mRNA stability prediction*

Shashank Pathak and Guohui Lin[0000−0003−4283−3396]

Department of Computing Science, University of Alberta, Canada
{spathak2, guohui}@ualberta.ca

Abstract. SARS-CoV-2 required the rapid development of vaccines. While conventional vaccine design techniques struggled due to the extensive genomic characteristics of the virus, the advent of mRNA-based vaccines presented a promising alternative. In this study we introduce a two-stage pipeline for mRNA stability prediction, by harnessing the potential of bidirectional GRU (bi-GRU) and the RNA secondary structure prediction tools: Vienna, Contrafold and RNAstructure. The performance measured in the mean column-wise root mean square error (MCRMSE) can be achieved at as low as 0.3031. Additionally, the secondary structure predicted from the three bioinformatics tools were observed statistical significantly better for stability prediction than the secondary structure coming with the dataset, on more than half of the sequences; yet the three tools themselves seemingly work equally well, i.e., statistical insignificantly different.

Keywords: mRNA stability · degradation · secondary structure · deep learning.

1 Introduction

The outbreak of COVID-19, caused by SARS-CoV-2 virus, had a profound global consequences, deeply impacting public health, economies, and daily life [2]. As the virus evolved, distinct variants emerged, each with unique transmissibility, severity, and immunity traits [8,29,30]. Genetic mutations in variants, pose challenges for vaccine development. Due to these rapid mutations, mRNA-based vaccine development gained focus over traditional vaccine design processes. Previously, vaccine development processes in the 20^{th} century, often taking decades to develop, relied on techniques like inactivated viruses, protein sub-units, and viral vectors [18]. Such methods were laborious, requiring virus cultivation, antigen purification, and extensive testing [21], making them slower and incompetent to mRNA based technology.

In contrast to traditional vaccine development, mRNA vaccines proved to be highly effective along with swift development, during the covid-19 pandemic [21]. The size of SARS-CoV-2 is quite large as compared to other viruses like influenza and smallpox. It puts limitations in standard vaccine processes, as they involve

* Supported by the NSERC Canada.

extensive cultivation and inactivation of the virus. Preserving the genetic information of large sized viruses is crucial for maintaining the structural integrity of vital antigens (for example the spike protein in SARS-CoV-2). Thus it puts forward a big challenge for non mRNA-based vaccines. Pfizer-BioNTech and Moderna were the frontiers in manufacturing production ready mRNA vaccines against SARS-CoV-2 but were limited by high mutation rates in the virus. Both Moderna and Pfizer-BioNtech vaccines demonstrated reduced efficacy against the Delta and Omicron variants, raising concerns about the vaccines' adaptability to newer variants [9].

According to central dogma of biology, genetic information flows from DNA to mRNA and then to protein [1]. mRNA serves as a blueprint for protein synthesis. This involves two main steps: transcription (DNA to RNA) and translation (RNA to protein). mRNA-based vaccines introduce a snippet of this genetic blueprint into the body, instructing it to produce a non-threatening version of the virus's spike protein, which is utilized by immune system to create antibodies. The complete mRNA vaccine design consists of 5'CAP, 5'UTR (untranslated region), ORF/CDS region, 3'UTR, and Poly(A) tail. The ORF/CDS is responsible for encoding target antigen protein, here spike protein of coronavirus. Other components of mRNA equally contribute and hold importance for efficient translation [10]. mRNA molecules however have inherent thermal instability *in vitro* and *in vivo*, reducing the efficacy of mRNA vaccines. It's still a challenge to identify which region is more prone to such instability. Secondary structure of mRNA is highly correlated to it's thermal stability.

In this work, emphasize is laid on mRNA stability due to its susceptibility to degradation under standard temperatures and various environmental conditions [28]. mRNA's vulnerability to in-line hydrolysis cleavage, which disrupts its backbone, necessitating it's preservation in constrained environmental condition. Since secondary structure of mRNA is influenced by both local and global nucleotide in the sequence, recurrent neural networks (RNNs), a deep learning algorithm, can be an appropriate model for learning long term sequential nonlinearities and context [24]. Given mRNA's sequential nature, we propose a novel pipeline for predicting degradation rates at each position of the sequence, using the RNN based GRU [3] model and then utilizing the trained model to compare and find stable secondary structures predicted from three different bioinformatics tools: Vienna, Contrafold and RNAstructure [13,7,22], and the secondary structure provided in [6]. For distinction, the secondary structure coming with the Open Vaccine dataset is referred to as the provided *test-data* secondary structure. By integrating GRU with each of the three secondary structure prediction tools, our proposed pipeline efficiently identifies and compares stable mRNA candidates, aiming to streamline in-vitro vaccine design. The adopted bi-GRU model is a basic RNN variant, prioritizing less trainable parameters over complex state-of-the-art models. The main reason for incorporating such a simple model is to focus on the comparative analysis on potentially employing the secondary structure prediction tools into the stability prediction pipeline.

The rest of the work is organized into five sections. Related work is presented in Section 2. Background knowledge regarding mRNA secondary structure and its prediction tools is contained in Section 3. The Open Vaccine dataset and the methodology are explained in Section 4. Section 5 contains results and discussion. The last section 6 concludes the work.

2 Related work

Artificial intelligence specifically deep learning has been instrumental in varied range of tasks in sciences including their impact on vaccines [17]. The introduction of the Open Vaccine dataset [6] advanced the machine learning community to address mRNA degradation. Early efforts in 2020 [25] used graph convolution network achieving low MCRMSE 2 loss values. Work done in [15] explored traditional machine learning models, with LGBM (Gradient Boosting), using a specific label encoding method, has achieved the smallest MCRMSE.

The comparative studies in [14] emphasized the superiority of recurrent neural networks over tree-based methods. [4] used ensemble GRU and LSTM models, leveraging feature-engineered BPPs dataset. Their LSTM+GRU+LSTM model stood out, reflecting a noticeable MCRMSE difference. A novel GCN-GRU model in [20] marked the state-of-the-art performance with a superior MCRMSE on a public dataset and their own private dataset. RNAdegformer [12] introduced a CNN with transformer network, converting RNA sequences into k-mer sequences, capturing both local and long-term dependencies. Their LOFO (leave one feature out) analysis revealed the nucleotide sequence as the most crucial feature.

The previous works have only utilized deep learning models for predicting degradation rates but not focused on utilizing it for investigating how secondary structure affects their stability. Our work primarily centers on a comparative study to pinpoint the importance of the secondary structure, that better predicted secondary structure would help predict the mRNA sequence stability.

3 Background

In the race to develop vaccines against SARS-CoV-2, mRNA-based vaccines have gained traction due to their speed and cost-effectiveness. Their performance, however, is intertwined with mRNA stability, deeply influenced by its secondary structure.

3.1 Secondary structure and mRNA stability

mRNA molecules, pivotal for vaccine development, are naturally unstable due to thermal susceptibilities. Their secondary structure, defined by specific base pairings, dictates their stability. Targeting at specific protein, designing the mRNA molecules with the optimal structure is essential for effective mRNA vaccines [27].

3.2 Secondary structure prediction tools

Several tools predict the RNA secondary structure quite accurately, providing insights into base pairings and molecular conformation. The three most known tools are:

- **Vienna [13]:** Renowned for using energy minimization, Vienna identifies structures achieving the minimum free energy (MFE).
- **Contrafold [7]:** Unlike Vienna, Contrafold adopts a log-linear model trained on established RNA structures, catering to both stability and sequence-specific aspects.
- **RNAstructure [22]:** Focusing on thermodynamics, RNAstructure offers a suite of algorithms for comprehensive RNA structure predictions.

Vienna, Contrafold, and RNAstructure have same running time complexity of $O(n^3)$ and space complexity of $O(n^2)$ [13,22,7].

4 Data and Method

4.1 Dataset

The dataset, sourced from the OpenVaccine Challenge Data [6] by Stanford University's Das Lab and Eterna, aims to predict mRNA base degradation rates. It comprises 3,000 RNA sequences, each either 107 or 130 bases long. For length-107 sequences, the first 68 are labeled; for length-130 sequences, the first 91 are. Labels of each base are the five degradation rates under five degradation conditions, respectively: **reactivity**, **deg pH10**, **deg Mg pH10**, **deg 50C**, and **deg Mg 50C**, representing various temperatures and chemical environments. That is, the degradation rate measurements are primarily based on conditions of pH = 10 and 37^{o}C, with and without Magnesium. In summary, an RNA sequence in the dataset has the features of a unique sequence ID, the sequence length (either 107 or 130), the nucleotide sequence, secondary structure defined by base pairings, SNR filter(signal to noise ratio), the number of labeled bases, and for each labeled base five degradation rates.

4.2 Methodology

Data preprocessing Data Pre-processing is a crucial step to transform the raw data into tensors to be utilized by the deep learning model. The first step involved looking for missing data for all 3000 samples. No missing data was found through *isnull()* operation in pandas. In the second step, to ensure non noisy samples are fed into the model, SNR filtering criteria was used. Sequences having SNR value less than 1 were removed. The dataset was split into training and test set in the third step. The training and test set consist of 1589 and 600 RNA sequences respectively. RNA sequences in the test set were passed through Vienna, Contrafold and RNAstructure to predict the secondary structures corresponding to each sequence, in the fourth step. Thus, 4 (3 predicted and provided from

test-data) different sets of secondary structure per sequence were created. In the last step, nucleotide sequence and the four sets of base pairing structure for each sequence in the training dataset were tokenized to convert non-numerical categorical feature into numeric tensors. The tokenization dictionary was utilized by the embedding layer later as a look up table.

The two phases, the model training and the performance testing, of our pipeline are depicted in Figures 1 and 2, respectively.

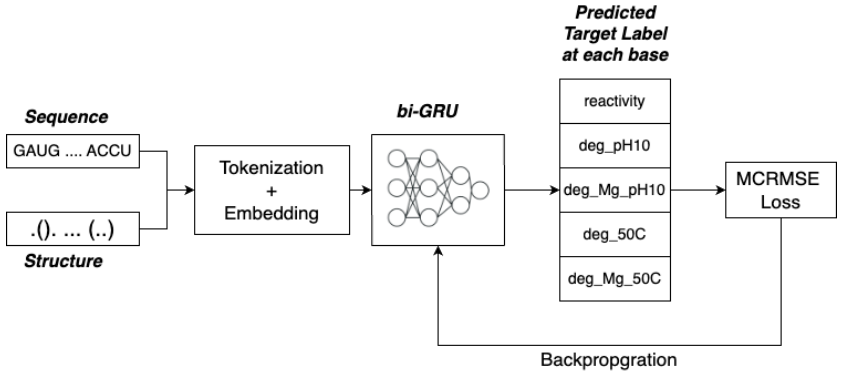


Fig. 1. The first stage of the pipeline — the model training.

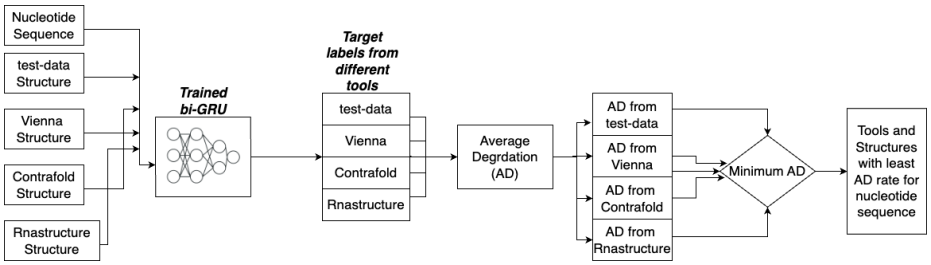


Fig. 2. The second stage of the pipeline — nucleotide sequence is paired with each of the four different sets of secondary structure and fed into the trained model to predict the degradation rates for each base.

Model RNA sequence and base pairing structure are sequential, encapsulating sequential context. Recurrent neural networks (RNN) suffer from vanishing gradient problem and it becomes bottleneck to capture longer sequences context. Gated recurrent unit (GRU) uses two gates to resolve the bottleneck. GRU was used in this task due to its simpler architecture, making it easy to train and

computationally less complex as compared to LSTM and transformer network. Tokenized nucleotide sequence and base pairing structure were the input to the model and output were the five degradation rates at each base of the RNA sequence. The hyper-parameter values used in the model were as follows: embedding dimension = 200, input size = 400, number of GRUs layers = 3, hidden layer dimension = 128, and bidirectional = True, making it a bi-GRU model. It helps to capture bi-directional context of the RNA sequence and its structure.

Post inference processing After training at inference, each sequence presents degradation value vectors of dimension either 68×5 or 91×5, contingent on its length. Given the intricacy in direct comparison of these values, predicted degradation values for each RNA sequence across the four test sets, as shown in 2, were averaged.

For the k-th sequence in the Vienna test set, for example, the average was computed via Equation (1).

$$Vienna\_k = \frac{1}{seq\_scored} \times \sum_{i=1}^{seq\_scored} \left(\frac{1}{5} \times \sum_{j=1}^{5} sequence\_degradation_{i,j} \right). \quad (1)$$

Upon post-processing, a statistical evaluation of the four averaged degradation values for each RNA sequence was done to pinpoint the most robust secondary structure among the quartet stemming from the diverse test sets.

Method pipeline In this study, we propose a workflow to predict the mRNA sequence stability, for identifying the best mRNA candidate vaccine sequences. The proposed method combines the potential of the deep learning model (bi-GRU) and the RNA secondary structure prediction tools. The pipeline involves two stages: In the first stage, depicted in Fig. 1, the bi-GRU model is trained on training dataset for predicting five degradation rates at each base of an RNA sequence. In the second stage, depicted in Fig. 2, the trained model is used to predict degradation rates at each base for each sequence, using four different sets of secondary structure, three predicted from Vienna [13], Contrafold [7] and RNAstructure [22] and one provided in the test dataset. The predicted degradation rates over an RNA sequence are summarized into a single score to rank the used secondary structure.

5 Results

5.1 Training and inference

The training dataset consists of tokenized 1589 sequences after performing all the data preprocessing steps as described in Section 4.2. The tokenized sequence was passed through an embedding layer [5] to learn contextual relationships between different bases in the sequence. The bi-GRU model was used learn sequential

non-linear context through nucleotide sequence and base pairing structure. To predict the 5 target degradation rates a linear layer was used. It takes in the last hidden layer state of the bi-GRU model and outputs a tensor of dimension $n \times 5$, where n is the nucleotide sequence length. To train a deep learning model, a loss function is used to learn parameters of the model through back-propagation. The loss function used for training was MCRMSE (mean column-wise root mean squared error, see Eq. (2)). There are multiple target labels, and hence MCRMSE is an appropriate metric for multi-labelled prediction task.

$$MCRMSE = \frac{1}{N_t}\sum_{j=1}^{N_t}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{ij} - \hat{y_{ij}})^2}, \tag{2}$$

where $N_t = 5$ represents the count of target labels ground truth values, y_{ij} is ground truth value of j^{th} label for the i^{th} sequence. $\hat{y_{ij}}$ is predicted value from the model for j^{th} label of the i^{th} sequence. $n = 68$ or 91 depending upon the RNA sequence length.

The model was trained for 70 epochs and 4-fold cross-validation for performing hyper-parameter tuning. The training and validation loss was logged for each fold. Fig. 3 plots the training and validation MCRMSE losses for the 4-th fold. The training loss for the model was **0.2991**, and the validation loss was **0.3031**. Due to the dropout layer, the validation loss is slightly higher than the training loss, showing the model training was good, leading to a robust model. Fig. 1 depicts the complete training procedure. Overall, the model performance was close to the transformer model (RNAdegformer [12]) and other ensemble learning methods [4].

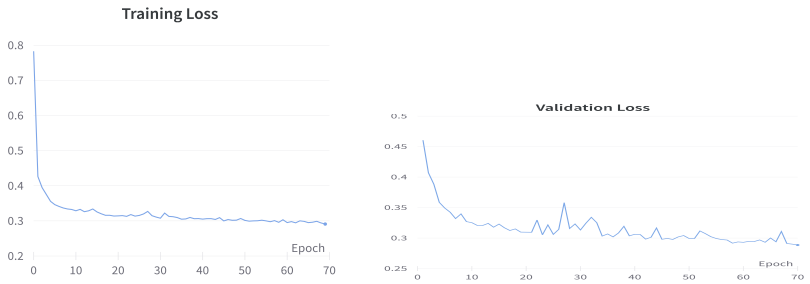
At inference the trained model was used for predicting the degradation rates at each position of the sequences together with four different sets of secondary structure. For test sets replication, Vienna, Contrafold, and RNAstructure RNA secondary structure prediction tools were installed and Arnie was used as the interface. The complete installation and setup were automated through a bash script.

Table 1. The 4-fold cross-validation MCRMSE losses.

Folds	Training Loss	Validation Loss
Fold-1	0.3156	0.3368
Fold-2	0.2944	0.2917
Fold-3	0.2957	0.2952
Fold-4	0.2910	0.2886

As depicted in Fig. 2, at inference each nucleotide sequence was paired with one of the four sets of secondary structure either predicted by the tools or provided in the test-data. Hence for each sequence we obtained four tensors of dimension $n \times 5$, where n is the sequence length. Each of these four tensors cor-

responding to an RNA sequence, were condensed into a single numerical score MCRMSE using Eq. (2).



(a) (b)

Fig. 3. Plots of the training MCRMSE loss (a) and the validation MCRMSE loss (b) over 70 epochs. The loss curves are for the fold-4 of the 4-fold-cross-validation on which the model achieved the minimum training and validation losses.

5.2 Comparative lens on secondary structure prediction tools

RNA secondary structure prediction is a computationally complex challenge due to large numbers of possible structures and long-range interactions [26]. In the stage- 2 of our pipeline, comparative analysis was performed between three RNA secondary prediction tools: Vienna, Contrafold, and RNAstructure, also comparing against the provided secondary structure for the test-data.

The test-dataset consists of 600 RNA sequences each with four different sets of secondary structure, three predicted and provided. For each sequence with one set of secondary structure, we average over all the predicted degradation values to give a single score for the set of secondary structure. Whichever set achieves the minimum score among the four is claimed the winner. Across 600 sequences, Vienna wins 222, Contrafold wins 241, and RNAstructure wins 222; while the provided secondary structure wins on only 153 sequences. Note that there are ties among the four.

Secondary structure prediction analysis The variability of the average degradation values across the four sets of secondary structure was analyzed on these 600 sequences. We first sorted the sequences in their non-decreasing order of the predicted average degradation values using the provided secondary structure, and use it as the baseline to plot the average degradation values using the predicted secondary structure (and plot them in Fig. 4). From these plots, it can be observed that, the variability in distribution of average degradation values were similar for Vienna and RNAstructure. Contrafold had a slightly larger

difference in their average degradation values against the test-data-method. Indeed, since Vienna and RNAstructure both predict secondary structure based on minimum free energy structures (MFE) [23], their degradation values follow a similar distribution in comparison with Contrafold, which uses a different conditional log-linear model [7].

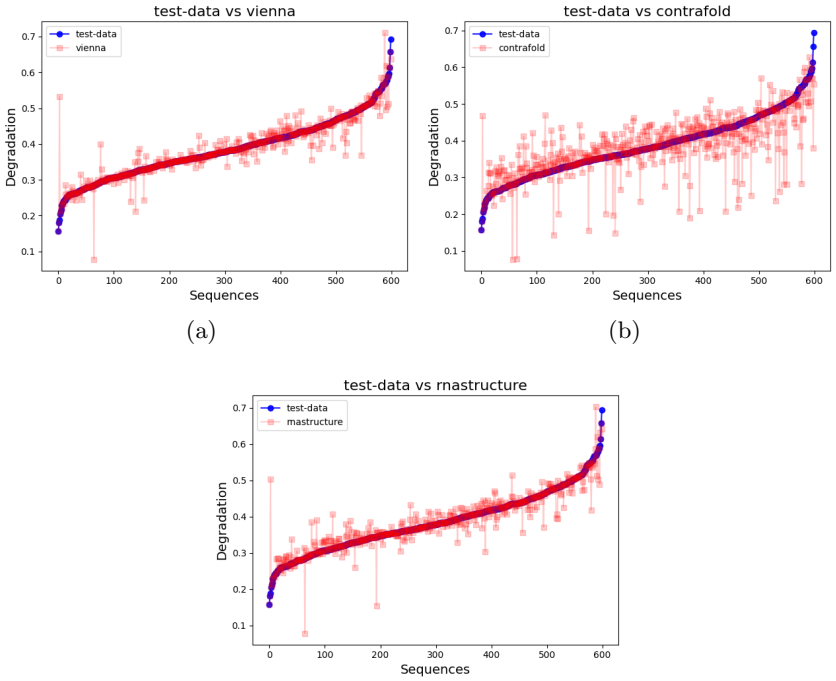


Fig. 4. The average degradation value (ADV) over 600 sequences in the test dataset. Using the secondary structure from the dataset, these sequences are sorted in non-decreasing order of their ADVs; the ADVs predicted using the predicted secondary structures by the three tools are plotted using this sequence order. (a) The ADVs by Vienna against by using the secondary structure from the dataset. (b) The ADVs by Contrafold against by using the secondary structure from the dataset. (a) The ADVs by RNAstructure against by using the secondary structure from the dataset.

Statistical testing The chi-square [19] test for independence was performed to evaluate statistical independence among RNA secondary structure prediction tools and the provided secondary structure. The test for independence was performed using the average predicted degradation values for the 600 sequences.

A pairwise comparison between each of the secondary structure prediction tools (Vienna, Contrafold, RNAstructure) and the provided secondary structure is done. All three p-values were less than 0.05, suggesting the predicted secondary structure differs statistical significantly from the provided. On the other hand, the pairwise comparison between the three prediction tools gave three p-value all greater than 0.075, implying no significant difference among the three sets of predicted secondary structure.

6 Conclusion

The effectiveness of mRNA vaccines is entangled with the stability of secondary structure of RNA sequences. The pipeline introduced in this work intends to address this pivotal factor, to enhance the reliability and potential success of mRNA vaccine candidates.

The first stage of pipeline goes through an extensive training phase to learn to predict degradation rates, which holds prime importance in the mRNA vaccine design process. The bi-directional aspect of the bi-GRU model ensures it captures information from both the 5' and 3' ends of RNA sequence, providing a comprehensive understanding of the degradation dynamics. This can help to make informed decisions about which RNA sequences to pursue in vaccine formulations, as sequences with high degradation rates may result in lower protein expression [11]. Following the prediction of degradation rates, the second stage of our pipeline integrates existing tools to predict secondary structures of given RNA sequences. The success of the pipeline is demonstrated through a comparative analysis on an Open Vaccine dataset. Notably, the secondary structure prediction tools were found to be capable statistical significantly.

To conclude, the stabilization of mRNA stands as a critical cornerstone for the future of mRNA vaccines, and our work would provide a streamline to advance the stability prediction to speed up the mRNA-vaccine development.

References

1. Akash, S., Arefine, M.F.: A short review on central dogma of molecular biology
2. Assefa, Y., Gilks, C.F., Reid, S., van de Pas, R., Gete, D.G., Van Damme, W.: Analysis of the COVID-19 pandemic: Lessons towards a more effective response to public health emergencies. Globalization and Health **18**, 10 (2022)
3. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078 (2014)
4. Chze, O.N., Abdullah, A.A.: COVID-19 mRNA vaccine degradation prediction by using deep learning algorithms. In: 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). pp. 444–450 (2022)
5. Dahouda, M.K., Joe, I.: A deep-learned embedding technique for categorical features encoding. IEEE Access **9**, 114381–114391 (2021)

6. Das, R.: OpenVaccine: COVID-19 mRNA vaccine degradation prediction (2020), https://kaggle.com/competitions/stanford-covid-vaccine
7. Do, C.B., Woods, D.A., Batzoglou, S.: CONTRAfold: RNA secondary structure prediction without physics-based models. Bioinformatics **22**, e90–e98 (2006)
8. Dong, Y., Dai, T., Wang, B., Zhang, L., Zeng, L.H., Huang, J., Yan, H., Zhang, L., Zhou, F.: The way of SARS-CoV-2 vaccine development: Success and challenges. Signal Transduction and Targeted Therapy **6**, 387 (2021)
9. Gardner, B.J., Kilpatrick, A.M.: Estimates of reduced vaccine effectiveness against hospitalization, infection, transmission and symptomatic disease of a new SARS-CoV-2 variant, Omicron (B.1.1.529), using neutralizing antibody titers. MedRxiv pp. 2021–12 (2021)
10. Gong, H., Wen, J., Luo, R., Feng, Y., Guo, J., Fu, H., Zhou, X.: Integrated mRNA sequence optimization using deep learning. Briefings in Bioinformatics **24**, bbad001 (2023)
11. Hargrove, J.L., Schmidt, F.H.: The role of mRNA and protein stability in gene expression. The FASEB Journal **3**, 2360–2370 (1989)
12. He, S., Gao, B., Sabnis, R., Sun, Q.: RNAdegformer: Accurate prediction of mRNA degradation at nucleotide resolution with deep learning. Briefings in Bioinformatics **24**, bbac581 (2023)
13. Hofacker, I.L.: Vienna RNA secondary structure server. Nucleic Acids Research **31**, 3429–3431 (2003)
14. Imran, S.A., Islam, M.T., Shahnaz, C., Islam, M.T., Imam, O.T., Haque, M.: COVID-19 mRNA vaccine degradation prediction using regularized LSTM model. In: IEEE International Women in Engineering Conference on Electrical and Computer Engineering (WIECON-ECE 2020). pp. 328–331 (2020)
15. Ing, S.H., Abdullah, A.A., Mashor, M.Y., Mohamed-Hussein, Z.A., Mohamed, Z., Ang, W.C.: Exploration of hybrid deep learning algorithms for COVID-19 mRNA vaccine degradation prediction system. International Journal of Advances in Intelligent Informatics **8** (2022)
16. Iserentant, D., Fiers, W.: Secondary structure of mRNA and efficiency of translation initiation. Gene **9**, 1–12 (1980)
17. Keshavarzi Arshadi, A., Webb, J., Salem, M., Cruz, E., Calad-Thomson, S., Ghadirian, N., Collins, J., Diez-Cecilia, E., Kelly, B., Goodarzi, H., et al.: Artificial intelligence for COVID-19 drug discovery and vaccine development. Frontiers in Artificial Intelligence p. 65 (2020)
18. Le, T.T., Andreadakis, Z., Kumar, A., Román, R.G., Tollefsen, S., Saville, M., Mayhew, S., et al.: The COVID-19 vaccine development landscape. Nature Reviews Drug Discovery **19**, 305–306 (2020)
19. McHugh, M.L.: The chi-square test of independence. Biochemia Medica **23**, 143–149 (2013)
20. Muneer, A., Fati, S.M., Akbar, N.A., Agustriawan, D., Wahyudi, S.T.: iVaccineDeep: Prediction of COVID-19 mRNA vaccine degradation using deep learning. Journal of King Saud University – Computer and Information Sciences **34**, 7419–7432 (2022)
21. Pardi, N., Hogan, M.J., Porter, F.W., Weissman, D.: mRNA vaccines – a new era in vaccinology. Nature Reviews Drug discovery **17**, 261–279 (2018)
22. Reuter, J.S., Mathews, D.H.: RNAstructure: Software for RNA secondary structure prediction and analysis. BMC Bioinformatics **11**, 1–9 (2010)
23. Seetin, M.G., Mathews, D.H.: RNA structure prediction: An overview of methods. Bacterial Regulatory RNA: Methods and Protocols pp. 99–122 (2012)

24. Sherstinsky, A.: Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena **404**, 132306 (2020)
25. Singhal, A.: Application and comparison of deep learning methods in the prediction of RNA sequence degradation and stability (2021)
26. Waterman, M.S., Smith, T.F.: RNA secondary structure: A complete mathematical analysis. Mathematical Biosciences **42**, 257–266 (1978)
27. Wayment-Steele, H.K., Kim, D.S., Choe, C.A., Nicol, J.J., Wellington-Oguri, R., Watkins, A.M., Parra Sperberg, R.A., Huang, P.S., Participants, E., Das, R.: Theoretical basis for stabilizing messenger RNA through secondary structure design. Nucleic Acids Research **49**, 10604–10617 (2021)
28. Zhang, N.N., Li, X.F., Deng, Y.Q., Zhao, H., Huang, Y.J., Yang, G., Huang, W.J., Gao, P., Zhou, C., Zhang, R.R., et al.: A thermostable mRNA vaccine against COVID-19. Cell **182**, 1271–1283 (2020)
29. Zhao, F., Zai, X., Zhang, Z., Xu, J., Chen, W.: Challenges and developments in universal vaccine design against SARS-CoV-2 variants. NPJ Vaccines **7**, 167 (2022)
30. Zhou, W., Wang, W.: Fast-spreading SARS-CoV-2 variants: challenges to and new design strategies of COVID-19 vaccines. Signal Transduction and Targeted Therapy **6**, 226 (2021)