mathscr{V}_{siRNA} \tag{20}$$

Here, h_i^{l+1}, which are all from siRNA nodes, are concatenated from h_i^l and $\sum m_{ij}$, where all nodes exhibit geometric relations.

As Eqs. (17) and (20), T_R computes the learnable weights based on the distance between $X_i \in \mathbb{R}^{3\times3}$ and $X_j \in \mathbb{R}^{3\times3}$, which are the coordinates of backbone atoms of RNA. T_R is detailed in Eq. (21), where w_i and w_j are two learnable weights as attribute embedding of different atoms, and A_i and A_j are two learnable attribute matrices as weights of different atoms. Furthermore, D_{ij},characterized by w_i and w_j, is shown in Eq. (22), calculating the pairwise distance between backbone atoms in v_i and v_j.

$$R_{ij} = A_i^\top \left(w_i w_j^l \odot D_{ij}\right) A_j \tag{21}$$

$$D_{ij} = \|X_i - X_j\|_2 \tag{22}$$

Finally, the output of T_R is normalized using Frobenius norm the Frobenius norm, with an added constant $\epsilon = 1$ for numerical stability.

Related method

For siRNA-mRNA prediction, we first retrained the HinSAGE-based method GNN4siRNA and the Transformer-based method AttSiOff according to their open-source code, using the same dataset as our method, including the 5-fold cross-validation split. GNN4siRNA features siRNA nodes and mRNA nodes with sequence and thermodynamic features, while AttSiOff extracts features from siRNA and local target mRNA sequences using a self-attention mechanism, incorporating both pre-trained RNA-FM embeddings and prior knowledge-based characteristics.[17,18] Both GNN4siRNA and AttSiOff were trained with the same hyperparameters in their original code, except for the length of siRNA sequences, which was uniform at 19 in AttSiOff. Notably, for GNN4siRNA, we built an external test set and a real-world drug design dataset in a different graph of the 5-fold cross-validation dataset, which better maps to actual applications. In addition to GNN4siRNA and AttSiOff, we tested the external test set and the real-world drug design dataset using machine learning-based methods s-BioPredsi, i-Score, and DSIR, utilizing the i-Score Designer website due to the unavailability of the source code.[5]

For modified siRNA prediction, due to the unavailability of the source code for the deep learning-based method Cm-siRPred, we recreated the cross-attention CNN method based on the details provided in the existing literature. Additionally, to better showcase the advantages of feature extraction with accurate secondary structures in our method, we also implemented machine learning methods using the flattened features of Cm-siRPred, including Random Forest based on scikit-learn and XGBoost based on xgboost.[42,43] All of these methods were used for comparison of different feature extraction techniques against our method.

Evaluation metrics

To evaluate several methods, we choose Pearson correlation coefficient (PCC) and Spearman rank correlation coefficient (SPCC) as metrics to assess the correlation between the actual and model predicted efficacies. On the other hand, Mean Absolute Error (MAE) and Mean Squared Error (MSE) are metrics used to evaluate the discrepancy or error between the actual and model predicted values. The formulas of each metrics are as follows Eqs. (23)–(26):

$$PCC(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma_x}\right)\left(\frac{Y_i - \bar{Y}}{\sigma_Y}\right) \tag{23}$$

$$SPCC(X,Y) = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \quad (24)$$

$$MAE = \frac{\sum |y_i - \hat{y}|}{n} \quad (25)$$

$$MSE = \frac{\sum (y_i - \hat{y})^2}{n} \quad (26)$$

Here, X, Y represents the predicted values and labels and σ_x, σ_Y are standard deviation on each data, d represents the difference between the X rank and Y rank for each pair of data. In MAE and MSE, y_i denotes predicted efficacy while $\hat{y}$ denotes actual efficacy. In addition to regression metrics, we calculate the Area Under the Receiver Operating Characteristic Curve (AUC) for further evaluation. The Brier score, which is analogous to MSE but is specifically tailored for binary classification with predictions ranging from 0 to 1, is also computed for siRNA drugs dataset evaluation.

Results

ENsiRNA

To demonstrate the performance of our multimodality method, we conducted experiments on siRNA-mRNA in four key areas: (1) Evaluation using 5-fold cross-validation; (2) Evaluation on an external siRNA dataset; (3) Evaluation on a real-world drug design dataset; and (4) Ablation experiments.

Performance on 5-fold cross-validation. We first assessed the learning efficacy of ENsiRNA compared to GNN4siRNA and AttSiOff through a 5-fold cross-validation process. The results are illustrated in Figure 2A and detailed in Table S4, indicating that ENsiRNA outperformed the other two deep learning methods. Specifically, ENsiRNA achieved mean PCC, SPCC, MAE, MSE and AUC scores of 0.7274, 0.7262, 0.1413, 0.0313 and 0.8626, respectively. This reflects improvements of 13.60% in PCC, 12.87% in SPCC and 6.36% in AUC, along with reductions of 10.11% in MAE and 20.76% in MSE compared to the next best method GNN4siRNA. The boxplots in Figure 2A further illustrate that ENsiRNA exhibits better performance across folds, suggesting better generalization performance. Notably, although GNN4siRNA (the HinSAGE method) slightly outperformed AttSiOff (the CNN method) when applied to 5-fold cross-validation without OOD problem, it still lagged behind our method.

Performance on external test set. To further evaluate the generalizability of ENsiRNA, we compared our method against two deep learning methods (GNN4siRNA and AttSiOff) and three machine learning-based methods (s-BioPredsi, i-Score, and DSIR). The results from the external test set are shown in Figure 2B and C, confirming that ENsiRNA not only fits well with internal data but also performs effectively on external data, demonstrating its robust generalization capabilities. ENsiRNA achieved PCC, SPCC, MAE, MSE and AUC scores of 0.6587, 0.6587, 0.1276, 0.0256 and 0.8324, which represents an improvement of 13.02% for PCC, 13.22% for SPCC and 5.74% in AUC compared to the next best method DSIR, along with a reduction of 14.25% in MAE and 27.68% in MSE compared to the next best method s-BioPredsi. Notably, s-BioPredsi and DSIR failed to predict two siRNAs on the i-Score Designer website, as indicated in Figure 2B and C, which does not include these calculations.

The performance results of the three deep learning methods on the complete dataset are presented in Figure S3 with error bars and statistical significance tests.

Evaluation on real-world drug design dataset. To assess the practical application value in drug design, we conducted an experiment predicting real-world drug design dataset from the patent of FDA-approved drug Givosiran. Givosiran is an FDA-approved siRNA drug targeting ALAS1. In the development of Givosiran, several sequences and modifications were tested in vitro and in vivo to confirm the best siRNA sequence and their modification. For the siRNA-mRNA prediction experiment, we selected data from the process of selecting siRNA sequences. The results are depicted in Figure 2D and E, where ENsiRNA achieved PCC, SPCC, MAE, MSE and AUC scores of 0.7032, 0.7034, 0.1436, 0.0364 and 0.8604, representing improvements of 13.16% in PCC, 13.86% in SPCC and 6.98% in AUC compared to the next best method DSIR, along with reductions of 8.88% in MAE compared to the next best method s-BioPredsi and 19.29% in MSE compared to the next best method i-Score.

Similarly, both s-BioPredsi and DSIR were unable to predict one siRNA on the i-Score designer website, the comparison of deep learning methods for the complete dataset, including the one siRNA, is illustrated in Figure S4 with error bars and statistical significance tests. Although the deep learning methods GNN4siRNA and AttSiOff can adapt to different input positions compared to machine learning methods, they lack sufficient generalization when applied to external datasets, which raises concerns about their practical use. Conversely, these results indicate that ENsiRNA has substantial potential for practical applications, further supporting its impressive generalization capabilities. It is important to note that due to the ODD problem, the Transformer based method AttSiOff outperformed the HinSAGE method GNN4siRNA on the external test set and real-

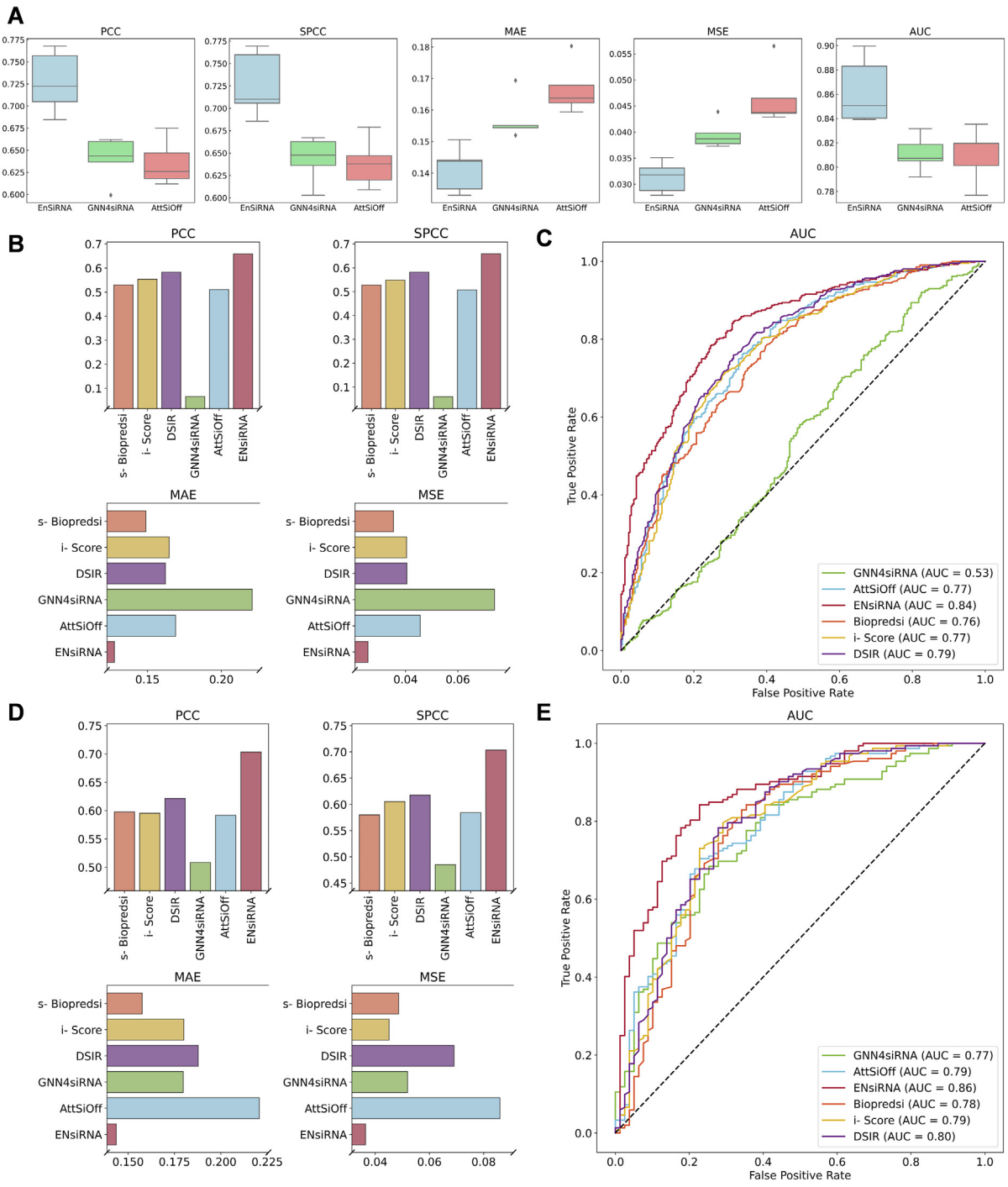

Figure 2. (**A**) The performance comparison on 5-fold cross-validation of mRNA-siRNA model. (**B**) The performance comparison of mRNA-siRNA model on external test dataset (700). (**C**) The Area Under the ROC Curve of mRNA-siRNA models on external test dataset (700). (**D**) The performance comparison of mRNA-siRNA model on real-world drug design dataset (231). (**E**) The Area Under the ROC Curve of mRNA-siRNA models on real-world drug design dataset (231).

world drug design dataset, which had a different distribution from the 5-fold cross validation.

Ablation experiment of ENsiRNA. In this section, we demonstrate the importance of specific components in our method through an ablation study based on 5-fold cross-validation. As mentioned above, the structure of siRNA represents typical geometric data based on secondary and tertiary structures playing an

important role in siRNA efficacy. To demonstrate the improvement gained by using geometric GNN method to extract structural information in coordinates, we designed an ablation experiment to train a model where m_{ij} updated without h_{ij}^{l}. Furthermore, due to the limited training data, we enhanced GNNs by integrating features from a pretrained RNA language model with node features in our method.

To visually represent the performance of the baseline model and models with ablation of specific components across multiple evaluation metrics, we use radar charts in Figure 3A and detailed in Table S5. Results from Figure 3A indicate that when the structure feature was removed (orange line), the model performance decreased in terms of PCC, SPCC, and AUC. Specifically, the removal of structure resulted in a 6.97% decrease in PCC, 6.68% decrease in SPCC, and 2.68% decrease in AUC. Similarly, excluding the RNA PLM (green line) resulted in a 7.48% decrease in PCC, 7.37% decrease in SPCC, and 2.64% decrease in AUC. These results underscore the significance of structural and RNA PLM features in improving the performance of our model.

Furthermore, we conducted an ablation study focusing on the energy features of single and duplex chains to demonstrate the significance of thermodynamic parameters at various stages of siRNA silencing. Initially, we removed $fenergy_i$, which encompasses both single-chain and duplex-chain energy features. Subsequently, we removed the duplex-chain energy feature and the single-chain energy feature individually. As illustrated in radar charts Figure 3B and detailed in Table S5, the model's performance declined the most when both energy features were absent (red line), resulting in decreases of 7.09% in PCC and 6.64% in SPCC. While the model without either single-chain (green line) or duplex-chain energy (orange line) experienced similar but smaller decreases—with PCC reductions of 6.67% and 6.93% and SPCC reductions of 6.17% and 6.71%—the absence of single-chain energy had a greater effect. These results indicate that optimal model performance is achieved only when both single-chain and duplex-chain energy features are present.

ENsiRNA-mod

Secondly, we conducted four parts of experiments on modified siRNA prediction: (1) Evaluation on 5-fold cross-validation and internal test set; (2) Evaluation on external siRNA drug datasets; (3) Evaluation on real-world drug design datasets; and (4) Ablation experiments.

Performance on 5-fold cross-validation and internal test set. We demonstrated the efficacy of ENsiRNA-mod compared to similar methods through a 5-fold cross-validation. Our reconstructed version of Cm-siRPred yielded comparable outcomes in the 5-fold cross-validation of the original paper, which we call CNN in the following text. Simultaneously, we flattened the features of Cm-siRPred and applied them to Random Forest and XGBoost using scikit-learn to further evaluate features without considering secondary structure. We proceeded to evaluate our method against identical dataset partitions.

The 5-fold cross-validation results are shown in Figure 4A and detailed in Table S6, indicating that ENsiRNA-mod outperformed the CNN method. Specifically, ENsiRNA-mod achieved mean PCC, SPCC, MAE, MSE and AUC scores of 0.8405, 0.8414, 0.0965, 0.0215, and 0.9188, respectively. This represents an improvement of 2.15% in PCC and 1.86% in SPCC, along with reductions of 9.90% in MAE and 10.04% in MSE compared to the CNN approach. As shown in the boxplots in Figure 4A, while performance distributions for ENsiRNA-mod and CNN exhibit some overlap, our

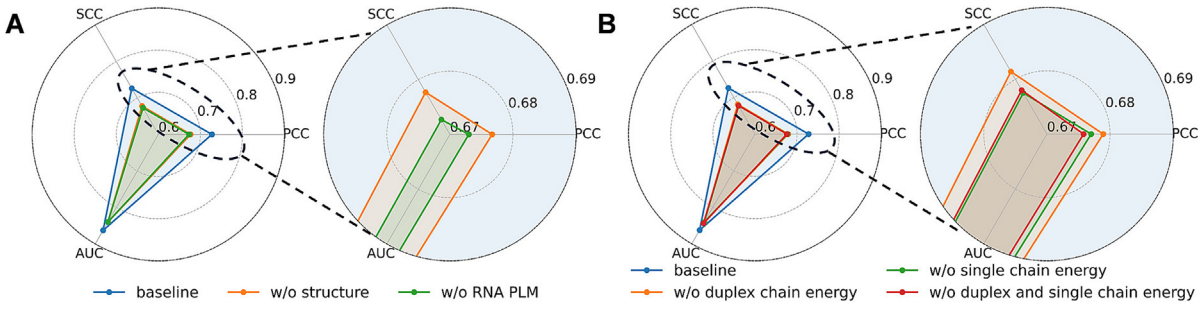

Figure 3. (**A**) The performance comparison of PCC, SPCC, and AUC in ablation of structure (w/o structure) and RNA pretrained language model (w/o RNA PLM). (**B**) The performance comparison of PCC, SPCC, and AUC in ablation of single-chain energy feature (w/o single chain energy), duplex-chain energy (w/o duplex chain energy) and the both (w/o duplex and single chain energy).

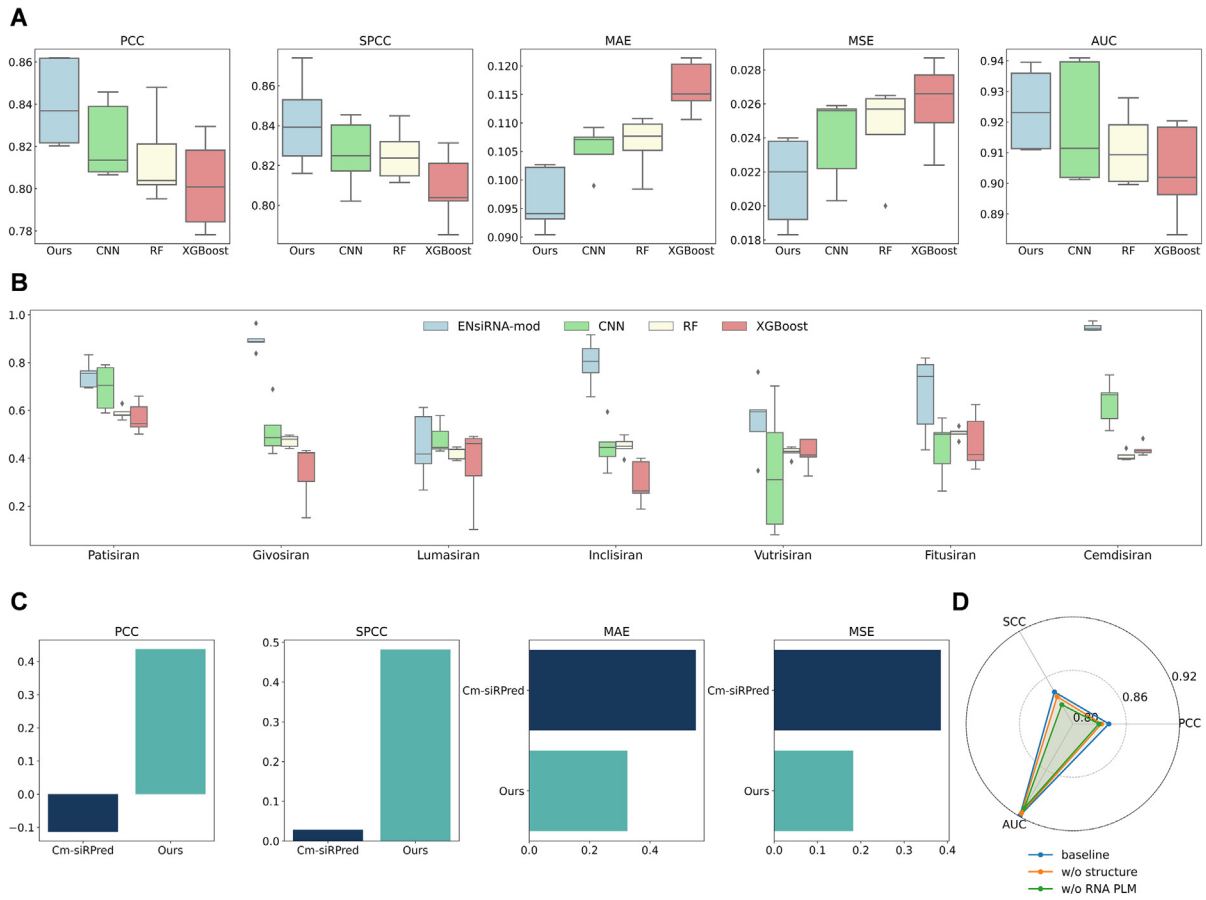

Figure 4. (**A**) Figure depicted the performance comparison on 5-fold cross-validation of modified siRNA model, where Ours means our method ENsiRNA-mod, CNN means the recreation version of Cm-siRPred, RF means Random Forest. (**B**) Figure depicted the prediction distribution siRNA drugs dataset of modified siRNA model. (**C**) Performance comparison in real-world drug design dataset experiments, where Cm-siRPred here is tested from website. (**D**) Figure depicted the performance comparison in ablation experiments of siRNA structure (w/o structure) and RNA pretrained language model (w/o RNA PLM).

method maintains a consistently higher median performance across all metrics, indicating improved predictive capability.

Additionally, we compared the performance of deep learning methods on the internal test set from the same source, as shown in Table S7. Here, ENsiRNA-mod maintained superior performance across most metrics, showcasing the robustness of our method.

Performance on siRNA drug dataset. To better evaluate the generalizability of our method, we compared it with related methods using the siRNA drug dataset as an external test set. We also calculated the Levenshtein distances between the siRNA sequences in the siRNA drug dataset derived from 5-fold cross-validation to confirm the external nature of this dataset.

Unlike previous approaches, we tested all 5-fold models on the siRNA drug dataset. The predictions from these models are shown in Figure 4B and detailed in Table S8. Additionally, we computed the Brier scores for all methods. Our method achieved the lowest Brier score of 0.1128, compared to scores of 0.2669 for CNN, 0.2871 for Random Forest, and 0.3554 for XGBoost. The results of ENsiRNA-mod demonstrate robustness, while the other methods exhibited instability, significantly influenced by the training data splits. This suggests that while both deep learning methods may fit the training data well, their feature extraction techniques fail to effectively capture the relationships between modifications, sequence, and structure of siRNA, thereby limiting their generalizability to external datasets and real-world applications.

Evaluation on real-world drug design dataset. To evaluate the potential for real-world application in modified siRNA design, we designed an experiment predicting siRNA with different modification positions after confirming their sequences and modification type. The dataset is also collected from the patent of Givosiran, which

is depicted in Figure 4C. Because the test data here is unseeable for Cm-siRPred, we compared Cm-siRPred on the website with our method. Although the performance of ENsiRNA-mod is not as strong as that demonstrated in the 5-fold cross-validation and the internal test set, it still outperforms Cm-siRPred despite the challenges posed by the wide compound space and limited training data. ENsiRNA-mod achieves a PCC of 0.4376 and an SPCC of 0.4821, while Cm-siRPred failed to produce predictions with a significant positive correlation, yielding a PCC of −0.1138 and an SPCC of 0.0286. These results showcase that Cm-siRPred could not distinguish the minor differences in modification changes, whereas our method performs better.

Ablation experiment of ENsiRNA-mod. In this section, we validate the necessity of specific components in our method through an ablation study based on 5-fold cross-validation. This study includes the structural information of siRNA and the embeddings from a pretrained RNA language model. The results are shown in radar chart Figure 4D and detailed in Table S9. We observed that the ablation of structural information (orange line) and pretrained RNA language model (green line) resulted in a decrease of 0.99% and 1.33% in PCC, as well as a decrease of 0.72% and 1.96% in SPCC, respectively. These findings, in conjunction with the previous ablation study, demonstrate the importance of multimodal features.

Discussion

To elucidate the role of ENsiRNA, we collected embeddings of the final global feature, mRNA global node feature, and siRNA global node feature from the external test set, which includes siRNAs at varying positions on a specific mRNA. Then, we created a t-SNE graph displaying three levels of features. As shown in Figure 5, all siRNA embeddings cluster into positive or negative groups. We observe that the features of siRNA nodes and global features are more distinctly separated than those of mRNA nodes, suggesting that while both siRNA and mRNA are significant in siRNA-mRNA prediction, siRNA plays a more pivotal role. While there are still limitations in fully understanding how the model works and challenges in classifying feature contributions due to the multimodality strategy, this observation motivated us to further investigate how structural information influences the predictive performance of ENsiRNA.

To better understand the influence of structural information in ENsiRNA, we conducted a detailed case study comparing predictions between ENsiRNA and existing method. We examined siRNA T631 (antisense sequence: 'CGCGCUUCUCGUUGGGGUC') from the external test set, which has an experimental efficacy of 0.32. ENsiRNA accurately predicted its efficacy at 0.44 (MAE = 0.11), whereas AttSiOff, another deep learning approach utilizing the same pretrained RNA language model, substantially overestimated it to 0.81 (MAE = 0.49). We compared T631 with its adjacent siRNA T632 (antisense sequence: 'UCGCGCUUCUCGU UGG GGU', efficacy = 0.86) to elucidate this disparity (Figure 6A).

First, we analyzed the structural differences between high-efficacy (efficacy > 70%) and low-efficacy (efficacy < 30%) siRNAs in the 5-fold cross-validation dataset and the two example cases. We simplified the structure representation to C4′ atoms, calculating the distances between nucleic acids within the same strand (intra-chain) and between complementary strands (inter-chain) following our graph construction. As shown in Figure S5A, the 5-fold cross-validation dataset shows that high-efficacy siRNAs exhibit larger inter-chain distances at the 5′-end of antisense strand (position 1), consistent with previous studies demonstrating that lower internal stability at the 5′ end ensures proper siRNA unwinding and

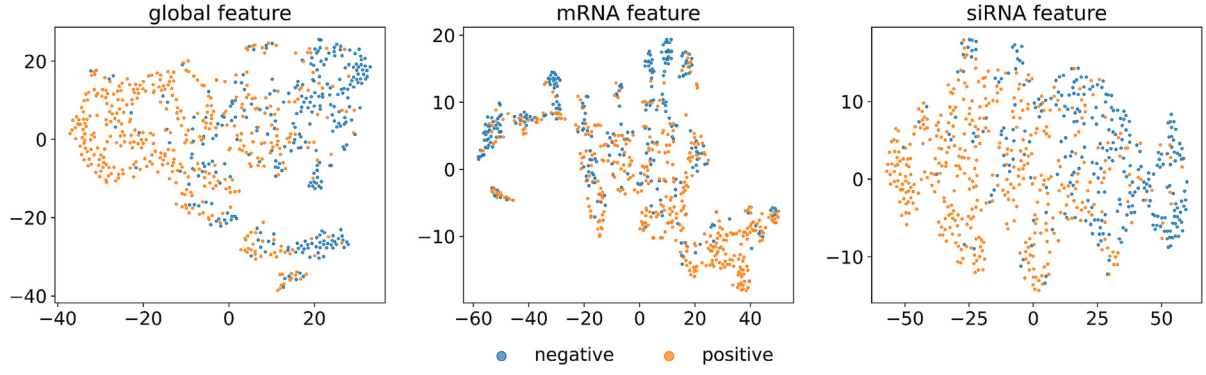

Figure 5. Figure depicted the t-SNE graph of different feature embeddings of siRNA in ENsiRNA, including the global feature, the feature of mRNA global node and the feature of siRNA global node.

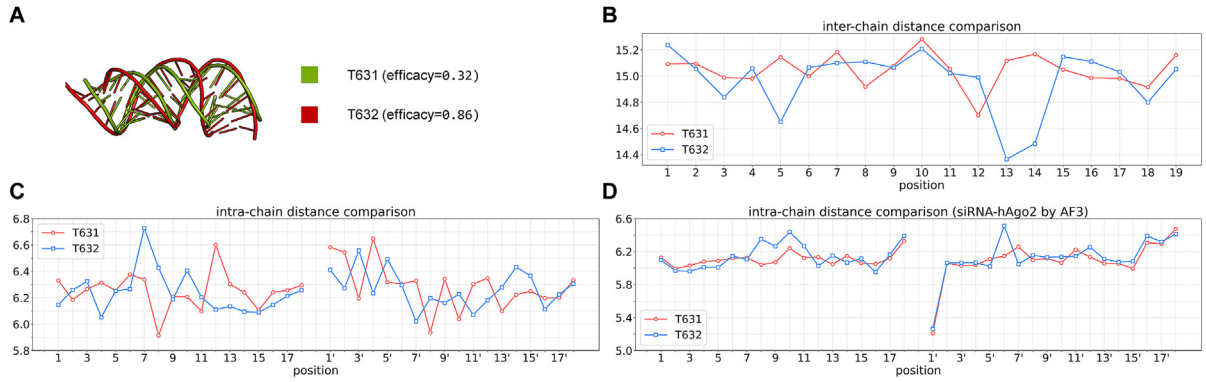

Figure 6. (**A**) Figure depicted three-dimensional structures of T631 (low efficacy) and T632 (high efficacy) predicted by FARFAR2. (**B**) Figure depicted the inter-chain distance comparison of T631 and T632 structure predicted by FARFAR2, where positions 1–19 represent nucleotides from 5′ to 3′ of the antisense strand. (**C**) Figure depicted the intra-chain distance comparison of T631 and T632 predicted by FARFAR2, where positions 1–18 represent distance between nucleotides from 5′ to 3′ of the sense strand, and positions 1′–18′ represent distance between nucleotides from 5′ to 3′ of the antisense strand. (**D**) Figure depicted the intra-chain distance comparison of T631 and T632 from T631-hAgo2 complex and T632-hAgo2 complex predicted by Alphafold3, where positions 1–18 represent distance between nucleotides from 5′ to 3′ of the sense strand, positions 1′–18′ represent distance between nucleotides from 5′ to 3′ of the antisense strand, hAgo2 refers to human Argonaute-2 protein.

RISC entry.[33,44,45] This pattern is also observed in the comparison between T631 and T632 (Figure 6B).

Beyond previous findings, we discovered distinct patterns in intra-chain in both 5-fold cross-validation dataset (Figure S5B) and two cases (Figure 6C): 5′-end distances of the antisense strand (position 1′), where high-efficacy siRNAs show shorter distances compared to low-efficacy ones. To further investigate this distinction, we employed AlphaFold3 to explore siRNA-human Ago2 structures for both T631 and T632,[46] which structures are shown in Figure S6A and B.[47] We can observe that the antisense strand's 5′ end of both T631 and T632 dock into Argonaute (AGO)'s 5′ binding pocket, as described in 48. These patterns are reflected in the notably shorter intra-chain distance at position 1′ compared to other positions in Figure 6D. Based on these observations, we hypothesized that the notably shorter distance at the antisense strand's 5′ end (position 1′) facilitates the transition from siRNA duplex to siRNA-hAgo2 binding conformations.

In conclusion, while FARFAR2-generated structures may not be completely accurate, the conformations obtained through all-atom scoring function minimization represent reasonable low-energy states. Given that both ENsiRNA and AttSiOff utilize the same RNA language model features, we propose that the structural properties of siRNA provide more informative features, enabling ENsiRNA to achieve superior prediction accuracy compared to AttSiOff. While our dataset may contain additional unexplored structural patterns related to siRNA efficacy beyond the scope of this study, these findings lay the groundwork for future investigations. Our ongoing research will focus on elucidating the mechanistic relationships between siRNA efficacy and the structural characteristics of siRNA, mRNA, and Ago2.

While we have demonstrated the advantages of our method, there remain potential areas for improvement. Firstly, although we have augmented our dataset by incorporating patent data and integrating it with a pretrained RNA Transformer model, the lack of sufficient data still limits generalization, particularly for modified siRNAs. In this context, a larger, open-source database specifically for modified siRNA would improve the generalization of our model. To support this, we have made our codebase and model fully open-source, allowing the research community to access, retrain, and fine-tune it with additional data. This open-science approach encourages model refinement and broadens its application across research domains, helping address data limitations and promote collaborative improvements.

Secondly, we have noted that the improvement in siRNA-mRNA prediction is greater than in modified siRNA prediction, which may be due to the unmodified structure we used. In this study, we constructed the siRNA structure based on FARFAR2, because the currently locally available deep learning methods do not support for duplex strand RNA.[49] However, FARFAR2 is not sufficiently accurate and fast, only supporting a limited number of modified nucleic acids. The further development of siRNA folding techniques, particularly deep learning approaches such as accurate RNA-protein complex and their dynamic structure transformation, would potentially enhance our method's performance[50].

Thirdly, due to the cross-species training data and the conserved RISC mechanism among mammals, we hypothesized that this method could be applied to different mammalian species. We included a preliminary test in the supplementary data (Table S10); however, further biological experiments are needed for validation.[35]

Finally, due to the limited availability and imbalance between modified and unmodified counterparts, as well as the sequence selection bias in modified siRNA data, we divided our approach into two steps: ENsiRNA and ENsiRNA-mod. Although this design aligns with real-world applications, we believe it could be further improved to predict modified siRNA in conjunction with mRNA and energy, especially as more data accumulates on mRNA.

In summary, ENsiRNA's strong performance and adaptability highlight its practical value in drug design and siRNA therapeutic applications. Future work should prioritize expanding the dataset and improving siRNA folding techniques to enhance model generalization and prediction accuracy, particularly for modified siRNA.

Data availability

The codes, models and dataset are available at https://github.com/tanwenchong/ENsiRNA.

CRediT authorship contribution statement

Wenchong Tan: Writing – original draft, Visualization, Validation, Software, Methodology, Conceptualization. **Mingshu Dai:** Validation, Investigation, Data curation. **Shimin Ye:** Validation, Investigation, Data curation. **Xin Tang:** Visualization, Methodology. **Dawei Jiang:** Writing – review & editing, Methodology. **Dong Chen:** Resources. **Hongli Du:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Funding

This work was supported by the National Key R&D Program of China (2024YFF1206603, 2018YFC0910201), and the Key R&D Program of Guangdong Province (2019B020226001).

DATA AVAILABILITY

Data will be made available on request.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jmb.2025.169131.

Received 12 November 2024;
Accepted 1 April 2025;
Available online 5 April 2025

Keywords:
siRNA;
deep learning;
multimodality;
geometric graph neural network

References

1. Almeida, R., Allshire, R.C., (2005). RNA silencing and genome regulation. *Trends Cell Biol.* **15**, 251–258.
2. Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., Plasterk, R.H., (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans. Genes Dev.* **15**, 2654–2659.
3. Dana, H., Chalbatani, G.M., Mahmoodzadeh, H., Karimloo, R., Rezaiean, O., Moradzadeh, A., et al., (2017). Molecular mechanisms and biological functions of siRNA. *Int. J. Biomed. Sci.* **13**, 48.
4. Guo, S., Zhang, M., Huang, Y., (2023). Three 'E'challenges for siRNA drug development. *Trends Mol. Med.*.
5. Ichihara, M., Murakumo, Y., Masuda, A., Matsuura, T., Asai, N., Jijiwa, M., et al., (2007). Thermodynamic instability of siRNA duplex is a prerequisite for dependable prediction of siRNA activities. *Nucleic Acids Res.* **35**, e123.
6. Davis, S.M., Sousa, J., Vangjeli, L., Hassler, M.R., Echeverria, D., Knox, E., et al., (2020). 2′-O-Methyl at 20-mer guide strand 3′ termini may negatively affect target silencing activity of fully chemically modified siRNA. *Mol. Therapy-Nucleic Acids* **21**, 266–277.
7. Hu, B., Zhong, L., Weng, Y., Peng, L., Huang, Y., Zhao, Y., et al., (2020). Therapeutic siRNA: state of the art. *Signal Transduct. Target. Ther.* **5**, 101.
8. Wu, S.Y., Yang, X., Gharpure, K.M., Hatakeyama, H., Egli, M., McGuire, M.H., et al., (2014). 2′-OMe-phosphorodithioate-modified siRNAs show increased loading into the RISC complex and enhanced anti-tumour activity. *Nature Commun.* **5**, 3459.
9. Crooke, S.T., Wang, S., Vickers, T.A., Shen, W., (2017). Liang X-h. Cellular uptake and trafficking of antisense oligonucleotides. *Nature Biotechnol.* **35**, 230–237.
10. Friedrich, M., Aigner, A., (2022). Therapeutic siRNA: state-of-the-art and future perspectives. *BioDrugs* **36**, 549–571.
11. Kraynack, B.A., Baker, B.F., (2006). Small interfering RNAs containing full 2′-O-methylribonucleotide-modified sense strands display Argonaute2/eIF2C2-dependent activity. *RNA* **12**, 163–176.
12. Alshaer, W., Zureigat, H., Al Karaki, A., Al-Kadash, A., Gharaibeh, L., Ma'mon, M.H., et al., (2021). siRNA: mechanism of action, challenges, and therapeutic approaches. *Eur. J. Pharmacol.* **905**, 174178.

13. Ahn, I., Kang, C.S., Han, J., (2023). Where should siRNAs go: applicable organs for siRNA drugs. *Exp. Mol. Med.* **55**, 1283–1292.
14. Huesken, D., Lange, J., Mickanin, C., Weiler, J., Asselbergs, F., Warner, J., et al., (2005). Design of a genome-wide siRNA library using an artificial neural network. *Nature Biotechnol.* **23**, 995–1001.
15. Vert, J.-P., Foveau, N., Lajaunie, C., Vandenbrouck, Y., (2006). An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinf.* **7**, 1–17.
16. Han, Y., He, F., Chen, Y., Liu, Y., Yu, H., (2018). SiRNA silencing efficacy prediction based on a deep architecture. *BMC Genomics* **19**, 59–65.
17. La Rosa, M., Fiannaca, A., La Paglia, L., Urso, A., (2022). A graph neural network approach for the analysis of siRNA-target biological networks. *Int. J. Mol. Sci.* **23**, 14211.
18. Liu, B., Yuan, Y., Pan, X., Shen, H.-B., Jin, C., (2024). AttSiOff: a self-attention-based approach on siRNA design with inhibition and off-target effect prediction. *Med-X* **2**, 5.
19. Vaswani, A., (2017). Attention is all you need. *Adv. Neural Inf. Proces. Syst.*.
20. Dar, S.A., Gupta, A.K., Thakur, A., Kumar, M., (2016). SMEpred workbench: a web server for predicting efficacy of chemicallymodified siRNAs. *RNA Biol.* **13**, 1144–1151.
21. Martinelli, D.D., (2024). From sequences to therapeutics: using machine learning to predict chemically modified siRNA activity. *Genomics*110815.
22. Liu, T., Huang, J., Luo, D., Ren, L., Ning, L., Huang, J., et al., (2024). Cm-siRPred: predicting chemically modified siRNA efficiency based on multi-view learning strategy. *Int. J. Biol. Macromol.* **264**, 130638.
23. Han, J., Cen, J., Wu, L., Li, Z., Kong, X., Jiao, R., et al. (2024). A survey of geometric graph neural networks: data structures, models and applications. arXiv preprint arXiv:240300485.
24. Zhang, X., Zhang, O., Shen, C., Qu, W., Chen, S., Cao, H., et al., (2023). Efficient and accurate large library ligand docking with KarmaDock. *Nature Comput. Sci.* **3**, 789–804.
25. Kong, X., Huang, W., Liu, Y. (2023). End-to-end full-atom antibody design. arXiv preprint arXiv:230200203.
26. Satorras, V.G., Hoogeboom, E., Welling, M., (2021). E (n) equivariant graph neural networks. *In: International conference on machine learning*. PMLR, pp. 9323–9332.
27. Watkins, A.M., Rangan, R., Das, R., (2020). FARFAR2: improved de novo rosetta prediction of complex global RNA folds. *Structure* **28**, 963–976.e6.
28. Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., et al., (2011). ViennaRNA Package 2.0. *Algorith. Mol. Biol.* **6**, 1–14.
29. Mysara, M., Garibaldi, J.M., ElHefnawi, M., (2011). MysiRNA-designer: a workflow for efficient siRNA design. *PLoS One* **6**, e25642.
30. Harborth, J., Elbashir, S.M., Vandenburgh, K., Manninga, H., Scaringe, S.A., Weber, K., et al., (2003). Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev.* **13**, 83–105.
31. Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., et al., (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res.* **32**, 936–948.
32. Vickers, T.A., Koo, S., Bennett, C.F., Crooke, S.T., Dean, N.M., Baker, B.F., (2003). Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents: a comparative analysis. *J. Biol. Chem.* **278**, 7108–7118.
33. Khvorova, A., Reynolds, A., Jayasena, S.D., (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–216.
34. Katoh, T., Suzuki, T., (2007). Specific residues at every third position of siRNA shape its efficient RNAi activity. *Nucleic Acids Res.* **35**, e27.
35. Guo, B., Zhang, B., Zheng, L., Tang, T., Liu, J., Wu, H., et al., (2014). Therapeutic RNA interference targeting CKIP-1 with a cross-species sequence to stimulate bone formation. *Bone* **59**, 76–88.
36. Dar, S.A., Thakur, A., Qureshi, A., Kumar, M., (2016). siRNAmod: a database of experimentally validated chemically modified siRNAs. *Sci. Rep.* **6**, 20031.
37. Li, Y., Zhang, C., Feng, C., Pearce, R., Lydia Freddolino, P., Zhang, Y., (2023). Integrating end-to-end learning with deep geometrical potentials for ab initio RNA structure prediction. *Nature Commun.* **14**, 5745.
38. Mao, F.Y., Tu, M.-J., Traber, G.M., Yu, A.-M., (2024). Comparison of three computational tools for the prediction of RNA tertiary structures. *Noncoding RNA* **10**, 55.
39. Bernard, C., Postic, G., Ghannay, S., Tahi, F., (2024). Has AlphaFold 3 reached its success for RNAs? *bioRxiv*2024.06.13.598780.
40. Chen, J., Hu, Z., Sun, S., Tan, Q., Wang, Y., Yu, Q., et al. (2022). Interpretable RNA foundation model from unannotated data for highly accurate RNA structure and function predictions. arXiv preprint arXiv:220400300.
41. Shen, T., Hu, Z., Sun, S., Liu, D., Wong, F., Wang, J., et al., (2024). Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nature*, 1–12.
42. Chen, T., Guestrin, C., (2016). Xgboost: a scalable tree boosting system. *In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785–794.
43. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al., (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830.
44. Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., Khvorova, A., (2004). Rational siRNA design for RNA interference. *Nature Biotechnol.* **22**, 326–330.
45. Lisowiec-Wąchnicka, J., Bartyś, N., Pasternak, A., (2019). A systematic study on the influence of thermodynamic asymmetry of 5′-ends of siRNA duplexes in relation to their silencing potency. *Sci. Rep.* **9**, 2477.
46. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., et al., (2024). Accurate structure prediction of biomolecular interactions with AlphaFold. *Nature* **3**, 1–3.
47. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., et al., (2023). UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* **32**, e4792.
48. Iwakawa, H.-O., Tomari, Y., (2022). Life of RISC: formation, action, and degradation of RNA-induced silencing complex. *Mol. Cell* **82**, 30–43.
49. Shen, T., Hu, Z., Peng, Z., Chen, J., Xiong, P., Hong, L., et al. (2022). E2Efold-3D: end-to-end deep learning method for accurate de novo RNA 3D structure prediction. arXiv preprint arXiv:220701586.
50. Jing, B., Berger, B., Jaakkola, T., (2024). AlphaFold meets flow matching for generating protein ensembles. *arXiv*.