TIP-Net: A Thermodynamics-Informed Deep Learning Framework for CRISPR-Cas9 gRNA On-Target Efficiency Prediction

Kanika Garg[1*], Rajiv Singh[2], and Nishchal Kumar Verma[3]

[1] Department of CSE-AI, Inderprastha Engineering College, Ghaziabad, U.P., India
[2] Department of Computer Science, Banasthali Vidyapith, Rajasthan, India
[3] Department of Electrical Engineering, Indian Institute of Technology, Kanpur, India

gargkanika7@gmail.com, jkrajivsingh@gmail.com, nishchal@iitk.ac.in*

Abstract. Predictive modeling of CRISPR-Cas9 guide RNA (gRNA) on-target editing efficiency is critical for achieving robust results in genome editing. However, current deep learning models such as DeepHF, CRISPRon, and C-RNNCrispr rely almost entirely on one-hot encoding of the gRNA sequence, largely overlooking the biophysical properties governing RNA:DNA hybridization and guide folding. We introduce TIP-Net, an architecture that integrates convolutional sequence encoding with RNA:DNA interaction thermodynamics through a physics-gated residual fusion model, incorporating duplex melting temperature (Sugimoto nearest-neighbor parameters), seed-region thermodynamics, and guide self-structure computed via the Nussinov algorithm as features. To study the effect of thermodynamic fusion and the Physics-Consistency Regularization (PCR) loss on the CRISPRon dataset, we present an ablation analysis of TIP-Net across three settings: sequence-only, sequence plus physics (concatenated), and the full TIP-Net model. We find that physics-based fusion improves biological interpretability and is complementary to sequence-based deep learning approaches.

Keywords: CRISPR-Cas9 · gRNA efficiency prediction · deep learning thermodynamics · physics-informed neural networks · attention pooling.

1 Introduction

CRISPR-Cas9 has emerged as the most widely used method for genome editing, but the editing efficiency of sgRNAs varies substantially depending on their sequence [1]. Accurately predicting on-target efficiency has therefore become important, and numerous machine learning and deep learning models have been developed for this purpose, including Azimuth [1], DeepHF [2], C-RNNCrispr [3], and CRISPRon [4].

The vast majority of CRISPR prediction systems use sequence-based encoders (CNN/RNN) with little or no information about biophysics [4,5,6]. Although recent attention mechanisms enhance sequence encoding, they lack biophysical priors, which have been shown to be important for editing efficiency [7,8].

We propose TIP-Net (Thermodynamics-Informed Prediction Network), which explicitly incorporates a sequence-encoding CNN branch alongside a physics branch

© The Author(s) 2026
A. Shastri et al. (eds.), *Proceedings of the Fourth Workshop on Applications of Computational Intelligence and Interactive Systems (WACIIS 2026)*, Advances in Intelligent Systems Research 217,
https://doi.org/10.2991/978-94-6239-832-0_12

that computes (a) the RNA:DNA duplex melting temperature based on Sugimoto's nearest-neighbor parameters [7], (b) the melting temperature of the PAM-proximal seed region, and (c) a Nussinov-algorithm-based [9] self-folding score that quantifies hairpin propensity. These features are fused with the sequence embedding through physics-gated residual fusion, while training incorporates a Physics-Consistency Regularization (PCR) loss that penalizes predictions contradicting the expected negative correlation between guide structure and efficiency.

2 Related Work

Sequence-only deep learning predictors. Early CNN/RNN-based models such as DeepHF [2], C-RNNCrispr [3], and CRISPRon [4] predict efficiency from one-hot-encoded gRNA sequences, with CRISPRon additionally introducing a single scalar for binding energy. Transformer- and graph-based models, including TransCrispr [5], CRISPR-OTE [6], DeepCRISTL [13], Graph-CRISPR [14], and protein-language-model transfer learning [15] extend encoder capacity and improve generalization [10,11,16], yet thermodynamics and secondary structure remain embedded within the encoder rather than treated as a separate branch, as in TIP-Net.

Thermodynamics and physics-informed learning. Nearest-neighbor parameters for DNA-DNA [8] and RNA-DNA [7] duplexes, together with the Nussinov self-folding algorithm [9], form the traditional basis for estimating melting temperature and secondary structure from sequence alone. Physics-informed neural networks [12] embed physical constraints into the training objective rather than the input space alone .

3 Methodology

3.1 Problem Formulation

We consider a 30−nucleotide gRNA context sequence $x \in \{A, C, G, T\}^{30}$, one−hot encoded as $X \in R^{30\times4}$. The task is to regress the measured on−target editing efficiency $y \in R$.

3.2 Architecture Overview

TIP-Net (Fig. 1) consists of two parallel branches whose representations are fused before a regression head.

Sequence branch. Three parallel 1D convolutions with kernel sizes $k \in \{3, 5, 7\}$ first generate motif representations at different widths; these are concatenated and fused via a 1×1 convolution, followed by a residual convolutional block. Rather than average−pooling over sequence positions, an attention−pooling module computes attention weights $\alpha_t = \mathrm{softmax}_t(v^T \tanh(W_p h_t + b_p))$, giving $e_{seq} = \Sigma_t \alpha_t h_t \in R^{64}$.

Physics branch. The physics branch computes: (i) T_{hybrid}, the hybridization melting temperature of the full 30-mer sequence, calculated from the Sugimoto nearest-neighbor parameter table [7]. (ii) T_{seed}, the melting temperature of the PAM-proximal 12-nt seed region. (iii) S_{fold}, the Nussinov self-folding score [9] (the maximum number of non-crossing base pairs the sequence forms with itself, normalized by sequence length. and (iv) GC content, terminal (3′) GC content, and a 29-element positional GC-stacking profile. Scalar features are passed through a small MLP, while the positional profile is passed through a 1D convolution and their concatenation is projected into $e_{phys} \in \mathbb{R}^{64}$.

Physics-gated fusion. Rather than simple concatenation, a learned gate determines how much physical information is incorporated into the sequence embedding:

$$g = \sigma(W_g[e_{seq}; e_{phys}] + b_g), \quad e_{fused} = e_{seq} + g \odot e_{phys} \tag{1}$$

The fused representation e_{fused} is passed to a two-layer MLP regression head that produces the prediction $\hat{y}$.

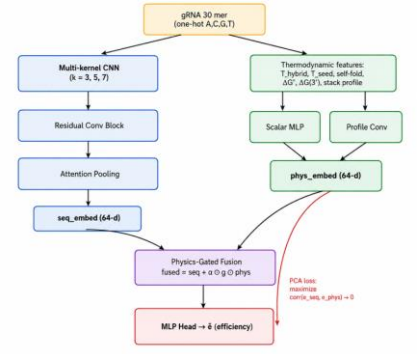

Fig. 1. TIP-Net architecture. A sequence CNN branch and a thermodynamic feature branch are combined via physics-gated residual fusion.

3.3 Physics-Consistency Regularization (PCR) Loss

Guide self-structure is expected to compete with target binding; therefore, higher self-fold scores should not correlate with higher predicted efficiency. Let $\rho(\hat{y}, s)$ denote the Pearson correlation, within a training batch, between the predictions $\hat{y}$ and the standardized self-fold scores s. The loss is then defined as

$$L = (1/N)\, \Sigma_i\, (\hat{y}_i)^2 + \lambda \cdot \mathrm{ReLU}(\rho(\hat{y}, s)), \quad \lambda = 0.05. \tag{2}$$

Penalization occurs only for positive correlations, so the model remains free to learn negative or zero relationships, making the prior a flexible constraint rather than a rigid one.

3.4 Training Protocol

Sequence data are stratified into 64/16/20% training/validation/testing splits prior to reverse complementation, ensuring that no augmented mirror-image sequence from the test set leaks into training. Label min-max scaling parameters are determined from the training set alone. The model is trained using the Adam optimizer (weight decay 5×10^{-4}), a batch size of 32, early stopping on validation RMSE (patience 15), and gradient-norm clipping at 1.0.

Reproducibility details. Sequence branch consists of parallel 1D convolutions with 24 channels each, having kernels of size 3, 5, 7 and concatenated to form 64 channels with residual layer and dropout. Physics branch reduces 35 feature representations (6 scalar plus 29 dimensional profile) to 32 dimensional space and then concatenates to sequence representation. The 64 dimensional output from the gated fusion block is used in the regression block (64→32→1). Adam optimizer is used for training (lr=8×10^{-4}, wd=5×10^{-4}), ReduceLROnPlateau scheduler, batch size=32 and early stopping with patience=15.

4 Experimental Setup

Dataset. We use CRISPRon [4], an existing dataset of approximately 23,902 SpCas9 gRNAs with indel-frequency efficiency scores, one of the largest available benchmarks for on-target efficiency prediction.

Ablation variants.

(a) Sequence only: only the sequence branch and the regression head.

(b) Sequence and physics: the physics branch is combined via additive concatenation, without gating or the PCR loss.

(c) TIP-Net (full): physics-gated fusion (Eq. 1) with the PCR loss (Eq. 2).

Metrics. We report RMSE, MAE, coefficient of determination (R^2), and Spearman rank correlation. Spearman correlation is included because gRNA efficiency research emphasizes ranking guides correctly, which is often more important than predicting the exact efficiency value.

5 Results

5.1 Ablation Study

Table 1 reports test-set performance for the three TIP-Net variants.

Table 1. Ablation study on the CRISPRon held-out test set

Variant	RMSE↓	MAE↓	R^2↑	Spearman↑
Sequence only (baseline)	14.72	11.63	0.42	0.640
Sequence + Physics (concat)	14.52	11.46	0.43	0.646
TIP-Net (gated + PCR)	**14.42**	**11.33**	**0.45**	**0.654**

5.2 Biological Sanity Check

The Nussinov self-fold score shows only weak positive correlation with on-target efficiency on the validation set (Fig. 2), indicating it alone does not explain efficiency. However, it contributes meaningfully as a structural feature when combined with sequence and thermodynamic parameters in TIP-Net.

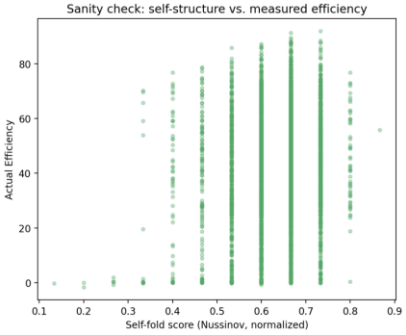

Fig. 2. TIP-Net biological sanity check on the held-out test set.

6 Discussion

The ablation study demonstrates that gated fusion outperforms concatenation (RMSE = 14.43 versus RMSE = 14.53), which suggests that even with the built-in sequence awareness of the CNN encoder, the network learns to recognize some thermodynamics-related information. A logical next step would be calculating the actual ΔG values.

Limitations. No statistical significance testing has yet been performed for the ablation improvements. All results are based on a single split of a single dataset (CRISPRon); neither validation on an independent dataset nor direct comparison with recent state-of-the-art models [14,15,17] was attempted, though both are natural next steps beyond the scope of this paper. The GC-stacking profile remains a heuristic approximation of the per-step ΔG value.

7 Conclusion

TIP-Net combines a CNN sequence encoder with a biophysical branch capturing duplex stability, seed-region stability, and self-folding propensity, integrated through a physics-gated fusion architecture and a newly proposed Physics-Consistency Regularization loss. By incorporating more than a single thermodynamic scalar, TIP-Net improves on prior work, and a systematic three-way ablation study quantifies the contribution of each component. Future work includes computing exact nearest-neighbor ΔG values in place of the heuristic stacking-energy profile, extending the physics branch to off-target prediction, and generalizing across datasets such as DeepHF and endogenous efficiency measurements.

References

1. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al.: Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology 34(2), 184–191 (2016)
2. Wang, D., Zhang, C., Wang, B., Li, B., Wang, Q., Liu, D., Wang, H., Zhou, Y., Shi, L., Lan, F., et al.: Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. Nature Communications 10, 4284 (2019)
3. Zhang, G., Dai, Z., Dai, X.: C-RNNCrispr: Prediction of CRISPR/Cas9 sgRNA activity using convolutional and recurrent neural networks. Computational and Structural Biotechnology Journal 18, 344–354 (2020)
4. Xiang, X., Corsi, G.I., Anthon, C., et al.: Enhancing CRISPR-Cas9 gRNA efficiency prediction by data integration and deep learning. Nature Communications 12, 3238 (2021)
5. Wan, Y., Jiang, S.: TransCrispr: Transformer based hybrid model for predicting CRISPR/Cas9 single guide RNA cleavage efficiency. IEEE/ACM Transactions on Computational Biology and Bioinformatics (2023)
6. Xie, J., et al.: CRISPR-OTE: A convolutional and bidirectional LSTM-based framework for gRNA on-target efficiency prediction. Computers in Biology and Medicine (2023)
7. Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., Sasaki, M.: Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry 34(35), 11211–11216 (1995)
8. SantaLucia, J.: A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proceedings of the National Academy of Sciences 95(4), 1460–1465 (1998)
9. Nussinov, R., Pieczenik, G., Griggs, J.R., Kleitman, D.J.: Algorithms for loop matchings. SIAM Journal on Applied Mathematics 35(1), 68–82 (1978)
10. Konstantakos, V., Nentidis, A., Krithara, A., Paliouras, G.: CRISPR-Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. Nucleic Acids Research 50(7), 3616–3637 (2022)
11. Sherkatghanad, Z., Abdar, M., Charlier, J., Makarenkov, V.: Using traditional machine learning and deep learning methods for on- and off-target prediction in CRISPR/Cas9: a review. Briefings in Bioinformatics 24(3), bbad131 (2023)
12. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707 (2019)
13. Elkayam, S., Tziony, I., Orenstein, Y.: DeepCRISTL: deep transfer learning to predict CRISPR/Cas9 on-target editing efficiency in specific cellular contexts. Bioinformatics 40(8), btae481 (2024)
14. Jiang, Y., Li, B., Xiong, J., Liu, X.: Graph-CRISPR: a gene editing efficiency prediction model based on graph neural network with integrated sequence and secondary structure feature extraction. Briefings in Bioinformatics 26(4), bbaf410 (2025)
15. Hou, Y., Li, Y., Zheng, R., Zhang, F., Guo, F., Li, M., Zeng, M.: Leveraging protein language models for cross-variant CRISPR/Cas9 sgRNA activity prediction. Bioinformatics 41(7), btaf385 (2025)
16. Saeed, M., Arham, M., Zafar, I., Jamal, A., Hussian, M., Usman, M., Bahwerth, F.S., Noman, M., Hossain, M.B.: Harnessing deep learning models for guide RNA optimization and off-target prediction in CRISPR systems. Biotechnology Journal 21(6) (2026)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

