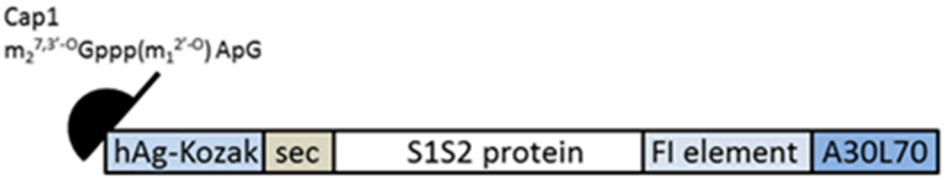
TABLE OF CONTENTS

LIST OF FIGURES

3.2.S.1.2. STRUCTURE, OMICRON (BA.4/BA.5)

The active principle in the BNT162b2 Omicron BA.4/BA.5 (herein referred to as Omicron) drug substance (DS) is a single-stranded, 5'-capped mRNA that is translated into the respective protein (the encoded antigen). Figure 3.2.S.1.2-1 illustrates the general structure of the antigen-encoding RNA, which is determined by the respective nucleotide sequence of the DNA used as template for *in vitro* RNA transcription. In addition to the codon-optimized sequence encoding the antigen, the RNA contains common structural elements optimized for mediating high RNA stability and translational efficiency (5'-cap, 5'UTR, 3'-UTR, poly(A) -tail; see below). Furthermore, an intrinsic signal peptide (sec) is part of the antigen-encoding regions and is translated as N-terminal tag.

Figure 3.2.S.1.2-1. General structure of the Omicron BA.4/BA.5 mRNA

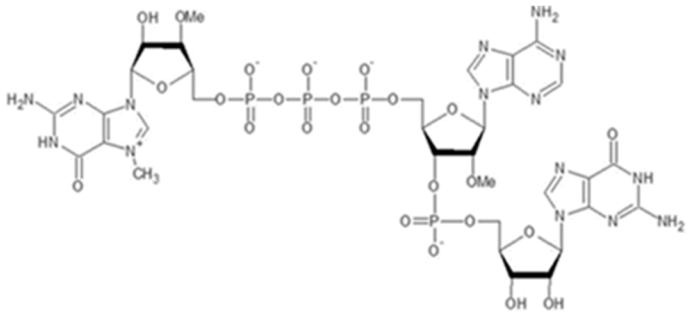


Schematic illustration of the general structure of the BNT162b2 Omicron BA.4/BA.5 mRNA drug substance with 5'-cap, 5'- and 3'untranslated regions (hAg-Kozak and FI element, respectively), coding sequence for variant of concern and intrinsic signal peptide (sec) as well as poly(A)-tail (A30L70). Individual elements are not drawn to scale compared to their respective sequence lengths.

mRNA cap

A cap1 structure $m_2^{7,3'\text{-}O}Gppp(m_1^{2'\text{-}O})ApG$ is utilized as specific capping structure at the 5′-end of the RNA drug substance (Figure 3.2.S.1.2-2).

Figure 3.2.S.1.2-2. 5′-cap analog ($m_2^{7,3'\text{-}O}Gppp(m_1^{2'\text{-}O})ApG$) for production of RNA containing a cap1 structure



The cap1 structure (i.e., containing a 2′-O-methyl group on the penultimate nucleoside of the 5′-end of the RNA chain) is incorporated into the RNA drug substance by using a respective cap analog during *in vitro* transcription. For RNAs with modified uridine nucleotides, the cap1 structure is superior to other cap structures, since cap1 is not recognized by cellular factors such as IFIT1[1] and, thus, cap1-dependent translation is not inhibited by competition with eukaryotic translation initiation factor 4E[2]. In the context of IFIT1 expression, mRNAs with a cap1 structure give higher protein expression.

In addition, use of the cap1 structure leads to low amounts of uncapped transcripts[3]. In general, the T7 Polymerase prefers a guanosine as priming nucleoside with the highest transcription efficiencies as compared to other starting nucleosides[4]. Capping structures with a guanosine moiety compete with GTP for incorporation in the mRNA resulting in uncapped transcripts. The $m_2^{7,3'\text{-}O}$Gppp($m_1^{2'\text{-}O}$)ApG cap analog rescues transcription efficiency from templates starting with adenosines, because the ApG moiety of cap1 allows transcription initiation at the second position, a guanosine, thereby giving mainly capped mRNAs.

Modified Uridine

The RNA does not contain any uridines; instead of uridine the modified N1-methylpseudouridine is used in RNA synthesis. Several reports have demonstrated that such a substitution often strongly enhances translation of *in vitro* transcribed mRNA sequences by reducing its immunogenicity[5,6,7]. Accordingly, the drug substance is synthesized in the presence of N1-methylpseudouridine triphosphate ($^{m1}\Psi$TP) instead of uridine triphosphate (UTP).

[1] Habjan M, Hubel P, Lacerda L, et al. Sequestration by IFIT1 Impairs Translation of 2′O-unmethylated Capped RNA. 2013. PLOS Pathog;9(10):e1003663

[2] Diamond MS. IFIT1: A dual sensor and effector molecule that detects non-2'-O methylated viral RNA and inhibits its translation. 2014. Cytokine Growth Factor Rev;25(5):543-50.

[3] Trilink Patent auf CC413 cap. Accessed at https://patentimages.storage.googleapis.com/4c/83/15/99418d175a3be2/WO2017053297A1.pdf

[4] Kuzmine I, Gottlieb PA, Martin CT. Binding of the priming nucleotide in the initiation of transcription by T7 RNA polymerase. 2003. J Biol Chem;278(5):2819-23.

[5] Kariko K, Muramatsu H, Welsh FA, et al. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. 2008. Mol Ther;16(11):1833-40.

[6] Andries O, Mc Cafferty S, De Smedt SC, et al. N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. 2015. J Control Release; 217:337-44.

[7] Richner JM, Himansu S, Dowd KA, et al. Modified mRNA Vaccines Protect against Zika Virus Infection. 2017. Cell;168(6):1114-25.e10

RNA sequence

The general sequence elements of the BNT162b2 Omicron BA.4/BA.5 mRNA drug substance, as depicted in Figure 3.2.S.1.2-1, are given below. The full sequence is given in Figure 3.2.S.1.2-3.

hAg-Kozak (nucleotides 1 to 53): 5'-UTR sequence of the human alpha-globin mRNA with an optimized 'Kozak sequence' to increase translational efficiency[8].

Sec (nucleotides 54 to 101, intrinsic to S protein): Sec corresponds to the intrinsic S1S2 protein signal peptide (sec), which guides translocation of the nascent polypeptide chain into the endoplasmic reticulum.

S protein omicron (nucleotides 54 to 3863; including sec 54 to 101 and stop codons 3858 to 3863): Codon-optimized sequences encoding the respective antigen of SARS-CoV-2 protein has following point mutations/deletions (reference for numbering Genbank ID QHD43416.1): T19I, ΔLPP24-26, A27S, ΔHV69-70, G142D, V213G, G339D, S371L, S373P, S375F, T376A, D405N, K417N, N440K, L452R, S477N, T478K, E484A, F486V, Q498R, N501Y, Y505H, D614G, H655Y, N679K, P681H, N764K, D796Y, Q954H, N969K, KV986-987PP.

FI element (nucleotides 3864 to 4158): The 3′-UTR is a combination of two sequence elements derived from the "amino terminal enhancer of split" (AES) mRNA (called F) and the mitochondrial encoded 12S ribosomal RNA (called I). These were identified by an *ex vivo* selection process for sequences that confer RNA stability and augment total protein expression[9].

A30L70 (nucleotides 4159 to 4268): The circular plasmid, described in Section 3.2.S.2.3 Control of Materials – Source, History and Generation of Plasmids BNT162b2 [Omicron (BA.4/BA.5) Variant], provides a template for an mRNA transcript that contains two poly(A) tracts of 30 and approximately 70 adenosine residues joined by a linker [10].

[8] Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. 1987. Nucleic Acids Res;15(20):8125-48.

[9] Orlandini von Niessen AG, Poleganov MA, Rechner C, et al. Improving mRNA-Based Therapeutic Gene Delivery by Expression-Augmenting 3′ UTRs Identified by Cellular Library Screening. 2019. Mol Ther;27(4):1-13.

[10] BioNTech Patent auf STABILISIERUNG VON DNA-SEQUENZEN ZUR POLY(A)SEQUENZ-CODIERUNG. Accessed at https://data.epo.org/publication-server/pdf-document?pn=3167059&ki=B1&cc=EP&pd=20190626

Figure 3.2.S.1.2-3. RNA nucleotide Sequence of the BNT162b2 (Omicron (BA.4/BA.5) Variant) drug substance

The nucleotide sequence 5'→3':

Cap-

```
AGAAUAAACU AGUAUUCUUC UGGUCCCCAC AGACUCAGAG AGAACCCGCC ACCAUGUUCG   60
UGUUCCUGGU GCUGCUGCCU CUGGUGUCCA GCCAGUGUGU GAACCUGAUC ACCAGAACAC  120
AGUCAUACAC CAACAGCUUU ACCAGAGGCG UGUACUACCC CGACAAGGUG UUCAGAUCCA  180
GCGUGCUGCA CUCUACCCAG GACCUGUUCC UGCCUUUCUU CAGCAACGUG ACCUGGUUCC  240
ACGCCAUCUC CGGCACCAAU GGCACCAAGA GAUUCGACAA CCCCGUGCUG CCCUUCAACG  300
ACGGGGUGUA CUUUGCCAGC ACCGAGAAGU CCAACAUCAU CAGAGGCUGG AUCUUCGGCA  360
CCACACUGGA CAGCAAGACC CAGAGCCUGC UGAUCGUGAA CAACGCCACC AACGUGGUCA  420
UCAAAGUGUG CGAGUUCCAG UUCUGCAACG ACCCCUUCCU GGACGUCUAC UACCACAAGA  480
ACAACAAGAG CUGGAUGGAA AGCGAGUUCC GGGUGUACAG CAGCGCCAAC AACUGCACCU  540
UCGAGUACGU GUCCCAGCCU UUCCUGAUGG ACCUGGAAGG CAAGCAGGGC AACUUCAAGA  600
ACCUGCGCGA GUUCGUGUUU AAGAACAUCG ACGGCUACUU CAAGAUCUAC AGCAAGCACA  660
CCCCUAUCAA CCUCGGCCGG GAUCUGCCUC AGGGCUUCUC UGCUCUGGAA CCCCUGGUGG  720
AUCUGCCCAU CGGCAUCAAC AUCACCCGGU UUCAGACACU GCUGGCCCUG CACAGAAGCU  780
ACCUGACACC UGGCGAUAGC AGCAGCGGAU GGACAGCUGG UGCCGCCGCU UACUAUGUGG  840
GCUACCUGCA GCCUAGAACC UUCCUGCUGA AGUACAACGA GAACGGCACC AUCACCGACG  900
CCGUGGAUUG UGCUCUGGAU CCUCUGAGCG AGACAAAGUG CACCCUGAAG UCCUUCACCG  960
UGGAAAAGGG CAUCUACCAG ACCAGCAACU UCCGGGUGCA GCCCACCGAA UCCAUCGUGC 1020
GGUUCCCCAA UAUCACCAAU CUGUGCCCCU UCGACGAGGU GUUCAAUGCC ACCAGAUUCG 1080
CCUCUGUGUA CGCCUGGAAC CGGAAGCGGA UCAGCAAUUG CGUGGCCGAC UACUCCGUGC 1140
UGUACAACUU CGCCCCCUUC UUCGCAUUCA AGUGCUACGG CGUGUCCCCU ACCAAGCUGA 1200
ACGACCUGUG CUUCACAAAC GUGUACGCCG ACAGCUUCGU GAUCCGGGGA AACGAAGUGC 1260
GGCAGAUUGC CCCUGGACAG ACAGGCAACA UCGCCGACUA CAACUACAAG CUGCCCGACG 1320
ACUUCACCGG CUGUGUGAUU GCCUGGAACA GCAACAAGCU GGACUCCAAA GUCGGCGGCA 1380
ACUACAAUUA CAGGUACCGG CUGUUCCGGA AGUCCAAUCU GAAGCCCUUC GAGCGGGACA 1440
UCUCCACCGA GAUCUAUCAG GCCGGCAACA AGCCUUGUAA CGGCGUGGCA GGCGUGAACU 1500
GCUACUUCCC ACUGCAGUCC UACGGCUUUA GGCCCACAUA CGGCGUGGGC CACCAGCCCU 1560
ACAGAGUGGU GGUGCUGAGC UUCGAACUGC UGCAUGCCCC UGCCACAGUG UGCGGCCCUA 1620
AGAAAAGCAC CAAUCUCGUG AAGAACAAAU GCGUGAACUU CAACUUCAAC GGCCUGACCG 1680
GCACCGGCGU GCUGACAGAG AGCAACAAGA AGUUCCUGCC AUUCCAGCAG UUUGGCCGGG 1740
AUAUCGCCGA UACCACAGAC GCCGUUAGAG AUCCCCAGAC ACUGGAAAUC CUGGACAUCA 1800
CCCCUUGCAG CUUCGGCGGA GUGUCUGUGA UCACCCCUGG CACCAACACC AGCAAUCAGG 1860
```

```
UGGCAGUGCU GUACCAGGGC GUGAACUGUA CCGAAGUGCC CGUGGCCAUU CACGCCGAUC 1920
AGCUGACACC UACAUGGCGG GUGUACUCCA CCGGCAGCAA UGUGUUUCAG ACCAGAGCCG 1980
GCUGUCUGAU CGGAGCCGAG UACGUGAACA AUAGCUACGA GUGCGACAUC CCCAUCGGCG 2040
CUGGAAUCUG CGCCAGCUAC CAGACACAGA CAAAGAGCCA CCGGAGAGCC AGAAGCGUGG 2100
CCAGCCAGAG CAUCAUUGCC UACACAAUGU CUCUGGGCGC CGAGAACAGC GUGGCCUACU 2160
CCAACAACUC UAUCGCUAUC CCCACCAACU UCACCAUCAG CGUGACCACA GAGAUCCUGC 2220
CUGUGUCCAU GACCAAGACC AGCGUGGACU GCACCAUGUA CAUCUGCGGC GAUUCCACCG 2280
AGUGCUCCAA CCUGCUGCUG CAGUACGGCA GCUUCUGCAC CCAGCUGAAA AGAGCCCUGA 2340
CAGGGAUCGC CGUGGAACAG GACAAGAACA CCCAAGAGGU GUUCGCCCAA GUGAAGCAGA 2400
UCUACAAGAC CCCUCCUAUC AAGUACUUCG GCGGCUUCAA UUUCAGCCAG AUUCUGCCCG 2460
AUCCUAGCAA GCCCAGCAAG CGGAGCUUCA UCGAGGACCU GCUGUUCAAC AAAGUGACAC 2520
UGGCCGACGC CGGCUUCAUC AAGCAGUAUG GCGAUUGUCU GGGCGACAUU GCCGCCAGGG 2580
AUCUGAUUUG CGCCCAGAAG UUUAACGGAC UGACAGUGCU GCCUCCUCUG CUGACCGAUG 2640
AGAUGAUCGC CCAGUACACA UCUGCCCUGC UGGCCGGCAC AAUCACAAGC GGCUGGACAU 2700
UUGGAGCAGG CGCCGCUCUG CAGAUCCCCU UUGCUAUGCA GAUGGCCUAC CGGUUCAACG 2760
GCAUCGGAGU GACCCAGAAU GUGCUGUACG AGAACCAGAA GCUGAUCGCC AACCAGUUCA 2820
ACAGCGCCAU CGGCAAGAUC CAGGACAGCC UGAGCAGCAC AGCAAGCGCC CUGGGAAAGC 2880
UGCAGGACGU GGUCAACCAC AAUGCCCAGG CACUGAACAC CCUGGUCAAG CAGCUGUCCU 2940
CCAAGUUCGG CGCCAUCAGC UCUGUGCUGA ACGAUAUCCU GAGCAGACUG GACCCUCCUG 3000
AGGCCGAGGU GCAGAUCGAC AGACUGAUCA CAGGCAGACU GCAGAGCCUC CAGACAUACG 3060
UGACCCAGCA GCUGAUCAGA GCCGCCGAGA UUAGAGCCUC UGCCAAUCUG GCCGCCACCA 3120
AGAUGUCUGA GUGUGUGCUG GGCCAGAGCA AGAGAGUGGA CUUUUGCGGC AAGGGCUACC 3180
ACCUGAUGAG CUUCCCUCAG UCUGCCCCUC ACGGCGUGGU GUUUCUGCAC GUGACAUAUG 3240
UGCCCGCUCA AGAGAAGAAU UUCACCACCG CUCCAGCCAU CUGCCACGAC GGCAAAGCCC 3300
ACUUUCCUAG AGAAGGCGUG UUCGUGUCCA ACGGCACCCA UUGGUUCGUG ACACAGCGGA 3360
ACUUCUACGA GCCCCAGAUC AUCACCACCG ACAACACCUU CGUGUCUGGC AACUGCGACG 3420
UCGUGAUCGG CAUUGUGAAC AAUACCGUGU ACGACCCUCU GCAGCCCGAG CUGGACAGCU 3480
UCAAAGAGGA ACUGGACAAG UACUUUAAGA ACCACACAAG CCCCGACGUG GACCUGGGCG 3540
AUAUCAGCGG AAUCAAUGCC AGCGUCGUGA ACAUCCAGAA AGAGAUCGAC CGGCUGAACG 3600
AGGUGGCCAA GAAUCUGAAC GAGAGCCUGA UCGACCUGCA AGAACUGGGG AAGUACGAGC 3660
AGUACAUCAA GUGGCCCUGG UACAUCUGGC UGGGCUUUAU CGCCGGACUG AUUGCCAUCG 3720
UGAUGGUCAC AAUCAUGCUG UGUUGCAUGA CCAGCUGCUG UAGCUGCCUG AAGGGCUGUU 3780
GUAGCUGUGG CAGCUGCUGC AAGUUCGACG AGGACGAUUC UGAGCCCGUG CUGAAGGGCG 3840
UGAAACUGCA CUACACAUGA UGACUCGAGC UGGUACUGCA UGCACGCAAU GCUAGCUGCC 3900
CCUUUCCCGU CCUGGGUACC CCGAGUCUCC CCCGACCUCG GGUCCCAGGU AUGCUCCCAC 3960
CUCCACCUGC CCCACUCACC ACCUCUGCUA GUUCCAGACA CCUCCCAAGC ACGCAGCAAU 4020
GCAGCUCAAA ACGCUUAGCC UAGCCACACC CCCACGGGAA ACAGCAGUGA UUAACCUUUA 4080
GCAAUAAACG AAAGUUUAAC UAAGCUAUAC UAACCCCAGG GUUGGUCAAU UUCGUGCCAG 4140
CCACACCCUG GAGCUAGCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAGC AUAUGACUAA 4200
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 4260
AAAAAAAA                                                         4268
```

Sequence length: 4269, which includes “Cap-“ to denote the presence of the 5’-cap analog
G: 1058; C: 1305; A: 1108; U: 797
A = Adenine; C = Cytosine; G = Guanine; U = N1-methylpseudouridine