How to design an mRNA vaccine or therapy

An mRNA typically encodes the following components:

Promoter. A promoter sequence is required to bind the phage RNA polymerases that initiate mRNA *in vitro* transcription. Typically, a T7 promoter is used, but other promoters for SP6 and T3 polymerases are also commonly used. If using co-transcriptional capping, a modified transcription initiation dinucleotide (AG) is often needed to favour incorporation of the 5' cap analog**.**

5' Untranslated Region (5'UTR). The 5' untranslated region regulates the initiation of translation by the ribosome. Most highly expressed genes and mRNA vaccines have short 5' UTRs with few secondary structures around the translation start site, and a consensus Kozak sequence preceding the start codon. For constitutive protein expression in mammalian cells, we use the alpha globin 5' UTR that facilitates strong expression in most cell types and animals.

Open Reading Frame (ORF). The ORF encodes the protein of interest. To optimise translation the protein, the ORF is often codon-optimised, with rare codons being replaced with synonymous codons often used by target organism, ensuring a reservoir of tRNAs is available to facilitate rapid translation. In addition, the mRNA sequence may be depleted of uridine content by preferential selection of codons with few uridines. Reduced uridine content can reduce activation of Toll-like receptors and reduce the reactogenicity of the mRNA[20].

3' Untranslated Region (3'UTR). The 3' UTR of native mRNAs often contains regulatory sequences that are bound by cellular proteins and regulate mRNA stability. The 3' UTR sequence may be modified to abolish microRNA binding sequences that reduce mRNA stability. 3'UTRs from highly expressed and stable mRNAs, such as alpha globin, are often used due to their short length and stability. To optimise mRNAs for constitutive expression in non-mammalian cell types (e.g., insect cell culture, plants, zebrafish), we recommend selecting short 5' and 3' UTR sequences from highly expressed and stable housekeeping genes. Another consideration for UTR selection is that cell-type-specific UTRs may confer selective expression in target cell types[39,40].

Poly(A) tail. The poly(A) tail plays a key role in mRNA translation, with longer poly(A) tails associated with improved mRNA stability. The use of alternative poly(A) tails, including segmented, modified, or branched sequences have also been described where they impart benefits on mRNA stability and expression.

Step-by-step example

BASE scientists use a software toolkit called *mRNArchitect* to design mRNA vaccines and therapies. This software uses an optimization strategy based on the *DNAChisel* framework to generate and assemble mRNA sequences.

1. Open the *mRNArchitect* website at http://www.basefacility.org.au/software

2. The Sequence Input panel allows you to input the sequences for the different components of an mRNA. Paste the wild-type Firefly luciferase protein sequence (in either nucleotide or amino acid nomenclature) into the *Coding Sequence* field.

3. Select the checkboxes for *Human alpha globin* in the *5'UTR* and *3'UTR* fields. For this protocol, it is not necessary to add a poly(A) tail, as this will be incorporated along with the T7 promoter during PCR amplification.

4. The *Parameters* panel allows the modification of key variables that impact mRNA sequence optimization. We recommend initially using the default parameters; however, these variables can be adjusted as needed for your application. Please refer to the HELP section in *mRNArchitect* for a detailed discussion of each parameter.

5. Select RUN to start the sequence optimization. When complete, the results will be available under the OUTPUT tab, where the optimized sequence(s) can be viewed and the results can be downloaded.

6. The optimized mRNA sequence can be copied and submitted for synthesis by a third-party provider (e.g., IDT, GeneArt, Genscript, etc.).