Design of in vitro Transcribed mRNA Vectors for Research and Therapy

Tusup, Marina; French, Lars E; De Matos, Mara; Gatfield, David; Kundig, Thomas; Pascolo, Steve

Abstract: The use of in vitro transcribed messenger RNA (ivt mRNA) for vaccination, gene therapy and cell reprogramming has become increasingly popular in research and medicine. This method can be used in vitro (transfected in cells) or administered naked or formulated (lipoplexes, polyplexes, and lipopolylexes that deliver the RNA to specific organs, such as immune structures, the lung or liver) and is designed to be an immunostimulatory or immunosilent agent. This vector contains several functional regions (Cap, 5’ untranslated region, open reading frame, 3’ untranslated region and poly-A tail) that can all be optimised to generate a highly efficacious ivt mRNA. In this study, we review these aspects and report on the effect of the ivt mRNA purification method on the functionality of this synthetic transient genetic vector.

DOI: https://doi.org/10.2533/chimia.2019.391

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-173348
Journal Article
Published Version

Originally published at:
Tusup, Marina; French, Lars E; De Matos, Mara; Gatfield, David; Kundig, Thomas; Pascolo, Steve (2019). Design of in vitro Transcribed mRNA Vectors for Research and Therapy. CHIMIA International Journal for Chemistry, 73(5):391-394.
DOI: https://doi.org/10.2533/chimia.2019.391
Design of in vitro Transcribed mRNA Vectors for Research and Therapy

Marina Tusupab, Lars E. Frenchab, Mara De Matosc, David Gatfieldc, Thomas Kundig1, and Steve Pascoloa,b

Abstract: The use of in vitro transcribed messenger RNA (ivt mRNA) for vaccination, gene therapy and cell reprogramming has become increasingly popular in research and medicine. This method can be used in vitro (transfected in cells) or administered naked or formulated (lipoplexes, polyplexes, and lipopolyplexes that deliver the RNA to specific organs, such as immune structures, the lung or liver) and is designed to be an immunostimulatory or immunosilent agent. This vector contains several functional regions (Cap, 5′ untranslated region, open reading frame, 3′ untranslated region and poly-A tail) that can all be optimised to generate a highly efficacious ivt mRNA. In this study, we review these aspects and report on the effect of the ivt mRNA purification method on the functionality of this synthetic transient genetic vector.

Keywords: Cap · Gene therapy · Globin UTR · in vitro transcribed mRNA · ivt RNA · Vaccination

Steve Pascolo studied biology/biochemistry at Ecole Normale Supérieure (Paris, France) and obtained his PhD in 1998 from the Pasteur Institute / University Paris 6, where he created an immunologically humanised mouse model called HHD (mice deficient in the endogenous H-2 class I molecules and transgenic for a human HLA class I molecule). As a post-doctoral fellow at the University of Tübingen, Germany, he focused on immunisation methods using WT and HHD mice, testing different vaccination formats (e.g. peptides, plasmids, and ivt mRNAs), and discovered that RNA is a ‘danger signal’ that primes innate and adaptive immune responses. With the goal of translating this vaccination method to humans, Steve Pascolo co-founded CureVac GmbH in 2000, implemented the first large scale and GMP production of ivt mRNAs worldwide and initiated the first clinical studies in which patients with cancer received direct ivt mRNA injections. In 2006, Steve Pascolo moved to the University Hospital of Zurich (Switzerland) and continued to develop innovative immunotherapies against cancer (vaccination, immunomodulation, and chemotherapies). In 2017, the University of Zurich supported the creation of an mRNA platform. This platform allows researchers and clinicians to have access to high quality ivt mRNAs produced on demand for their research and pre-clinical projects.

1. Introduction

Vaccination (induction of an immune response against a protein) and gene therapy (expression of a therapeutic protein) can be performed using genetic vectors, such as plasmid DNA, in vitro transcribed mRNAs (ivt mRNA) or recombinant microorganisms (modified viruses or bacteria). Because of this technology’s superlative features (safety, efficacy, steerable immunostimulatory properties, and ease of production and preservation), ivt mRNA is emerging as the optimal vector format.[1,2] We have initiated the clinical use of directly injected ivt mRNAs through academic (University of Tübingen, Germany) and industrial (CureVac AG, Germany) co-developments. As a key milestone, we implemented the first large scale GMP (good manufacturing practice) production of an ivt mRNA (CureVac AG, Germany) worldwide[3] and led the first clinical studies in which patients with cancer were injected with ivt mRNAs.[3-6] This pioneering work has revealed for the first time that an ivt mRNA can be produced at a large scale in injectable quality, and that its direct administration in humans is safe and leads to the expected pharmacodynamics (expression of the encoded protein[6] and development of an immune response[7]). The most recently developed ivt mRNA-based approaches were reported to induce very strong immune responses in patients with cancer.[8,9] Key to this notably high efficacy are the formulation of the ivt mRNA, the site of injection, as well as the structure of ivt mRNA.[10] The present review will focus on this latter aspect.

2. Structure of the ivt mRNA

2.1 The 5′ Cap

Eukaryotic mRNAs contain a 5′ Cap called Cap 0 that consists of a methyl-7 guanine nucleotide connected to the RNA via a 5′ to 5′ triphosphate linkage. If the first base of the mRNA is methylated at the 2′ position, the Cap is called Cap 1. It is the optimal structure to stabilise the mRNA and recruit the eukaryotic initiation factor 4E (eIF4E) that binds to eIF4G and facilitates the subsequent assembly of the eIF4F complex. Alternatives to the 5′ Cap for the recruitment of the translation machinery are Cap-independent viral sequences, such as the 5′ untranslated region of the tobacco etch virus[11] or internal ribosomal entry site (IRES) sequences.[12] However, such sequences at the 5′ end of an ivt mRNA are not sufficient to achieve high levels of protein expression after transfection of the RNA in cells. Thus, the production of a 5′ capped ivt mRNA is recommended. Two methods for capping have been reported: (i) co-transcriptional capping, where a Cap analogue (or an anti-reverse Cap analogue ‘ARCA’ eventually optimised[13]) or a Cap dinucleotide ‘CleanCap®’ from Trilink) is present in the transcription reaction, and (ii) post-transcriptional capping, where the transcribed mRNA is capped using enzymes such as the vaccinia virus capping enzyme. In all cases, a Cap
0 or Cap 1 can be generated (the latter requires the utilisation of a Cap 2'-O-methyltransferase in addition to the vaccinia virus capping enzyme in the context of enzymatic capping). Based on our experience, the most functional ivt mRNA was produced by co-transcriptional capping using the CleanCap® Reagent AG (3'Ome) from Trilink (m7(3'OmeG)(5')ppp(5')(2'OmeA)pG). It has the dual advantages of being an anti-reverse Cap 1 and of not competing with GTP to initiate transcription. Thus, the GTP concentration in the transcription reaction can be maintained at high levels without reducing the percentage of capped mRNA (using usual Cap mononucleotide analogues, the GTP concentration must be reduced to favour initiation by Cap and not by G). The method's disadvantage is that the DNA template must contain a modified T7 promoter, where the C that is used to base pair with the first nucleotide of the ivt mRNA (usually a G) is a T (allowing incorporation of the A of the CleanCap® Reagent AG). Thus, new dedicated DNA constructs must be produced to use the CleanCap® Reagent AG.

2.2 The 5' UTR

Located between the 5' Cap and the first ATG codon (Fig. 1), the 5' UTR sequence must provide a good Kozak consensus sequence that favours initiation of the translation process at the expected ATG. An 80 base sequence from the globin 5'UTR is frequently used.[14] We recently reported the superior efficacy of a short sequence of 40 nucleotides corresponding to an aptamer that binds eIF4G.[15] The rationale was that in some cell types, the availability of eIF4E is a limiting factor for translation; therefore, the direct recruitment of eIF4G to the 5'UTR would likely promote translation. We now typically produce ivt mRNAs with this 5' aptamer sequence, as follows: CapAG - G A C T C A C T A T T T G T T T T T C G C G C C A G T T G C A A A A A G T G T C G C C A G C A T G (italicized bases represent the eIF4G aptamer and underlined bases represent the start codon).

Fig. 1. Schematic representation of optimised ivt mRNA. A 5' Cap (optimally a methyl-7 guanine nucleotide connected via a 5' to 5' triphosphate linkage to a 2'-O-methyl modified residue: Cap') is followed by a 5'UTR (optimally an aptamer that binds eIF4G: 'Apt') and an open reading frame (optimally a codon optimised sequence depicted in yellow) between a start (AUG) and a stop codon. The 3' end of the mRNA can contain a stabilising 3'UTR (for example a tandem repeat of the beta-globin 3'UTR 'i') and a poly-A tail (ideally of more than 90 residues).

2.3 Coding Sequence

Depending on the species, codon usage can vary and reflects the relative abundance of tRNAs; efficient translation relies on optimal codons that correspond to abundant tRNA species.[16] Rare codons in an open reading frame hinder the production of full-length proteins. These codons may also decrease the stability of the mRNA, as slow-moving ribosomes may trigger mRNA decay. Therefore, ivt mRNAs are typically produced from synthetic DNA templates that have been 'codon optimised'. However, the natural sequence of a gene may have evolved to favour proper protein production, particularly folding, by controlling the speed of translation; rare codons are used to slow translation between structural domains allowing time for each domain to fold properly. Therefore, the replacement of suboptimal codons with optimal codons may increase protein misfolding. Indeed, we observed that some codon-optimised ivt mRNAs encoding secreted proteins were less functional than ivt mRNAs with the WT sequences (data not shown). Notably, multiple software packages are available for codon optimisation that incorporate several features (e.g., predicted secondary RNA structure and avoiding consecutive repetition of identical codons). Thus, we advise codon optimisation for the production of mRNAs encoding non-mammalian proteins (optionally using two independent companies/software packages and comparing the functionality of the resulting ivt mRNAs), while, a priori, the wild-type coding sequence should be used to produce mRNAs coding mammalian proteins (a subsequent test of codon-optimised forms can be envisioned).

2.4 The 3'UTR and Poly-A Tail

In ivt mRNAs, the most frequently used 3'UTR is the globin 3'UTR.[14,17–20] An optimised version of this 3'UTR is a tandem repeat of the sequence,[21] which contains then approximately 300 bases. Because the presence of repeated sequences may hamper the synthesis of the synthetic DNA that is used as a matrix for transcription, Orlandini von Niessen et al. recently reported the identification of new stabilising 3'UTRs.[22] The researchers started from a library of long-lived mRNAs (derived from dendritic cells treated with block transcribed DNA for a few hours to D for 24 hours) and cloned fragments of those reverse transcribed molecules at the 3' end of an ORF encoding destabilised eGFP. This library was then transcribed into mRNAs, transfected into dendritic cells, and RNA was extracted at different time points, and reverse transcription followed by PCR amplification was performed. Using an iterative process of transfection/RNA extraction/amplification/transcription, optimal 3'UTR sequences were identified. The best sequences corresponded to UTRs from mRNAs with long half-lives (this result validated the approach). The combination of two of those sequences in the 3'UTR, one from mitochondrially encoded 12S rRNA 'mtmRNA1' and one from the amino-terminal enhancer of split mRNA 'AES', produced an ivt mRNA with the highest expression (the combined sequences produced a UTR of less than 300 bases). This mRNA's expression in vitro (several cell types) and in vivo (mice) was greater than two-fold higher than the expression of the usual ivt mRNA with a tandem repeat of the globin 3'UTR. The poly-A tail can be encoded by the DNA matrix directly such that the ivt mRNA contains a 3' stretch of at least 90 A bases. Importantly, extra C, G or U bases at the 3' end of the poly-A sequence should be avoided.[21] If the DNA template does not encode this 3'poly-A sequence, an incubation of the ivt mRNA in a reaction with a poly-A polymerase post-transcriptionally generates the poly-A tail.

2.5 Purification Methods

The RNA polymerases that are usually used to produce ivt mRNAs (T3, T7, and SP6) are very processive and specific for their promoters. However, ivt mRNAs visualised on a gel[13] or with an Agilent analysis may show some aberrant shorter or longer transcripts. Those transcripts might be derived from cryptic promoters or cryptic terminators of transcription (e.g., highly structured RNA regions). Therefore, we have implemented a dedicated HPLC-based purification method for large-scale GMP production.[6] This method allows identifying and isolating the mRNA with the correct size, hence removing potential by-products that are longer or shorter. In addition, the RNA polymerase may generate hairpins at the 3' end by synthesizing a complementary RNA sequence. A newly available T7 RNA polymerase that is optimised to generate RNA at 50 °C (Hi-T7™ RNA Polymerase from NEB Biolabs) was reported to reduce dsRNA by-product formation. We also noticed that this Hi-T7™ can produce full-length mRNA from hard-to-transcribe DNA templates (e.g., containing a GC-rich 5' region, Fig. 2). Meanwhile, for research-grade ivt mRNAs, usually the classical production using wild type polymerases and simple purification of ivt mRNAs using selective precipitation (LiCl that precipitates only long RNA molecules, not oligonucleotides, DNA or proteins) or columns (silica membranes) is sufficient.
comparison of both purification methods (Fig. 3) revealed that they are equivalent and present similar efficacy in terms of the amount of recovered RNA, OD260/OD280 ratio (ca. 2), of the observed transfection efficiencies (Fig. 3A) and reduced immunogenicity (for a completely 1-methyl-pseudoUridine -substituted mRNA, Fig. 3B). Thus, LiCl does not exert any deleterious effects on ivt mRNAs that impair their subsequent applications, and it is a simple, inexpensive and scalable method for purifying mRNAs from in vitro transcription reactions.

3. Conclusions

The production of ivt mRNAs for research and clinical interventions is easy and inexpensive. However, because this technology has many facets, including the structure of the ivt mRNA and its formulation, substantial expertise is helpful for producing optimal ivt mRNAs. To this end, the implementation of a centralised ivt mRNA platform, which we have used at the University of Zurich (https://www.cancer.uzh.ch/en/Research/mRNA-Platform.html), represents a significant asset for researchers and clinicians aiming to use ivt mRNAs for research or clinical interventions. Investigator-driven clinical trials have been initiated based on findings from academic studies, in particular for vaccinations that we have pioneered\(^{3–5,19,20,23,24}\), and for modified T-cells.\(^{25,26}\) Meanwhile, several dedicated biotechnology companies have initiated human studies, including BioNTech AG, CureVac AG, eTheRNA, Moderna Inc. Ongoing company-sponsored clinical trials include vaccinations (e.g., vaccines against solid cancers and viruses, such as rabies, HPV16 and HIV), and non-immunogenic protein expression (e.g., ivt mRNAs encoding cytokines that are injected into the tumour). Furthermore, the versatility of the ivt mRNA format allows the design of individualized therapeutics; mRNA vaccines encoding mutated peptides corresponding to patient-specific cancer mutations have been designed, produced and administered.\(^9\) In addition, ivt mRNAs can be combined with

![Fig. 2. Transcription of difficult DNA matrices. Transcriptions were performed using three PCR matrices with a GC content of 78% in the first 161 bases. The produced mRNAs were analysed on a denaturing Formaldehyde MOPS agarose gel. The expected size was over ca. 2500 bases. The canonical T7 RNA polymerase ('T7') at 37 °C produces abundant shorter transcripts while the Hi-T7 RNA Polymerase ('Hi-T7') at 50 °C generates mostly full-length transcripts.](image)

![Fig. 3. Effects of the purification method on the functionality of ivt mRNAs. A completely 1-methyl-pseudoUridine -substituted synthetic mRNA encoding luciferase was purified either with LiCl precipitation (2.5 M LiCl final concentration at –20 °C for one hour, followed by centrifugation and washing of the pellet with 75% ethanol), 'Pseudo LiCl' or MiniElute columns ('Pseudo MiniElute') from Qiagen, A. The ivt mRNA was formulated with MessengerMax (ThermoFisher) according to the manufacturer’s recommendations and added to 100 µL of cells (in each well: 200 ng of mRNA per 1 million fresh human peripheral blood mononuclear cells (PBMCs)) or 10 ng of mRNA per 100,000 human HEK cells or mouse CT26 tumour cells) in RPMI containing 10% foetal bovine serum and antibiotics. Cells that were not treated with the RNA were used as a negative control ('No RNA'). Twenty hours later, the luciferase activity in each well was recorded following the addition of 25 µL of BrightGlo (Promega) and measurement on a GloMax instrument (Promega). The results in A represent the average luciferase activity ± the standard deviation (from triplicates). They show that the two purification methods provide mRNAs with comparable expression capacities. B. The ivt mRNA was formulated with protamine to generate 130 nm nanoparticles\(^{29,30}\) and 10 µg of the preparation (5 µg mRNA plus 5 µg Protamine) were added to 200 µL of cells (1 million fresh human peripheral blood mononuclear cells) in RPMI containing 10% foetal bovine serum and antibiotics. Twenty hours later, the interferon-alpha and TNF-alpha levels in the supernatants were measured using ELISAs (from Mabtech and Bioscience, respectively). A luciferase mRNA generated with the four canonical bases ACGU and precipitated with LiCl ('U LiCl') was used as a positive control. As negative controls, cells without added RNA ('No RNA') or treated with only 5 µg Protamine ('Protamine') were used. The results in B represent the average cytokine concentrations ± standard deviations (from triplicates). They show that the two purification methods give 1-methyl-pseudoUridine substituted RNA with very reduced immunogenic capacities even at the high concentration used here in the cell culture (25 µg/mL final).]
siRNAs or aptamers to “bring a protein” and at the same time “block a protein or pathway”. Meanwhile, drugs have been used to enhance the efficacy of ivt mRNA-based therapies. Thus, ivt mRNAs are a highly versatile tool for research and clinical studies and are predicted to be the format used in many therapies in the future.

Conflict of Interest
The authors have no conflicts of interest to disclose.

Acknowledgements
This work was supported by the University Research Priority Project (URPP) “Translational Cancer Research” and by the Monique Dormonville de la Cour Stiftung.

Received: March 1, 2019

[1] S. Pascolo, J. Multidiscipl. Engin. Sci. Technol. 2017, 4, 6937.
[2] U. Sahin, K. Kariko, Ö. Türeci, Nat. Rev. 2014, 13, 759, doi:10.1038/nrd4278.
[3] S. Pascolo, Expert Opin. Biol. Ther. 2004, 4, 1285, doi:10.1517/14712598.4.8.1285.
[4] S. Pascolo, Methods Mol. Med. 2006, 127, 23, doi:10.1385/1-59745-168-1:23 [pii].
[5] S. Pascolo, Handb. Exp. Pharmacol. 2008, 221, doi:10.1007/978-3-540-72167-3_11.
[6] J. Probst, B. Weide, B. Scheel, B. J. Pichler, I. Hoerr, H. G. Rammensee, S. Pascolo, Gene Ther. 2007, 14, 1175.
[7] S. Pascolo, Expert Rev. Vaccines 2015, 14, 153, doi:10.1586/14760584.2015.1000871.
[8] L. M. Kranz, M. Diken, H. Haas, S. Kreiter, C. Loquai, K. C. Reuter, M. Meng, D. Fritz, F. Vascotto, H. Hefesha, C. Grunwitz, M. Vormehr, Y. Hümels, A. Selmi, A. N. Kuhn, J. Buck, E. Derhovanessian, R. Rae, S. Attig, J. Diekmann, R. A. Jabulowsky, S. Heesch, J. Hassel, P. Langguth, S. Grabbe, C. Huber, Ö. Türeci, U. Sahin, Nature 2016, 534, 396, doi:10.1038/ nature18300.
[9] U. Sahin, E. Derhovanessian, M. Miller, B.-P. Kloke, P. Simon, M. Löwer, B. Vallazza, T. Beissert, V. Bukur, A. N. Kuhn, Ö. Türeci, U. Sahin, Mol. Ther. 2011, 19, 990, doi:10.1038/mt.2010.289.
[10] M. Tusup, S. Pascolo, K. Ghoreschi, C. H. June, S. J. Schuster, Blood 2010, 115, 4333, doi:10.1182/blood-2010-04-292856.
[11] M. Tusup, L. E. French, E. Guenova, T. M. Kundig, S. Pascolo, Biomed. J. Sci. Tech. Res. 2018, doi:10.26717/BJSTR.2018.07.001487.
[12] D. R. Gallie, J. Virol. 2001, 75, 12141, doi:10.1128/JVI.75.24.12141-12152.2001.