Synthetic mRNAs for manipulating cellular phenotypes: an overview

Elgar Susanne Quabius¹,² and Guido Krupp³

¹ Department of Otorhinolaryngology, Head and Neck Surgery, Christian-Albrechts-University Kiel, Arnold-Heller-Str. 3, Building 27, D-24105 Kiel, Germany
² Institute of Immunology, Christian-Albrechts-University Kiel, Arnold-Heller-Str. 3, Building 17, D-24105 Kiel, Germany
³ AmpTec GmbH, Königstr. 4A, 22767 Hamburg, Germany

Availability of high quality synthetic mRNAs (syn-mRNAs) has enabled progress in their applications. Important structural features and quality requirements are discussed. Developments in the application of mRNA-mediated manipulation of cells are presented (i) mRNA-directed expression of antigens in dendritic cells for vaccination projects in oncogenesis, infectious disease and allergy prevention; (ii) reprogramming of human fibroblasts to induced pluripotent stem cells with their subsequent differentiation to the desired cell type; (iii) applications in gene therapy.

Introduction

Conventional approaches for engineered changes in cellular expression profiles employ mostly DNA or RNA based viral and non-viral vectors. But these methods carry high risks, due to genomic integration with permanent genetic alteration of cells, and safety and ethical concerns have been raised against the use of DNA-based vectors in human clinical therapy. Employment of in vitro synthesized mRNAs is an advantage if these crucial permanent changes are not needed and the associated problems can be avoided in applications if transient gene expression changes are sufficient or even an advantage. Contrarily, concerns about RNA degradation problems have halted widespread use. Recent advances in the availability of synthetic mRNAs (syn-mRNAs) have increased confidence in working with syn-mRNAs and this overview presents important features of these molecules and their potential in multiple application fields.

The field is presented in two recent reviews: Bernal [1] presents the different DNA- and RNA-based technologies for cellular reprogramming and cellular lineage-conversion with their potential clinical applications, whereas Li et al. [2] focus on mRNA technology; in addition to the positive aspects, they also point out some crucial issues, such as selection pressure for cells that are deficient in innate immune response and possible accumulation of mutations in tumor suppressor genes.

Advantages of synthetic mRNAs (syn-mRNAs)

Although Wolff et al. [3] have shown already in 1990 that direct injection of ‘naked’ messenger RNA (mRNA) into the skeletal muscle of a mouse resulted in expression of the encoded protein, development of mRNA vaccines was considered unrealistic because of the expected mRNA instability during storage and after application in vivo. These concerns were compounded by difficulties in manufacturing at large scale and lack of commercial suppliers of high quality mRNA products. These concerns could be resolved in recent years and the two current major applications of syn-mRNAs with clinical potential are outlined in Fig. 1.

Highly specific and efficient vaccination can be achieved by the Wolff approach with injection of ‘naked mRNA’, but the alternative route via ex vivo transformation of isolated autologous dendritic cells (DC) offers multiple advantages [4–8] and it is outlined in Fig. 1A. The structural mRNA requirements for these applications are shown in Fig. 2A.

More recently, Warren et al. [9] have demonstrated the successful mRNA-mediated reprogramming of human fibroblasts to induced pluripotent stem cells (iPSC) and in turn, their mRNA-mediated differentiation to the desired mature cell type. This route is outlined in Fig. 1B and the more challenging structural mRNA
Requirements for these applications are shown in Fig. 2B. Furthermore, very stringent purification procedures by HPLC are advantageous, and the standard procedure with silica-based spin column purification is not always sufficient, possibly the removal of minor amounts of short double stranded RNA fragments is critically important for further reduced immune response, Kariko et al. [10]. Possibly, future projects based on this mRNA approach could provide additional power to the clinical application of iPSC, considering the already initiated clinical study of the treatment for age-related macular degeneration using iPSC, where a protein-based route is used, Cyranoski [11].

In essence: activity and clearance of syn-mRNAs provide all advantages of conventional pharmaceutical drugs, but avoid biodegradation and other environmental problems of synthetic chemical drugs because syn-mRNAs do not contain non-biological structures.

However, for cell biology laboratories, the generation of syn-mRNA remains a severe limitation, especially if many different constructs are needed in small scales at early stages of a project. Another challenge comes later, if bigger amounts with high and highly reproducible quality are needed or if regulatory requirements demand standardized and controlled manufacturing processes in compliance with cGMP principles. Commercial suppliers who meet these requirements are rare, GMP-grade mRNAs are offered by two established companies [4]: Asuragen in the USA and CureVac in Europe. A flexible small company like AmpTec GmbH could be beneficial by offering all manufacturing options, combined with flexible services for multiple sequences in small amounts at early project stages and for large amounts at later stages.

**Design of syn-mRNAs**

Syn-mRNAs present the advantage that their production can be engineered to the required scale, and once the technology is established it works, in principle, for all possible RNA sequences, and there are essentially no size limitations. RNA sequences can be adapted to the specific needs, concerning stability and translation efficiency, which can be controlled or optimized by cis-acting structural elements in the RNA: the 5'-cap (structural and modified alternatives are commercially available), the poly(A)-tail (defined length and absence of 3'-terminal ‘masking’ sequences are possible), untranslated regions and the sequence of the coding region can be optimized for the cellular background, and if needed, nucleotide modifications can be inserted, Warren et al. [9]. For discussions of the general design aspects we refer to several recent reviews [5–8]. Requirements in the syn-mRNA structures and quality control aspects are discussed in more detail in section ‘Structural and quality requirements for syn-mRNAs’.

**Syn-mRNAs and cellular immune response: an advantage or a problem?**

Syn-mRNAs are recognized as non-self nucleic acids leading to adaptable activation of the innate immune system [12]. Non-self RNAs naturally originate from invading viruses or other cellular pathogens and encode non-cellular genes. This means, differentiation between self and non-self RNAs are fundamental cellular protective mechanisms against infections. Cells of the innate immune system possess multiple sensory receptors that are activated by RNA, including Toll-like receptors (TLR) 3, 7, and 8, as well as RIG-I-(retinoic acid-inducible gene I)-like receptors (RLR). They differ in their cellular localization and ligand specifications, and induce various complex signaling cascades that finally lead to an antiviral state [13,14]. This state means a translational shutdown, an up-regulation and production of interferons (IFNs) and/or cell death [15]. These sensors detect the externally introduced, in vitro transcribed mRNA-based therapeutics, and consequently induce an analogous antiviral state [16].

Dependent on the therapeutic or scientific application, this immune response can be beneficial or detrimental!

**Innate immune response as an advantage**

This immune response is beneficial in applications of mRNA-based vaccinations, because it potentiates the immune response of the vaccine itself. The syn-mRNA does not only encode the antigen but it also acts as an adjuvant by enhancing immunological responses and antigen presentation. Therefore, the use of syn-mRNA is a rapidly growing field for vaccination strategies in the battle against viral infections and cancer.

Several clinical applications based on syn-mRNA vaccination are under investigation with a focus on DC [17,18], DCs have a
FIGURE 2

General structural requirements for syn-mRNAs. (A1) Desired structure for syn-mRNA for vaccination: 5'-cap (the ARCA analogue) and a long, free 3'-poly(A), and no internal nucleotide modifications. (A2) Undesired features are crossed out: 5'-terminal phosphates and a masked 3'-terminal poly(A) with extra nucleotides, indicate as V (=G, C or U). (A3) undesired, but accepted 5'-OH, resulting from phosphatase treatment. (B1) Desired structure for syn-mRNA for cellular reprogramming: 5'-cap (the ARCA analogue) and a long, free 3'-poly(A), and 100% replacement of all pyrimidines. Please note that other pyrimidine modifications and modification levels can be used, see sections ‘Special highly modified mRNA constructs may overcome immunogenicity problems in cell culture systems’ and ‘Additional challenges in systemic gene therapy applications’. (B2) as B1, but with accepted 5'-OH, resulting from phosphatase treatment.
central function in presenting antigens through major histocompatibility complex (MHC) class I and II proteins and in eliciting both cellular and humoral immune responses. Immune responses can be induced by loading DCs with proteins, cDNAs or mRNAs. Although whole cell protein extracts can be used, the practicability is limited by risks of including tumorigenic compounds and the presence of many irrelevant antigens can cause autoimmune responses. Using specific proteins is challenging. Therefore, nucleic acid vaccinations present an interesting alternative. Furthermore, genetic vaccinations against infectious diseases eliminate the risk of mutation and uncontrollable proliferation of inactivated pathogens. In principle, DNA can be taken up and expressed by cells in vivo (or in vitro), its use as nucleic acid vaccine has disadvantages, as mentioned before: DNA can integrate into the host genome, causing inactivation of cellular genes or oncogenesis. As a further disadvantage in this application, DNA results in a long duration of expression of immunizing antigens, whereas it has been demonstrated that optimal vaccination can be achieved by a boost in antigen expression, as obtained by vaccination with mRNA [19]. As outlined in Fig. 1A, the ex vivo route of syn-mRNA-based immunotherapy starts with transfection of (mostly autologous) human DCs.

Although most studies using syn-mRNA vaccination strategy are directed towards immunotherapeutic treatment of different cancers, already in 1993 Martinon et al. have demonstrated the potential of a liposome-entrapped mRNA vaccine against influenza in a mouse model [20]. An extensive and updated study was published in 2012 by Petsch et al. [6]. Also studies with HIV-1-infected patients were initiated [21,22].

A helpful application could emerge with ‘mRNA vaccination as a safe approach for specific protection from type I allergy’ [23] using a broad panel of allergen-encoding syn-mRNA vaccines which elicit long-lasting protection from sensitization, and induce a type of immunity similar to the natural protective response that is acquired in the presence of allergen burden early in life.

**Innate immune response as a problem**

Apart from vaccination, induced innate immune effects are detrimental for most other applications. For the production of high amounts of recombinant proteins, translational shutdown is obviously counterproductive, and repetitive RNA transfer for improved productivity is obviously impossible after cell death.

However, low transformation efficiencies necessitate multiple RNA transfer steps for successful conversion of one cell-type to another, for example, in the generation of iPSC.

In scientific studies which address signaling pathways, the secretion of IFN will trigger signaling cascades that can cause a mix-up with the studied mRNA-related effects. In these cases, overcoming the hurdles of immunogenic side effects is crucial for success.

**Special highly modified mRNA constructs may overcome immunogenicity problems in cell culture systems**

It is frequently overlooked that eukaryotic mRNAs are extensively modified in vivo, and it has been shown that mRNA containing base-modified nucleosides lead to reduced signaling by the ssRNA sensor RIG-I, and reduced induction of protein kinase R (PKR), a global repressor of protein translation, and of the endosomal ssRNA sensors TLR7 and TLR8 [24–27].

In addition, it is known that 5'-terminal triphosphates in mRNA transcripts can trigger PKR [26] and RIG-I [28,29]. The standard synthetic IVT-RNA products have a 5'-triphosphate. Whereas the 5’ ends of cellular mRNAs are modified post-transcriptionally in the nucleus with a methylated m7GpppN cap structure which is important for mRNA splicing, stabilization, transport and for efficient protein synthesis by recruiting ribosomes. In essence, the cap structure is mandatory for normal mRNA function. Accordingly, also syn-mRNAs require a 5’-cap structure. Although post-transcriptional capping in vitro is a possibility for achieving very high capping efficiencies, the widely used strategy relies on co-transcriptional capping by the addition of excess amounts of synthetic cap structures. However, the chain-elongating GTP must be present and it competes with the cap structure in the initiation step, accordingly it is impossible to prevent that a significant fraction of uncapped in vitro transcription (IVT) products with 5’-triphosphates remains.

Recently, Warren et al. [9] addressed both aspects: (i) generating highly modified mRNAs with complete replacement of both pyrimidines, m5SC instead of C and Ψ (Pseudo-U) instead of U; and (ii) removing remaining 5’-triphosphates by treating the syn-mRNAs with a phosphatase. Furthermore, they used polymerase chain reaction (PCR) products as IVT templates which encode a long, free 3’-terminal poly(A) tail with A120. A very detailed presentation of the technical procedures was provided by Mandal and Rossi [30], including the identification of crucial steps and a trouble shooting guide, see also below, section ‘Quality control aspects for syn-mRNAs are illustrated in the following list for syn-mRNAs from AmpTec’ (sections ‘Molecular purity’ and ‘Bioburden’).

From a biophysics perspective, the high frequency introduction of 100% Ψ has effects on the properties of these highly modified syn-mRNAs. For example, one single Ψ at a crucial position in a tRNA anticodon has an important consequence: the tRNAΨA with its GΨA anticodon does not only recognize the Tyr codons UAC and UAU, but in addition, it can act as a natural suppressor of the UAG stop codon, possibly due to a more rigid local RNA structure and the stabilized central Ψ:A base pair [31].

Although these highly modified syn-mRNAs dramatically attenuated interferon signaling, residual upregulation of some interferon targets was still detected. The additional media supplementation with recombinant B18R protein, a Vaccinia virus decay receptor for type I interferons, allowed for high, dose-dependent levels of protein expression with high cell viability.

With the combination of RNA modifications with the B18R interferon inhibitor, Warren et al. [9] have developed a technology to overcome innate antiviral responses, and this enables highly efficient reprogramming of somatic cells to pluripotency and subsequently to direct the differentiation of pluripotent cells towards a desired cell type.

It is important to note that modified syn-mRNA-derived iPSC clones ‘more faithfully recapitulated the global transcriptional signature of human ESCs than retrovirally derived iPSCs’. In consequence, this route ‘may produce higher-quality iPSCs, possibly owing to the fact that they are transgene free’.

In addition to ‘conventional’ mRNAs that encode only one protein, Yoshiiko et al. [32] present a different approach by utilizing a single noninfectious (nonpackaging), self-replicating Venezuelan equine encephalitis (VEE) virus RNA replicon that is
engineered to express four reprogramming factors (OCT4, KLF4, and SOX2, with c-MYC or GLIS1). The required very long transcriptions (up to 15,000 nt) were obtained with standard commercial IVT kits. No internal nucleoside modifications were introduced and 5'-capping was achieved post-transcriptionally by using a high-efficiency enzymatic capping system. Expression of this unmodified VEE-RNA was achieved but it was critically dependent on high levels of the interferon inhibitor B18R.

As shown by Kariko et al. [10], dependence on the interferon inhibitor supplement is at least reduced, if very stringent purification procedures by HPLC are applied as an additional, final syn-mRNA production step. Possibly, the removal of minor amounts of short double stranded RNA fragments is critically important to achieve this further reduced immune response.

Interesting to note: although 100% replacements of uridine byΨ and cytidine by m5C were used by Warren et al. [9] for cellular reprogramming, Yoshioko et al. [32] have used reduced levels with only 25% of Ψ and 25% of m5C.

**Additional challenges in systemic gene therapy applications**

Although the combination of RNA modifications with the B18R interferon inhibitor is a viable possibility in cell culture systems, this requirement could be a severe limitation for emerging systemic applications in gene therapy and for engineered protein expression changes in cellular targets [33].

In mouse model systems this limitation was not evident and potential therapeutic applications for syn-mRNA in vivo could be demonstrated. Rudolph and Kormann [34,35] have used a mouse model for a lethal congenital lung disease that lacks surfactant protein B (SP-B) expression in the pulmonary epithelium. Local application of an aerosol with SP-B syn-mRNA resulted in high-level SP-B expression and survival of the treated animals.

In addition to this localized use, systemic expression of biologically active erythropoietin protein in mice was achieved by injection of erythropoietin-encoding syn-mRNA into the *tibalis anterior*, Kormann et al. [34], or into the peritoneal cavity, Kariko et al. [36].

Important note for syn-mRNA constructs: 100% replacements of uridine by Ψ and cytidine by m5C as used by Warren et al. [9] for cellular reprogramming are not always required: As already noted above (section ‘Special highly modified mRNA constructs may overcome immunogenicity problems in cell culture systems’), Yoshioko et al. [32] have used reduced levels with only 25% of Ψ and m5C; Kariko et al. [36] maintained 100% level of Ψ in a combination with unmodified cytidine; Rudolph and Kormann [34,35] have even shown potential advantages of using lower, 25% modification levels with a combination of s2U and m5C.

**Structural and quality requirements for syn-mRNAs**

Optimized syn-mRNA constructs combine these structural features (see also Fig. 2):

**5'-End**

Standard IVT-RNAs carry a 5'-triphosphate structure, whereas cellular mRNAs have a 5'-cap structure. In IVT reactions, the cap structure can be added by an efficient post-transcriptional enzymatic step, although capping is frequently introduced co-transcriptionally by including a free cap structure GpppG which competes with the standard nucleoside triphosphate pppG in the transcription initiation step. But the natural cap includes a methylated G, with the structure m7GpppG. Its incorporation in IVT products will lead to two different orientations: m7GpppGpN. . . and Gpppm7GpN . . ., and only 50% of the IVT product carry the first, biologically active orientation. Several options are possible to block the undesired elongation at m7G, with the Anti-Reverse Copy Analogue (ARCA) as the most popular one. Here the 3'-OH group next to m7G is replaced by –OCH3, which prevents elongation and directs IVT initiation to the desired opposite side. The ARCA structure can be abbreviated as m7G(m)pppG (commercially available for example from life technologies, NE Biosials, Trilink and Jena Bioscience). But incorporation is usually only about 80% and any remaining 5'-triphosphates have to be carefully eliminated by a phosphatase treatment of the purified IVT product.

**3'-End**

Stable, fully functional cellular mRNAs have a long, free 3'-terminal poly(A) tail. A designed plasmid template for IVT can include a long homopolymeric stretch of A’s but the maintenance of its exact length during amplification in *Escherichia coli* cells is difficult to control. In addition, the generation of IVT templates by cleaving at a standard restriction site following the poly(A) stretch, leaves several non-A nucleotides, which in effect are masking the poly(A) tail and intracellular mRNA efficiency can be ~two-fold reduced [5]. A possible bypass is the use of a class III restriction enzyme where cleavage site and recognition sequence are separated [5], but their shorter recognition sequences can be problematic if the same sequence is present within the mRNA sequence. AmpTec takes a different approach for IVT template synthesis, by generating PCR products which span the plasmid insert, as outlined in Fig. 3. These IVT templates avoid both problems, the length of the 5'-terminal homopolymeric T-stretch in the chemically synthesized PCR primer is well defined, the PCR amplification process is highly reproducible and there is no interference with internal mRNA sequences. As indicated in Fig. 2, the standard AmpTec primers introduce a 3'-terminal poly(A) tail with 120 A’s.

**Internal sequence**

Requirements for 5'- and 3'-terminal structures are the same for all mRNA applications, but for avoiding the innate immune response, internal pyrimidines have to be replaced. A substitution of CTP by m5CTP and UTP by ΨTP (Pseudo-UTP) can achieve this goal. In general, the modified NTPs are efficient substrates for T7 RNA polymerase, comparable with their unmodified counterparts; aberrant premature termination can be avoided and full-size IVT products are obtained (example in Fig. 4).

Although 100% replacement of uridine by Ψ and cytidine by m5C in mRNAs results in the described advantages, Rudolph and Kormann [34,35] have shown potential benefits of using lower modification levels.

**Quality control aspects for syn-mRNAs are illustrated in the following list for syn-mRNAs from AmpTec**

**Identity**

Syn-mRNA sequences can be freely designed by the user, and they are generated by the assembly of chemically synthesized sequence blocks, which are inserted into a plasmid backbone and propagated
by cloning in *E. coli*. The finally selected plasmid clone is fully sequenced and must have 100% identity with the expected sequence.

To generate a template for IVT reactions with T7 RNA polymerase, plasmid DNA is converted to an insert-specific PCR product which includes a long terminal stretch of 120 T's. Two options are available for 5'-capping: Post-transcriptional addition with a capping enzyme or co-transcriptional by including an excess of a cap analogue, which competes with GTP, the standard IVT initiator. With about 80% efficiency in co-transcriptional capping, any remaining 5'-terminal triphosphates must be removed by a phosphatase treatment, subsequent to the spin column purification of IVT products.

Capillary electrophoresis (Agilent Bioanalyzer 2100 or TapeStation 2200) is used to verify the correct mRNA size and prove that the expected syn-mRNA is the major product. Optionally, susceptibility to RNase can be used as additional verification of molecular identity.

 Routinely, users can select a set of important gene sequences which must be absent in the syn-mRNA product. With the use of capillary electrophoresis as analytical tool, the expected sizes of undesired RNAs must be significantly different from the syn-mRNA product, but the limits of detection are only in the % range. AmpTec employs reverse transcription polymerase chain reaction (RT-PCR) as an additional analytical approach to avoid dependence on size differences and for increased sensitivity: gene-specific RT-PCR assays for all user-selected sequences are combined with capillary electrophoresis (Agilent) of the resulting RT-PCR products. In this manner, detection sensitivities in the parts per million (ppm) range are achieved. These assays are performed at two stages in the manufacturing process: as PCR with the starting material, plasmid DNA, and as RT-PCR with the final product, syn-mRNA.

**Quantification**

Standard measurements by UV absorbance at 260 nm are performed in a defined buffer (10 mM TE buffer, pH 8). Variability due to hypochromicity (random effects in the formation of aberrant RNA conformations with double-stranded regions) is reduced by a heat renaturation step. The combination with capillary electrophoresis (Agilent) provides an independent quantification method.

As a special option, quantification by HPLC is offered: the RNA is completely digested with nuclease P1 and the resulting nucleoside monophosphates are analyzed. In addition to an absolute RNA quantification, this provides quantitative base composition ratios and enables a comparison with the sequence-predicted values; furthermore, the quantitative insertion of modified nucleotides can be monitored.

**Molecular purity**

Residual proteins, bacterial genomic DNA, bacterial RNA, and endotoxin must be below specified limits. The work-flow of AmpTec already includes the use of minimal amounts of bacterial-derived...
materials and the addition of plasmid DNA in IVT reactions is about a million-fold lower than in protocols where plasmid DNA is directly used as IVT template. This level of plasmid-derived material is about 100-fold lower, as compared with the work-flow described by Mandal and Rossi [30]. In addition to reducing carry-over of bacterial compounds, this also eliminates the need of using a linearized plasmid as template for Heel-PCR.

The presence of residual template DNA is checked. As explained above, AmpTec employs RT-PCR assays, combined with capillary electrophoresis (Agilent) of PCR products. An additional Minus-RT control can detect residual template DNA with a sensitivity in the ppm range.

**Further option:** In addition to analytic HPLC (see section ‘Quantification’), HPLC can be applied as a final purification step to achieve higher purity in general, and for increased sample homogeneity by removing undesired shorter or longer IVT products.

**References**


**Bioburden**

Endotoxin-level is monitored and sterility is controlled by standard microbiological assays (aerobic, anaerobic, fungal). This aspect was only indirectly addressed in the detailed work-flow of Mandal and Rossi [30] as potential problem of ‘RNA toxicity’ in their troubleshooting section.

**Concluding remarks and outlook**

This short overview has presented encouraging advancements in the generation of syn-mRNA molecules, paving the way for the now rapidly growing applications.

Challenges remain in the further optimization of the syn-mRNA construct itself and of standardized, efficient methods to obtain stable syn-mRNA preparations, and in developing novel formulations to achieve efficient cell entry and intracellular activity, ideally with the option to engineer different versions with selectivity/specificity for different cell types.