Case Study

DATAHOW

Reducing Material Costs in mRNA In Vitro Transcription Without Compromising Quality

How DataHow's hybrid modeling identified optimal process conditions, transferred insights across mRNA sequences, and boosted efficiency while reducing raw material use.

Challenge

mRNA manufacturing offers a faster, fully synthetic path from design to product, eliminating the need for cell-based systems and facilitating process development. **While platform process formats robustly deliver high yields of good quality mRNA, many compounds are used in excess.** Cost of goods and in particular materials account for ~90% of the cost to manufacture mRNA at scale.

In vitro transcription (IVT), the core step where mRNA is synthesized, is one of the main contributors to material costs. Enzymes and capping agents used during IVT are expensive and in platform process conditions they are added at levels that allow to meet critical quality attributes (CQAs) and yields regardless of the mRNA sequence. However, this also means that for many mRNA sequences these expensive compounds are added in excess.

The challenge was to determine whether IVT conditions could be fine-tuned for different mRNA sequences to reduce excess and material costs, while maintaining yield and product quality.

Objective

OBJ. 1: Assess hybrid modeling capability: Determine whether DataHow's hybrid modeling approach can capture how compound concentrations and process conditions influence product quality and yield.

OBJ. 2: Evaluate knowledge transfer potential: Test if insights from one mRNA process can be transferred to another sequence using model-based transfer learning and a limited number of experimental runs.

OBJ. 3: Increase process efficiency: Use model-driven optimization to identify process conditions that maintain CQA specifications and maximize yield, while significantly reducing material costs.

Approach

DataHow collaborated with **Curevac SE** to explore what development effort would be required to reduce excess while still meeting CQAs and yield requirements when using DataHow's AI approach for two mRNA sequences.

To have enough experimental data to train a model and test its abilities, a design of 20 experimental runs was created, varying both compound concentrations and process parameters. CureVac SE executed these runs, capturing evolution data on intact and fragmented mRNA, along with endpoint CQA measurements.

Using varying number of runs from the dataset, **a hybrid model, combining mechanistic understanding and machine-learning, was trained to capture the relationship between inputs and outcomes.** A transfer learning method was then applied to assess how efficiently the model could be adapted to a new mRNA sequence.

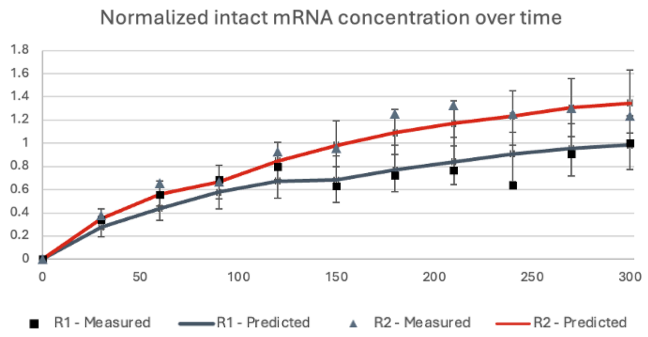
Finally, the optimized model was used to propose process conditions that would reduce material use while maintaining yield and CQA compliance.

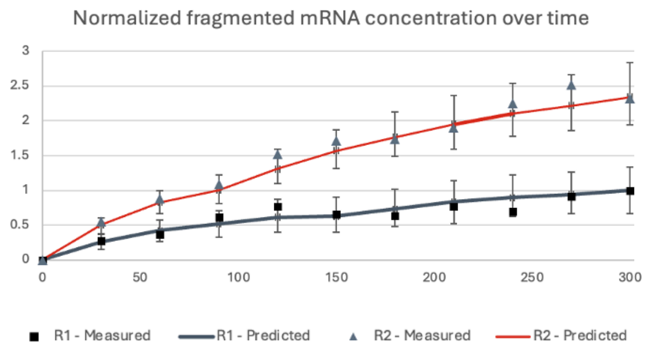
How Can Knowledge Be Transferred?

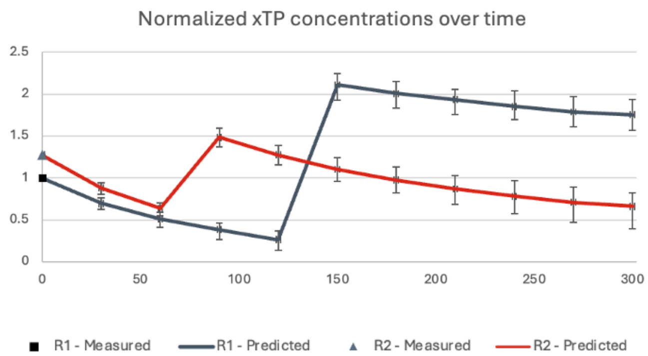
Knowledge transfer in bioprocessing is key to scaling innovation efficiently. In this study, DataHow explored multiple transfer learning strategies, including embeddings, meta-learning, and scaling approaches. These methods make it possible to **reuse knowledge about general process behavior across different mRNA sequences.**

In practice, this means that only **a limited number of new runs are required to adapt a model** to a new sequence, as the foundational process understanding remains valid. For example, while fragmentation behavior might vary with sequence, the overall process dynamics can be generalized.

1. Mechanistic Insights Without Extra Experiments

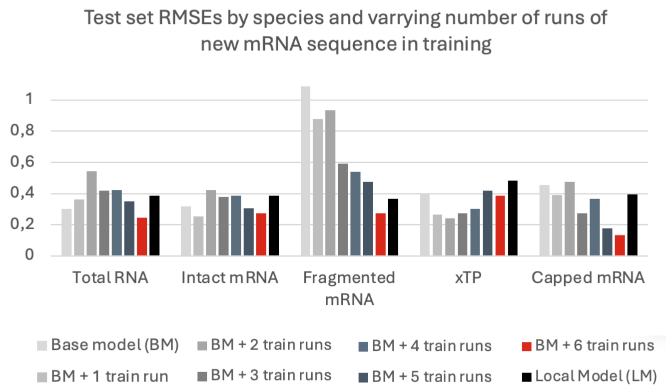






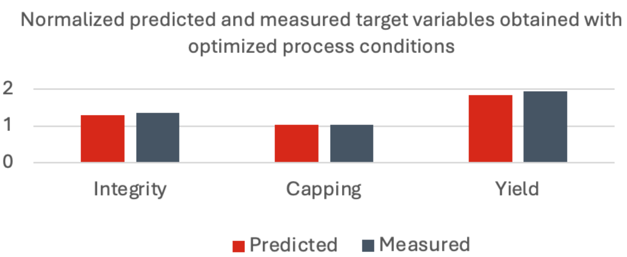
- The hybrid model accurately captures the impact of material and process parameter changes on process evolution.
- Process understanding related to xTP amounts and addition times can be easily generated without additional experimentation, as their **effects are mechanistically captured.**
- Different durations of the IVT process can easily be studied insilico, exploring for instance whether longer durations might give rise to better conversions or increased fragmentation.

2. Optimal Performance Achieved with Just Six Runs



- A clear improvement in model performance is observed when historical data is combined with data from the new process.
- On average, six runs appear to offer an optimal balance between performance and cost.
- The performance of variables influenced by changes in the mRNA sequence (such as fragmented or capped mRNA) improves when data from the new process is included.

3. Significant Reduction in Overall Raw Material Costs for IVT



- Maximizing yield while minimizing material costs and maintaining capped mRNA percentage and integrity above specified levels demonstrated that **simultaneous variations in process conditions and material additions can counterbalance potential decreases in individual variables.**
- Compared to the standard platform process, **the mass of one compound suffered a strongly significant reduction**, with improved process performance and minimal impact on quality attributes. This corresponds to a **significant reduction in overall raw material costs for the IVT step.**

Significant
Reduction in Overall Raw Material Costs for IVT

6
Runs to Achieve Optimal Performance

Strongly Significant
Mass Reduction in One Compound