TECHNISCHE UNIVERSITÄT DARMSTADT

Master's Thesis (M.Sc.)
Deep Generative Design of IRES Elements for Therapeutic RNA

Department 18
Electrical Engineering and Information Technology
Self-Organizing Systems Lab

Prof. Dr. Heinz Koeppl
Head of lab

Philipp Froehlich and Tom Brown
Project supervisor

S3|06 206
Merckstrasse 25
64283 Darmstadt

philipp.froehlich@tu-darmstadt.de
thomas.brown@tu-darmstadt.de
https://www.bcs.tu-darmstadt.de

September 10, 2026

In collaboration with Merck

Background: How can we design RNA sequences that support efficient protein production? *Internal ribosome entry sites (IRES)* are regulatory RNA sequences that act as internal entry points for the cell's machinery to read protein-coding instructions. They can influence how efficiently a protein is produced from linear or circular RNA (circRNA), making them useful components for therapeutic RNA design. The project investigates how generative AI can design new IRES sequences with high predicted activity in a specific RNA context.

Objective: In this thesis, you will combine experimental activity data with generative sequence models. Building on the IRES-LM/IRES-DM framework introduced by Chu et al. [2026], you will develop a model that predicts IRES-driven protein expression in a given experimental context. You will then use these predictions to guide a *diffusion model* towards new IRES sequences with high predicted activity.

Key tasks:

- **Data preparation and evaluation:** Prepare experimental datasets for linear RNA and circRNA and establish a rigorous evaluation protocol using independent test data.
- **Predict context-specific activity:** Develop a deep learning model that predicts IRES activity while accounting for the relevant RNA context.
- **Guide sequence generation:** Use the activity predictions to guide sequence generation and compare the approach with existing methods.
- **Analyze candidate sequences:** Evaluate generated sequences with independent models and examine their diversity, similarity to training data, and characteristic sequence and structural features.

Expected outcomes: A reproducible design pipeline and an evaluated set of new IRES candidates to inform future experimental studies. You will also assess the reliability and limitations of model-based improvements. Conducted in collaboration with Merck, the project connects generative AI with RNA design.

Prerequisites:

- Good programming skills in Python; experience with PyTorch is helpful.
- A foundation in machine learning and deep learning.
- Interest in generative models and biological sequences, and enthusiasm for data-driven experimentation.

Nice to have (optional): Familiarity with RNA biology or bioinformatics.

For further information, please contact Philipp Froehlich and Tom Brown.

References

Yanyi Chu, Di Yin, Dan Yu, Guangxue Xu, Junze Zhang, Xiaotong Wang, Yue Shen, Yupeng Li, Ning Zhao, Yi Zhu, et al. Programmable RNA translation through deep learning-driven IRES discovery and de novo generation. *Nature Machine Intelligence*, 8(4):559–574, 2026.