

## REVIEW

[View Article Online](#)  
[View Journal](#) | [View Issue](#)Cite this: *RSC Chem. Biol.*, 2024, 5, 418Pseudouridine and *N*1-methylpseudouridine as potent nucleotide analogues for RNA therapy and vaccine developmentLyana L. Y. Ho,<sup>ab</sup> Gabriel H. A. Schiess,<sup>c</sup> Pâmella Miranda,<sup>cd</sup> Gerald Weber<sup>id</sup><sup>c</sup> and Kira Astakhova<sup>id</sup><sup>\*a</sup>

Modified nucleosides are integral to modern drug development, serving as crucial building blocks for creating safer, more potent, and more precisely targeted therapeutic interventions. Nucleobase modifications often confer antiviral and anti-cancer activity as monomers. When incorporated into nucleic acid oligomers, they increase stability against degradation by enzymes, enhancing the drugs' lifespan within the body. Moreover, modification strategies can mitigate potential toxic effects and reduce immunogenicity, making drugs safer and better tolerated. Particularly, *N*1-methylpseudouridine modification improved the efficacy of the mRNA coding for spike protein of COVID-19. This became a crucial step for developing COVID-19 vaccine applied during the 2020 pandemic. This makes *N*1-methylpseudouridine, and its "parent" analogue pseudouridine, potent nucleotide analogues for future RNA therapy and vaccine development. This review focuses on the structure and properties of pseudouridine and *N*1-methylpseudouridine. RNA has a greater structural versatility, different conformation, and chemical reactivity than DNA. Watson–Crick pairing is not strictly followed by RNA that has more unusual base pairs and base-triplets. This requires detailed structural studies and structure–activity relationship analyses for RNA, also when modifications are incorporated. Recent successes in this direction are revised in this review. We describe recent successes with using pseudouridine and *N*1-methylpseudouridine in mRNA drug candidates. We also highlight remaining challenges that need to be solved to develop new mRNA vaccines and therapies.

Received 22nd January 2024,  
Accepted 10th March 2024

DOI: 10.1039/d4cb00022f

[rsc.li/rsc-chembio](https://rsc.li/rsc-chembio)

## Introduction

Deoxyribo- and ribonucleic acids (DNA and RNA) represent a class of natural compounds central to the life on Earth. Their main functions are to store, transmit genetic information and to allow species adopt, survive, and evolve.

Nucleosides are monomers of DNA and RNA (the latter illustrated in Fig. 1A). They have two distinct structural elements: nitrogen heterocycle, called nucleobase, and a ribose carbohydrate. Each element serves as a scaffold for multiple non-covalent interactions, vital to biological performance of nucleic acids. Nucleobase provides with  $\pi$ -stacking and hydrogen bonding interactions, forming nucleic acid helices and universal genetic code.

Given crucial biological importance of nucleic acids, interference with their synthesis and function creates powerful

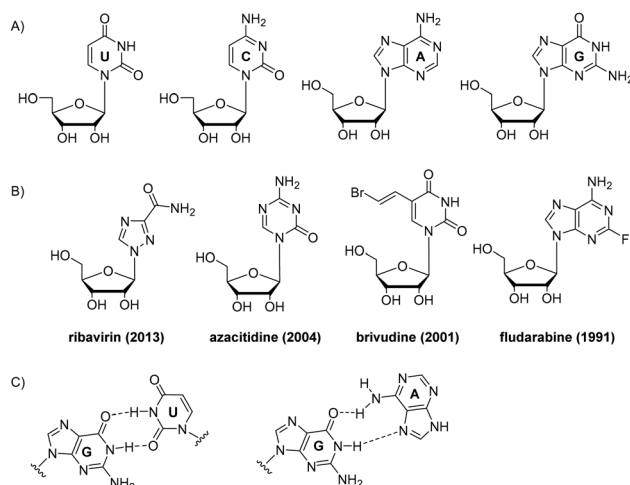


Fig. 1 Chemical structures of natural RNA nucleotides (A), examples of nucleoside drugs (year of FDA approval) (B), and wobble RNA base pairs: GU and GA (C).

<sup>a</sup> Technical University of Denmark, 2800 Kongens Lyngby, Denmark.  
E-mail: kiraas@kemi.dtu.dk

<sup>b</sup> The Hong Kong Polytechnic University, 11 Yuk Choi Rd, Hung Hom, Hong Kong

<sup>c</sup> Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

<sup>d</sup> Programa Interunidades de Pós-Graduação em Bioinformática, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

therapeutic interventions. This can be achieved *via* chemical modification of nucleosides.

Multiple nucleoside drugs with nucleobase modifications have been developed. Several examples are shown in Fig. 1B. Modified nucleoside monomers include drugs for treating viral infections (ribavirin, idoxuridine, trifluridine, brivudine), cancer (azacitidine, fludarabine, 6-mercaptopurine, decitabine), and immunosuppressive drug azathioprine, used in organ transplantation and autoimmune diseases.<sup>1</sup> Due to their altered nucleobase structure, these drugs interfere with DNA synthesis, transcription and/or translation leading to cell death.<sup>1</sup>

As to oligonucleotide drugs, it mostly ribose and phosphate backbone modifications that have been incorporated. Chemical modifications in these drugs result in enhanced stability, improved pharmacokinetics, reduced toxicity, increased target specificity, or resistance to degradation, thereby improving their therapeutic potential.<sup>2</sup>

Over the last two decades, there has been a growing interest in RNA drug development. RNA differs from DNA by having a 2'-hydroxyl group in ribose sugar, which gives rise to unique conformational features, specific hydration, and electrostatic properties of RNA (Fig. 1A).<sup>1</sup> Another structural feature of RNA is that uracil nucleobase does not contain the major groove 5-methyl group as that in DNA nucleotide thymidine. As a result of this, RNA duplexes are more compact than B-DNA helices. Ribose nucleotides tend to adopt the C3'-*endo* sugar pucker resulting in compact A-form RNA duplexes with eleven base pairs per turn.<sup>1</sup>

RNA has various alternative base pairings, such as G:U wobble pair and A:C pair, which differ from DNA. G:U pair has an exocyclic amino group that acts as a prominent hydrogen bond donor in the minor grooves, important for recognising proteins and other ligands (Fig. 1C).<sup>3</sup> G:U pair frequently occurs, such as in codon-anticodon interaction which forms the genetic code between tRNA and mRNA. The hydrogen bonding of G:U are strongly dependent on their next neighbour pairs, ranging from very weak bonding when in tandem configuration to strong if located at the end of a double strand.<sup>4</sup> G:A pairs are ubiquitous at RNA helices termini, in loops and folds of tertiary structure. Among G:A, G:G and A:A pairs, sheared G:A pairs that include a Hoogsteen interaction are most common (Fig. 1C).<sup>3</sup>

RNA nucleotides form multiple interactions with nucleobases and sugars. G-tetrad and G-quartet are similar for DNA and RNA but more diversity is found in RNA. Quadruple interaction can be formed in highly folded RNA molecules as well.<sup>3</sup>

RNA folding is complex and sensitive to temperature and buffer composition. Metal ions like Mg<sup>2+</sup> and K<sup>+</sup> stabilize RNA tertiary structure.<sup>3</sup> As a result of stable structure, misfolding of RNA could become a severe problem with *e.g.*, therapeutics and vaccine candidates.<sup>3</sup>

Pseudouridine ( $\Psi$ ) is an abundant post-transcriptional modification of mRNA which together with N1-methylpseudouridine (N1-Me- $\Psi$ ), became a milestone in developing mRNA drug candidates. This review describes  $\Psi$  and N1-Me- $\Psi$  nucleobase

analogues, their impact on RNA structure and appealing properties in RNA vaccines and therapeutic candidates. We also give highlights on most potent applications of modified mRNA achieved so far, discuss future directions and remaining challenges.

## Pseudouridine

The first known and one of the most abundant RNA modifications is the pseudouridine ( $\Psi$ ), or 5-ribosyluracil, which is a post-transcriptional modification and an isomer of uridine (U) (Fig. 2A). Pseudouridine contains a C-C base-sugar bond due to the uracil base is attached to the sugar by a C1'-C5 bond unlike a C1'-N1 glycosidic linkage (Fig. 2), enhancing the base rotation. Moreover, it has an additional ring nitrogen atom (N1 imino atom), which behaves as an additional hydrogen bond donor.<sup>5,6</sup> The replacement of U by  $\Psi$  promotes a C3'-*endo* sugar conformation and increases the local base stacking, thermodynamically stabilizing RNA duplexes.<sup>5,7,8</sup>

Pseudouridine ( $\Psi$ ) sites have been identified in different types of natural mRNA, such as tRNA, rRNA, snRNA and mRNA.<sup>9-13</sup> Intracellular  $\Psi$  formation is catalyzed by a class of enzymes known as pseudouridine synthases.<sup>14</sup> Pseudouridine synthases can be classified into two main families: stand-alone pseudouridine synthases and pseudouridine synthase domains within larger proteins. Stand-alone pseudouridine synthases include TruA found in bacteria and TruB found in bacteria and yeast. In eukaryotes, several pseudouridine synthase domains have been found. Eucaryotic H/ACA box small nucleolar ribonucleoproteins (snoRNPs) have a dyskerin (Cbf5) component which catalyzes pseudouridylation in rRNA, snRNA, and telomerase RNA. Nop10 is another component of H/ACA snoRNPs which is involved in the pseudouridylation activity.<sup>14</sup>

In humans, there are multiple enzymes catalyzing pseudouridylation. RPUSD (RNA pseudouridylate synthase domain-containing) family consists of several enzymes involved in  $\Psi$  formation. RPUSD1 and RPUSD2 are involved in mitochondrial RNA pseudouridylation. The PUS family consists of enzymes with pseudouridine synthase activity towards multiple RNA species, represented by PUS1 and PUS7.

Quantitative  $\Psi$  detection remains being challenging. Therefore its biological role is still not completely understood.

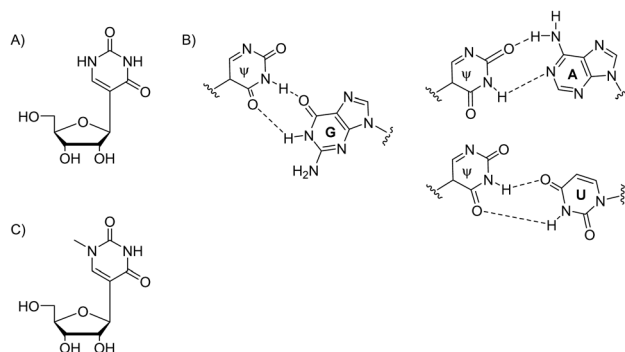


Fig. 2 Chemical structure of  $\Psi$  (A), its base pairing with G, A and U nucleobases (B); chemical structure of N1-methyl- $\Psi$  (C).



Incorporated into mRNA,  $\Psi$  increases its stability and modifies various cellular and biological processes, such as transcription, pre-mRNA splicing and mRNA translation.<sup>14</sup>  $\Psi$  was also found in stop codons.<sup>15</sup>

Pseudouridylation is a crucial post-transcriptional modification that influences the structural stability and function of transfer RNA (tRNA) as well. With nanopore sequencing it has been shown that  $\Psi$  modifications are present in diverse positions of tRNA.<sup>16</sup> As a result of a modification in the *hisT* gene, the tRNA *E. coli* and *Salmonella typhimurium* lacks  $\Psi$ 38,  $\Psi$ 39 and  $\Psi$ 40 positions and may have a reduced polypeptide chain elongation rate (20–25%) and longer cell division time (30%). In addition, inhibition of  $\Psi$ 38 and  $\Psi$ 39, as a result of *DEG1* gene disruption, in *S. cerevisiae* leads to reduced growth rate.<sup>9</sup>

Notably, pseudouridylation of tRNA affects its methylation at other positions. In tRNA<sup>Phe</sup> of *S. cerevisiae*,  $\Psi$ 55 positively influences the introduction of methylated nucleotides, m<sup>5</sup>U54 and m<sup>1</sup>A58.<sup>16</sup>

Aberrant pseudouridine modifications in tRNA can impact translation accuracy and efficiency, potentially contributing to disease pathogenesis.<sup>10,11,17–20</sup> For example, mutations in the *PUS1* gene, which encodes a pseudouridine synthase, are associated with mitochondrial myopathy and sideroblastic anaemia.<sup>19</sup> Abnormal pseudouridylation has been implicated in neurological disorders.<sup>20</sup> Altered tRNA modifications, including pseudouridylation, have been observed in various cancer types.<sup>9,10,17,18</sup>

Nucleic acid function is guided by non-covalent interactions between polynucleotides and with cellular protein machinery.<sup>3</sup> Therefore 2D and 3D structure of RNA is central to its interactions and function. Regarding impact on RNA structure, two hydrogen bond donors by N1 and N3 imine protons, and two hydrogen bond acceptors are present in  $\Psi$  (Fig. 2B). Therefore,  $\Psi$  acts as a universal nucleobase by enhancing stability when it pairs with A, G, U or C in a double helix. The relationship between  $\Psi$ -modified mRNA structure and stability is supported by multiple examples of *in vitro*-synthesised, more stable,  $\Psi$ -containing mRNAs, and the natural  $\Psi$ -containing mRNAs' increased half-life time, as in eukaryotic parasite *Toxoplasma gondii*.<sup>21,22</sup>

NMR structural studies found that  $\Psi$  stacks better than U due to the C1'–C5 glycoside bond in  $\Psi$  that has higher rotational lability than the C1'–N1 bond in U.<sup>5,23</sup>  $\Psi$  enhances local RNA base stacking in both single- and double-stranded conformations and promotes an increase in neighbouring stacking stabilization at nucleotide level.<sup>7,9</sup> In addition,  $\Psi$  induces a more rigid phosphodiester backbone in its vicinity, increasing the base stacking neighbouring, perhaps due to C–C isomerization in  $\Psi$ .<sup>9,14</sup> Moreover, the base stacking enhancement is considered the most important contribution of  $\Psi$  to RNA stability.<sup>7,24–27</sup>

To investigate more the impact of  $\Psi$  on RNA structure, thermodynamic studies were conducted.<sup>5</sup> The data revealed that  $\Psi$  stabilizes the RNA duplexes when U is replaced and  $\Psi$ –A,  $\Psi$ –G,  $\Psi$ –U and  $\Psi$ –C pairs are formed. The effect of pseudouridylation is more significant when it occurs at the internal

positions of the RNA structure, usually releasing more negative free energy. Enhanced thermodynamic stability by  $\Psi$  is around 0.5 kcal mol<sup>−1</sup>, except in 5'-C $\Psi$ G/3'-GAC and 5'-G $\Psi$ C/3'-CGG, nearly 2.5 and 1.5 kcal mol<sup>−1</sup> is recorded respectively.  $\Psi$  contribution is, in most cases, weaker than a typical hydrogen bond (2–10 kcal mol<sup>−1</sup>), which indicates that the enhanced stability is a result of a different mechanism.<sup>28</sup> This endorses the hypothesis that the contribution of  $\Psi$  to the stability arises from the C1'–C5 bond properties. Current NMR data<sup>5</sup> suggest that the N1H in the 5'-C $\Psi$ G/3'-GAC trimer is in a weak hydrogen bond while there is no evidence for hydrogen bonding in the 5'-G $\Psi$ C/3'-CGG trimer.<sup>5</sup>  $\Psi$  biological roles, such as non-sense suppression of stop codons and non-synonymous translation, might be better understood by identifying contexts where N1H is in an active hydrogen bond.<sup>29</sup> The 3D structures and computational studies of these outlier duplexes could be interesting to investigate due to little knowledge of the interactions that are favourable to stacking.

$\Psi$ –A,  $\Psi$ –G,  $\Psi$ –U and  $\Psi$ –C base pairs have shown enhancement of base stacking and also sequence context dependence.<sup>5,6,30</sup> High level quantum mechanical (QM) methods have shown a clear dependence of the change in the base stacking energies concerning the sequence context with a range from −1.59 to 0.23 kcal mol<sup>−1</sup>. However, the change from U–A to  $\Psi$ –A base pair not always stabilizes the stacking interactions of the duplex.<sup>6,30</sup> Yet experimental data of internal and terminal  $\Psi$ –A pair has shown greater stability in comparison to predicted data of U–A pair in the same duplexes, on average 1.7 and 1.0 kcal mol<sup>−1</sup>, respectively.<sup>30</sup>

Terminal  $\Psi$ –A, –C, –G and –U pairs show stability enhancement in both 5' and 3' terminals in a similar way to U–A, –C, –G and –U pairs, respectively. Curiously, those  $\Psi$  pairs at 3' terminal reach slightly higher energies than the U pairs, except  $\Psi$ –A pair.  $\Psi$ –C pair contributes with more stability than U–C pair only when added at 3' terminal. On the other hand, only  $\Psi$ –U pair shows a more stable behaviour than U pair in 5' terminal. Similarly,  $\Psi$ –G pairs have higher stabilization when at 3' terminal (3.22 kcal mol<sup>−1</sup>) in comparison to U–G pairs (2.44 kcal mol<sup>−1</sup>). Finally, any internal  $\Psi$  pair enhances the thermodynamic stability of RNA duplexes. Mainly  $\Psi$ –A and  $\Psi$ –G pairs have large energies, within a 0.4–0.8 kcal mol<sup>−1</sup>, when compared to A–U and G–U pairs, perhaps due to the favourable base stacking and hydrogen bonds of  $\Psi$ .<sup>5,31</sup>

### RNA N1-methylpseudouridine structure

N1-Methylpseudouridine (N1-methyl- $\Psi$ ) is another naturally occurring RNA nucleotide analogue<sup>32</sup> which has an extra hydrogen bond donor in the nucleobase (Fig. 2C). This N1-modified structure can be found in natural 18S rRNA and tRNA, not only in humans but also in archaea and eukaryotes. Research shows that RNAs of Thermococcales and Nanoarchaea<sup>33</sup> include N1-methyl- $\Psi$ .<sup>34,35</sup> N1-Methyl- $\Psi$  biosynthetic pathway begins with converting U to  $\Psi$ , being catalyzed by the aforementioned pseudouridine synthase Pus10.<sup>36,37</sup> S-Adenosylmethionine (SAM)-dependent pseudouridine N3-methyltransferase YbeA in eubacteria<sup>38,39</sup> and N1-specific pseudouridine methyltransferase



Nep1 are two enzymes that can further methylate  $\Psi$ . Nep1 catalyse the N1-specific pseudouridine methylation of position 1191 (*Saccharomyces cerevisiae* numbering, nucleotide 913 in *M. jannaschii*) in RNA.<sup>40,41</sup>

$\Psi$ - and N1-methyl- $\Psi$  modified RNA duplexes have different stability which has been confirmed with multiple thermal denaturation studies.<sup>5</sup> N1-Methyl- $\Psi$  differs from  $\Psi$  by having a methyl group to replace the extra hydrogen at N1-position, so it no longer has the universal base character like  $\Psi$  (Fig. 2). N1-Methyl- $\Psi$  only forms the traditional Watson-Crick pair but both  $\Psi$  and N1-methyl- $\Psi$  have the C5-C1' bond that allows rotation between the nucleobase and sugar, achieving better base-pairing, base-stacking and duplex stability.<sup>42</sup>

A recent molecular dynamics study shows that N1-methyl- $\Psi$  induces a higher stabilization effect of the dsRNA, due to stronger stacking and base pair interactions, than  $\Psi$ . Moreover, N1-methyl- $\Psi$ :A pair have a stronger binding interaction than both U:A and  $\Psi$ :A pairs in the majority of neighbours context.<sup>43</sup>

Overall, N1-methyl- $\Psi$  is more like uridine in translation coding but at the same time, behaves like  $\Psi$  that let the mRNA not trigger the immune response which is of critical importance to RNA vaccines and therapies, described below.

### mRNA drug development

mRNA holds a potential to treat a vast spectrum of diseases and to act as a vaccine, by producing the desired protein *in vivo* and acting as adjuvant.<sup>44</sup> Key steps in mRNA drug development include: sequence design, chemical modification, formulation, testing *in vitro* and *in vivo*, and finally, trials.

Currently there are two approved mRNA vaccines, by Moderna and Pfizer.<sup>45,46</sup> According to FDA, January 2024, there are 56 mRNA drugs in the clinical pipeline worldwide, with R&D mainly focused on vaccines, accounting for about 84%, while therapeutic drugs account for about 16%. Except for the mRNA COVID-19 vaccine, which is urgently marketed, most others are in the early stages.<sup>47</sup>

Over the last decade, therapeutic mRNAs and mRNA vaccines have been encountering several major obstacles. First, long single (ss) or double stranded (ds) RNAs in the cytosol are commonly derived from the genome of RNA viruses or intermediate products that are generated during viral replication, leading to immune response.<sup>48</sup> Introducing  $\Psi$  and N1-Me- $\Psi$  has been a breakthrough to overcome this challenge. Immune responses to RNA and their inhibition with  $\Psi$  are illustrated in Fig. 3. TLR7 and TLR8 recognize single stranded (ss) RNA. These TLRs preferentially recognize polyuridine (polyU) and guanosine/uridine-rich (GU-rich) sequences. TLR7 and TLR8 also recognize RNA degradation products and require free guanosines and uridines, respectively, for maximum activation.  $\Psi$  and N1-methyl- $\Psi$  could optimise mRNA performance by reduced immunogenicity and effective protein translation. Reduced binding to TLRs due to  $\Psi$ .

Delivery has been another main challenge to mRNA therapeutics and vaccines, that could not be solved solely with chemical modification of mRNA. mRNA is rapidly degraded *in vivo* and poorly uptaken by cells. Delivery systems like lipid

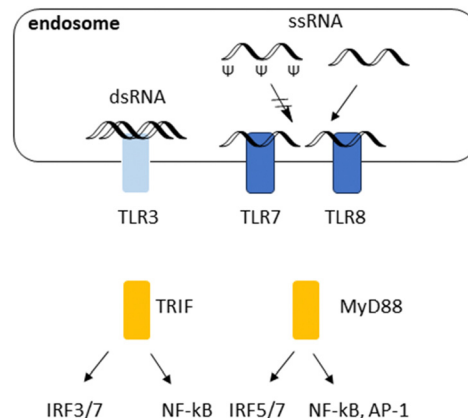


Fig. 3 RNA response mechanisms and impact of  $\Psi$  modification. Endosomal RNA sensor TLR3 recognizes dsRNA and subsequently, activates the TRIF-dependent pathway to activate IRF3/7 and NF- $\kappa$ B. ssRNA binds to TLR7 and TLR8, activating MyD88, NF- $\kappa$ B, IRF5/7 and AP-1. TLR, toll-like receptor; IRF, interferon regulatory factor; NF- $\kappa$ B, activation protein 1, AP-1.

nanoparticles (LNPs) have high efficiency, low toxicity and are applicable to various cell types. LNP-mRNA nanoparticle formulations have been successful and reached clinical applications.<sup>44–46</sup> LNPs have encapsulation efficiency reaching 90%, and can be targeted through surface modification.<sup>49,50</sup>

On-going research on LNP mRNA covers multiple potential applications. Main direction in current trials is cancer vaccines. Very recently Moderna and Merck published successful results of phase IIb study on melanoma cancer vaccine.<sup>51</sup> However it is a limiting factor that MC3 dLin-MC3-DMA (also known as C12-200) cationic formulation used in COVID-19 vaccine is patented and cannot be broadly used for other mRNA drug candidates. To overcome this, there are attempts to develop alternative formulations. Cationic lipid formulations using DOTAP (1,2-dioleoyl-3-trimethylammonium-propane), PEGylated lipids and DOPE (dioleoylphosphatidylethanolamine) are being tested.<sup>52,53</sup>

The delivery tools for mRNA are not limited to LNP. Naturally occurring vesicles, such as exosomes, can be used to carry mRNA.<sup>52</sup> These vesicles are derived from cells and besides mRNA encapsulation, can facilitate cell-specific targeting. Encapsulation efficiency of exosomes is lower than for LNP, however exosomes are less toxic and have higher target specificity.<sup>52</sup>

Polymeric nanoparticles and dendrimers have been actively explored for the mRNA delivery.<sup>53</sup> Polymeric nanoparticles, such as those made from polyethyleneimine (PEI), poly(lactic-co-glycolic acid) (PLGA), or chitosan, can form complexes with siRNA through electrostatic interactions. These nanoparticles protect mRNA from degradation and can enhance cellular uptake.<sup>53,54</sup> Other examples of reported nanoparticles for mRNA include inorganic nanoparticles, such as gold or silica nanoparticles, hybrid nanoparticles, virus-like nanoparticles and peptide-based nanoparticles.<sup>55–58</sup> They are extensively reviewed elsewhere.

Overall, there is no universal solution to all nucleic acid delivery tasks. Choosing the most suitable delivery system





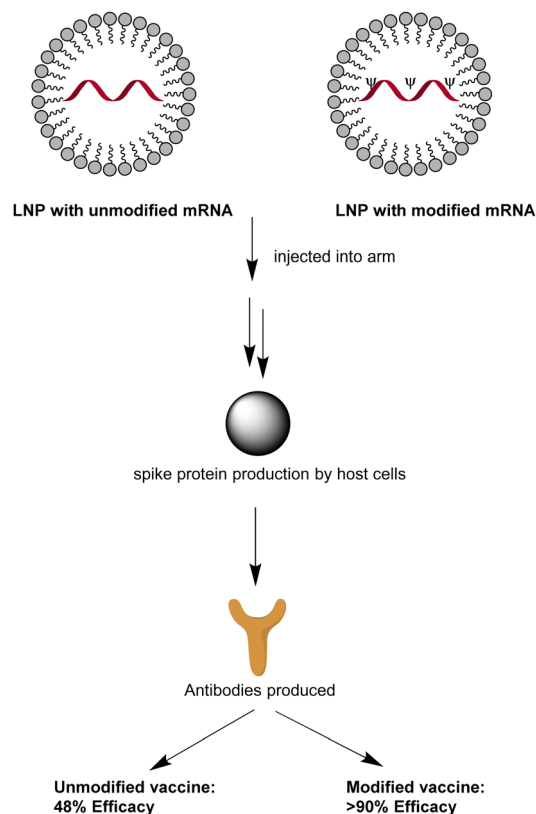


Fig. 4 SARS-CoV-2 mRNA vaccination.

depends on the specific requirements of the therapeutic application, including the targeted tissue, desired release kinetics, and safety considerations.

### N1-Methylpseudouridine impact on COVID-19 mRNA vaccine

A key principle of the mRNA-based vaccination is that under low dosage, non-modified mRNA encodes the antigen while acting as an adjuvant (Fig. 4). Restricted to a maximum of 12  $\mu\text{g}$  dosage by patients' tolerance in late-stage clinical trials, unmodified mRNA vaccine CVnCoV maintained only 48% efficacy, regardless of the disease severity.<sup>45</sup> In contrast, 30  $\mu\text{g}$  Pfizer-BioNTech or 100  $\mu\text{g}$  Moderna's mRNA vaccines could demonstrate around 95% high protection rate against COVID-19 after modification with N1-methyl- $\Psi$ .<sup>45</sup>

Vaccines composed of unmodified and modified mRNA compared after being injected into the muscle of the upper arm.<sup>45</sup> An immune response would be created and N1-methylpseudouridylated mRNA demonstrates over 90% efficacy against COVID-19 symptoms, much higher than the unmodified one which is beneath 50%.

Clearly, N1-methyl- $\Psi$  has different impact on product protein translation compared to  $\Psi$ . According to Kim *et al.* (2022), N1-methyl- $\Psi$  does not necessarily change the decoding accuracy in a reconstituted system.<sup>59</sup> Neither would it increase the probability of miscoded peptides, nor stabilise the mismatched RNA-duplex formed. It only has a slight tendency in increasing errors when reverse transcription occurs.

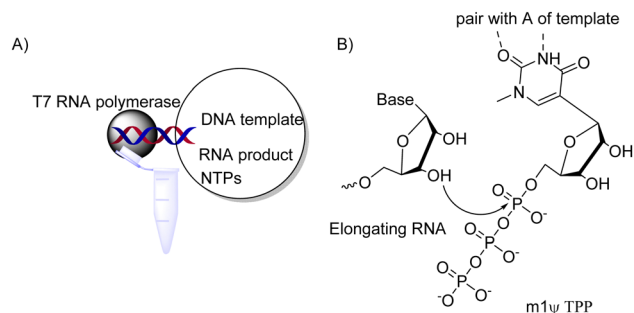


Fig. 5 Producing N1-methyl- $\Psi$ -modified mRNA by *in vitro* transcription. (A) Key components of enzymatic reaction and RNA product; NTP = nucleotide triphosphate;<sup>63</sup> (B) enzymatic 3'  $\rightarrow$  5' elongation of RNA during *in vitro* transcription.

### Remaining challenges in therapeutic mRNA development

mRNA development needs to be supported by research on mRNA structure since it is closely related to the function of mRNA. Nonetheless, mRNA structure prediction is still unreliable despite combining the usage of thermodynamic stability and evolutionary covariation information.<sup>60</sup> The conserved mRNA structure could be predicted by the combination of three features: using significant covariation, negative evolutionary information and a plethora of probabilistic folding algorithms which incorporate those positive covariations into a single structure.<sup>39</sup>

Another challenge is how to produce exogenous long mRNA while incorporating N1-methyl- $\Psi$  flexibly on a large scale. Getting inspiration from nature and utilizing enzymes is an effective method. In this approach, fragments can become the building blocks of plasmids for COVID-19 vaccine synthesis. According to past studies, T7 polymerase could effectively build RNAs longer than 20 000 nucleotides perfectly without error<sup>61,62</sup> while tolerating non-natural NTPs. m1 $\Psi$  triphosphate can be incorporated into RNA by polymerases providing long modified RNA molecules (Fig. 5).<sup>63</sup> This allows producing large quantities of long modified RNA, of tremendous benefit to RNA drug development and commercialization.

## Conclusions

Overall, mRNA therapy is a cutting-edge medical approach that uses synthetic mRNA molecules to prevent and treat various conditions. This technology has gained significant attention and recognition, particularly due to its role in the development of COVID-19 vaccines.<sup>45</sup> mRNA therapy has the potential to treat a wide range of diseases, including infectious diseases, cancer, genetic disorders, and more.<sup>64–66</sup> By introducing synthetic mRNA into the body, it's possible to instruct cells to produce specific proteins that can correct or combat disease. As to vaccine development, mRNA technology gained widespread attention during the COVID-19 pandemic when companies like Pfizer-BioNTech and Moderna developed highly effective COVID-19 vaccines using this approach.<sup>45</sup> The vaccines contain mRNA that encodes the spike protein of the SARS-CoV-2 virus,



which triggers an immune response, providing protection against the virus.<sup>45</sup>

One of the advantages of mRNA therapy is its rapid development and adaptability. Creating synthetic mRNA for a new target can be faster and more scalable than traditional vaccine or drug development processes. With LNP technology, efficient delivery can be achieved as well.<sup>64–66</sup>

Despite its promise, mRNA therapy faces challenges, including ensuring the stability and delivery of mRNA molecules to target cells, managing immune responses, and addressing potential side effects. These pitfalls can be overcome by using chemical analogues of RNA nucleotides. Pseudouridine ( $\Psi$ ) is an important modification that is conserved naturally in the RNA structure. Studies showed that  $\Psi$ -modification allows mRNA to resist intrinsic immune responses<sup>67</sup> and  $\Psi$ -derivatives could further improve mRNA properties, such as stability and efficacy of translation.

To apply this technology to vaccine and therapy development, deeper knowledge on the impact of  $\Psi$  and its analogues on RNA structure is required. Computational approaches, thermodynamics and structural investigations with *e.g.*, NMR would be significant next steps in this direction.

## Author contributions

The review is conceptualized by GHA Schiess and K. Astakhova. LLY Ho and K. Astakhova made the initial draft. All authors contributed to finalising the review.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was supported by the Danish Foundation for Independent Research (award 3105-00141B) to KAS; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) to GHAS and PM; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig) to GW.

## Notes and references

- 1 E. De Clercq and G. Li, Approved Antiviral Drugs over the Past 50 Years, *Clin. Microbiol. Rev.*, 2016, **29**(3), 695–747.
- 2 T. C. Roberts, R. Langer and M. J. A. Wood, Advances in oligonucleotide drug delivery, *Nat. Rev. Drug Discovery*, 2020, **19**(10), 673–694.
- 3 *Nucleic Acids in Chemistry and Biology*, ed. G. M. Blackburn, M. J. Gait, D. Loakes and D. Williams, Royal Society of Chemistry, 2006.
- 4 T. D. Amarante and G. Weber, Evaluating hydrogen bonds and base stacking of single, tandem and terminal GU mismatches in RNA with a mesoscopic model, *J. Chem. Inf. Model.*, 2016, **56**(1), 101–109.
- 5 E. Kierzek, M. Malgowska, J. Lisowiec, D. H. Turner, Z. Gdaniec and R. Kierzek, The contribution of pseudouridine to stabilities and structure of RNAs, *Nucleic Acid Res.*, 2014, **42**(5), 3492–3501.
- 6 I. Deb, Ł. Popenda and J. Sarzyńska, *et al.*, Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair, *Sci. Rep.*, 2019, **9**, 16278.
- 7 D. R. Davis, Stabilization of RNA stacking by pseudouridine, *Nucleic Acid Res.*, 1995, **23**(24), 5020–5026.
- 8 K. B. Hall and L. W. McLaughlin, Properties of a U1/mRNA 5' splice site duplex containing pseudouridine as measured by thermodynamic and NMR methods, *Biochemistry*, 1991, **30**(7), 1795–1801.
- 9 M. Charette and M. W. Gray, Pseudouridine in RNA: what, where, how, and why, *IUBMB Life*, 2000, **49**(5), 341–352.
- 10 X. Xue, Z. Wang, Y. Wang and X. Zhou, Disease Diagnosis Based on Nucleic Acid Modifications, *ACS Chem. Biol.*, 2023, **18**(10), 2114–2127.
- 11 P. Nombela, B. Miguel-López and S. Blanco, The role of m6A, m5C and  $\Psi$  RNA modifications in cancer: Novel therapeutic opportunities, *Mol. Cancer*, 2021, **20**(1), 1–30.
- 12 S. Tavakoli, M. Nabizadeh, A. Makhamreh, H. Gamper, C. A. McCormick, N. K. Rezapour, Y. M. Hou, M. Wanunu and S. H. Rouhanifard, Semi-quantitative detection of pseudouridine modifications and type I/II hypermodifications in human mRNAs using direct long-read sequencing, *Nat. Commun.*, 2023, **14**(1), 334.
- 13 Q. Dai, L. S. Zhang, H. L. Sun, K. Pajdzik, L. Yang, C. Ye, C. W. Ju, S. Liu, Y. Wang, Z. Zheng, L. Zhang, B. T. Harada, X. Dou, I. Irklyenko, X. Feng, W. Zhang, T. Pan and C. He, Quantitative sequencing using BID-seq uncovers abundant pseudouridines in mammalian mRNA at base resolution, *Nat. Biotechnol.*, 2023, **41**(3), 344–354.
- 14 S. H. Boo and Y. K. Kim, The emerging role of RNA modifications in the regulation of mRNA stability, *Exp. Mol. Med.*, 2020, **52**(3), 400–408.
- 15 J. Song, L. Dong, H. Sun, N. Luo, Q. Huang, K. Li, X. Shen, Z. Jiang, Z. Lv, L. Peng, M. Zhang, K. Wang, K. Liu, J. Hong and C. Yi, CRISPR-free, programmable RNA pseudouridylation to suppress premature termination codons, *Mol. Cell*, 2023, **83**(1), 139–155e9.
- 16 M. C. Lucas, L. P. Pryszcz, R. Medina, I. Milenkovic, N. Camacho, V. Marchand, Y. Motorin, L. Ribas de Paula and E. M. Novoa, Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing, *Nat. Biotechnol.*, 2024, **42**(1), 72–86.
- 17 N. M. Martinez, A. Su, M. C. Burns, J. K. Nussbacher, C. Schaening, S. Sathe, G. W. Yeo and W. V. Gilbert, Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing, *Mol. Cell*, 2022, **82**(3), 645–659e9.
- 18 J. A. Stockert, A. Gupta, B. Herzog, S. S. Yadav, A. K. Tewari and K. Yadav, Predictive value of pseudouridine in prostate cancer, *Am. J. Clin. Exp. Urol.*, 2019, **7**(4), 262–272.
- 19 E. Fernandez-Vizarra, A. Berardinelli, L. Valente, V. Tiranti and M. Zeviani, Nonsense mutation in pseudouridylate



- synthase 1 (PUS1) in two brothers affected by myopathy, lactic acidosis and sideroblastic anaemia (MLASA), *BMJ Case Rep.*, 2009, **2009**, bcr05.2009.1889.
- 20 T. Y. Lin, R. Smigiel, B. Kuzniewska, J. J. Chmielewska, J. Kosińska, M. Biela, A. Biela, A. Kościelniak, D. Dobosz, I. Laczmańska, A. Chramiec-Głąbik, J. Jeżowski, J. Nowak, M. Gos, S. Rzonca-Niewczas, M. Dziembowska, R. Płoski and S. Glatt, Destabilization of mutated human PUS3 protein causes intellectual disability, *Hum. Mutat.*, 2022, **43**(12), 2063–2078.
  - 21 M. J. Holmes, L. R. Padgett, M. S. Bastos and W. J. Sullivan Jr, m6A RNA methylation facilitates pre-mRNA 3'-end formation and is essential for viability of *Toxoplasma gondii*, *PLoS Pathog.*, 2021, **17**(7), e1009335.
  - 22 M. A. Nakamoto, A. F. Lovejoy, A. M. Cygan and J. C. Boothroyd, mRNA pseudouridylation affects RNA metabolism in the parasite *Toxoplasma gondii*, *RNA*, 2017, **23**(12), 1834–1849.
  - 23 N. B. Leontis and E. Westhof, Analysis of RNA motifs, *Curr. Opin. Struct. Biol.*, 2003, **13**(3), 300–308.
  - 24 D. R. Davis, Biophysical and conformational properties of modified nucleosides in RNA (nuclear magnetic resonance studies), in *Modification and Editing of RNA*, ed. H. Grosjean and R. Benne, 1998, pp. 85–102.
  - 25 D. R. Davis, C. A. Veltri and L. Nielsen, An RNA model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in tRNA<sup>Lys</sup>, tRNA<sup>His</sup> and tRNA<sup>Tyr</sup>, *J. Biomol. Struct. Dyn.*, 1998, **15**, 1121–1132.
  - 26 C. S. Yarian, M. M. Basti, R. J. Cain, G. Ansaria, R. H. Guenther, E. Sochacka, G. Czerwinska, A. Malkiewicz and P. F. Agris, Structural and functional roles of the N1- and N3-protons of  $\Psi$  at tRNA's position 39, *Nucleic Acids Res.*, 1999, **27**, 3543–3549.
  - 27 P. C. Durant and D. R. Davis, Stabilization of the anticodon stemloop of tRNA<sup>Lys,3</sup> by an A<sup>+</sup>–C base pair and by pseudouridine, *J. Mol. Biol.*, 1999, **285**, 115–131.
  - 28 T. W. Martin and Z. S. Derewenda, The name is bond—H bond, *Nat. Struct. Biol.*, 1999, **6**(5), 403–406.
  - 29 H. Monajemi and S. M. Zain, How stop codon pseudouridylation induces nonsense suppression, *Comput. Theor. Chem.*, 2021, **1204**, 113414.
  - 30 G. A. Hudson, R. J. Bloomingdale and B. M. Znosko, Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides, *RNA*, 2013, **19**(11), 1474–1482.
  - 31 N. Sugimoto, R. Kierzek, S. M. Freier and D. H. Turner, Energetics of internal GU mismatches in ribooligonucleotide helices, *Biochemistry*, 1986, **25**(19), 5755–5759.
  - 32 R. C. Brand, *et al.*, Biosynthesis of a Hypermodified Nucleotide in *Saccharomyces Carlsbergensis* 17S and HeLa-Cell 18S Ribosomal Ribonucleic Acid, *Biochem. J.*, 1978, **169**(1), 71–77.
  - 33 J. Urbonavicius, *et al.*, Acquisition of a bacterial RnaA-type tRNA(uracil-54, C5)-methyltransferase by Archaea through an ancient horizontal gene transfer, *Mol. Microbiol.*, 2007, **67**(2), 323–335.
  - 34 H. Pang, M. Ihara, Y. Kuchino, S. Nishimura, R. Gupta, C. R. Woese and J. A. McCloskey, Structure of a modified nucleoside in archaeobacterial tRNA which replaces ribosylthymine. 1-Methylpseudouridine, *J. Biol. Chem.*, 1982, **257**(7), 3589–3592.
  - 35 R. Gupta, Halobacterium volcanii tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs, *J. Biol. Chem.*, 1984, **259**(15), 9461–9471.
  - 36 P. Gurha and R. Gupta, Archaeal Pus10 proteins can produce both pseudouridine 54 and 55 in tRNA, *RNA*, 2008, **14**(12), 2521–2527.
  - 37 J. P. Wurm, *et al.*, Identification of the enzyme responsible for N1-methylation of pseudouridine 54 in archaeal tRNAs, *RNA*, 2012, **18**(3), 412–420.
  - 38 R. Rya, *et al.*, Identification of pseudouridine methyltransferase in *Escherichia coli*, *RNA*, 2008, **14**(10), 2223–2233.
  - 39 E. Purta, *et al.*, YbeA is the m<sup>3</sup> $\Psi$  methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA, *RNA*, 2008, **14**(10), 2234–2344.
  - 40 J. P. Wurm, *et al.*, The ribosome assembly factor NEP1 responsible for Bowen–Conradi syndrome is a pseudouridine-N1-specific methyltransferase, *Nucleic Acid Res.*, 2010, **37**(7), 2387–2398.
  - 41 B. Meyer, *et al.*, The Bowen–Conradi syndrome protein Nep1 (Emg1) has a dual role in eukaryotic ribosome biogenesis, as an essential assembly factor and in the methylation of  $\Psi$ 1191 in yeast 18S rRNA, *Nucleic Acids Res.*, 2011, **39**(4), 1526–1537.
  - 42 E. Westhof, Pseudouridines or how to draw on weak energy differences, *Biochem. Biophys. Res. Commun.*, 2019, **520**(4), 702–704.
  - 43 N. Dutta, I. Deb, J. Sarzynska and A. Lahiri, Structural and thermodynamic consequences of base pairs containing pseudouridine and N1-methylpseudouridine in RNA duplexes, *bioRxiv*, 2023.
  - 44 C. Wang, Y. Zhang and Y. Dong, Lipid nanoparticle-mRNA formulations for therapeutic applications, *Acc. Chem. Res.*, 2021, **54**(23), 4283–4293.
  - 45 P. Morais, H. Adachi and Y. T. Yu, The critical contribution of pseudouridine to mRNA COVID-19 vaccines, *Front. Cell. Dev. Biol.*, 2021, **9**, 789427.
  - 46 Y. N. Lamb, BNT162b2 mRNA COVID-19 Vaccine: First Approval, *Drugs*, 2021, **81**(4), 495–501.
  - 47 Web: <https://www.fda.gov>; accessed January 21st 2024.
  - 48 T. Uehata and O. Takeuchi, RNA Recognition and Immunity-Innate Immune Sensing and Its Posttranscriptional Regulation Mechanisms, *Cells*, 2020, **9**(7), 1701.
  - 49 M. Herrera-Barrera, R. C. Ryals, M. Gautam, A. Jozic, M. Landry, T. Korzun, M. Gupta, C. Acosta, J. Stoddard, R. Reynaga, W. Tschetter, N. Jacomino, O. Taratula, C. Sun, A. K. Lauer, M. Neuringer and G. Sahay, Peptide-guided lipid nanoparticles deliver mRNA to the neural retina of rodents and nonhuman primates, *Sci. Adv.*, 2023, **9**(2), eadd4623.
  - 50 D. Shi, S. Toyonaga and D. G. Anderson, *In Vivo* RNA Delivery to Hematopoietic Stem and Progenitor Cells via



- Targeted Lipid Nanoparticles, *Nano Lett.*, 2023, **23**(7), 2938–2944.
- 51 Web: <https://www.merck.com/news/moderna-and-merck-announce-mrna-4157-v940-an-investigational-personalized-mrna-cancer-vaccine-in-combination-with-keytruda-pembrolizumab-met-primary-efficacy-endpoint-in-phase-2b-keynote-94/>; accessed January 21st 2024.
  - 52 S. J. Tsai, N. A. Atai, M. Cacciottolo, J. Nice, A. Salehi, C. Guo, A. Sedgwick, S. Kanagavelu and S. J. Gould, Exosome-mediated mRNA delivery *in vivo* is safe and can be used to induce SARS-CoV-2 immunity, *J. Biol. Chem.*, 2021, **297**(5), 101266.
  - 53 N. Gretskeya, M. Akimov, D. Andreev, A. Zalygin, E. Belitskaya, G. Zinchenko, E. Fomina-Ageeva, I. Mikhalyov, E. Vodovozova and V. Bezuglov, Multicomponent Lipid Nanoparticles for RNA Transfection, *Pharmaceutics*, 2023, **15**(4), 1289.
  - 54 L. S. Mbatha, F. Maiyo, A. Daniels and M. Singh, Dendrimer-Coated Gold Nanoparticles for Efficient Folate-Targeted mRNA Delivery In Vitro, *Pharmaceutics*, 2021, **13**(6), 900.
  - 55 A. Suberi, M. K. Grun, T. Mao, B. Israelow, M. Reschke, J. Grundler, L. Akhtar, T. Lee, K. Shin, A. S. Piotrowski-Daspit, R. J. Homer, A. Iwasaki, H. W. Suh and W. M. Saltzman, Polymer nanoparticles deliver mRNA to the lung for mucosal vaccination, *Sci. Transl. Med.*, 2023, **15**(709), eabq0603.
  - 56 H. Shin, S. Kang, C. Won and D. H. Min, Enhanced Local Delivery of Engineered IL-2 mRNA by Porous Silica Nanoparticles to Promote Effective Antitumor Immunity, *ACS Nano*, 2023, **17**(17), 17554–17567.
  - 57 M. Segel, B. Lash, J. Song, A. Ladha, C. C. Liu, X. Jin, S. L. Mekhedov, R. K. Macrae, E. V. Koonin and F. Zhang, Mammalian retrovirus-like protein PEG10 packages its own mRNA and can be pseudotyped for mRNA delivery, *Science*, 2021, **373**(6557), 882–889.
  - 58 D. Grant-Serroukh, M. R. Hunter, R. Maeshima, A. D. Tagalakakis, A. M. Aldossary, N. Allahham, G. R. Williams, M. Edbrooke, A. Desai and S. L. Hart, Lipid-peptide nanocomplexes for mRNA delivery *in vitro* and *in vivo*, *J. Controlled Release*, 2022, **348**, 786–797.
  - 59 K. Q. Kim, *et al.*, N1-methylpseudouridine found within COVID-19 mRNA vaccines produces faithful protein products, *Cell Rep.*, 2022, **40**(9), 111300.
  - 60 E. Rivas, RNA structure prediction using positive and negative evolutionary information, *PLoS Comput. Biol.*, 2020, **16**(10), e1008387.
  - 61 M. Chamberlin, J. McGrath and L. Waskell, New RNA polymerase from *Escherichia coli* infected with bacteriophage T7, *Nature*, 1970, **228**, 227–231.
  - 62 M. Chamberlin, *et al.*, Isolation of bacterial and bacteriophage RNA polymerases and their use in synthesis of RNA *in vitro*, *Methods Enzymol.*, 1983, **101**, 540–568.
  - 63 I. H. Goldberg and M. Rabinowitz, The incorporation of 5-ribosyluracil triphosphate into RNA in nuclear extracts of mammalian cells, *Biochem. Biophys. Res. Commun.*, 1961, **6**, 394–398.
  - 64 C. J. Knudson, P. Alves-Peixoto, H. Muramatsu, C. Stotesbury, L. Tang, P. J. C. Lin, Y. K. Tam, D. Weissman, N. Pardi and L. J. Sigal, Lipid-nanoparticle-encapsulated mRNA vaccines induce protective memory CD8 T cells against a lethal viral infection, *Mol. Ther.*, 2021, **29**(9), 2769–2781.
  - 65 J. Ramos da Silva, K. Bitencourt Rodrigues and G. Formoso, *et al.*, Single immunizations of self-amplifying or non-replicating mRNA-LNP vaccines control HPV-associated tumors in mice, *Sci. Transl. Med.*, 2023, **15**(686), eabn3464.
  - 66 E. Kenjo, H. Hozumi, Y. Makita, K. A. Iwabuchi, N. Fujimoto, S. Matsumoto, M. Kimura, Y. Amano, M. Ifuku, Y. Naoe, N. Inukai and A. Hotta, Low immunogenicity of LNP allows repeated administrations of CRISPR-Cas9 mRNA into skeletal muscle in mice, *Nat. Commun.*, 2021, **12**(1), 7101.
  - 67 K. Karikó, *et al.*, Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA, *Immunity*, 2005, **23**, 165–175.

