

## Perspective

# News from Mars: Two-Tier Paradox, Intracellular PCR, Chimeric Junction Shift, Dark Matter mRNA and Other Remarkable Features of Mammalian RNA-Dependent mRNA Amplification. Implications for Alzheimer's Disease, RNA-Based Vaccines and mRNA Therapeutics

Vladimir Volloch<sup>1\*</sup> and Sophia Rits-Volloch<sup>2,3</sup>

<sup>1</sup>Department of Developmental Biology, Harvard School of Dental Medicine, USA

<sup>2</sup>Division of Molecular Medicine, Children's Hospital, USA

<sup>3</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, USA

## Abstract

Molecular Biology, a branch of science established to examine the flow of information from “letters” encrypted into DNA structure to functional proteins, was initially defined by a concept of DNA-to-RNA-to-Protein information movement, a notion termed the Central Dogma of Molecular Biology. RNA-dependent mRNA amplification, a novel mode of eukaryotic protein-encoding RNA-to-RNA-to-Protein genomic information transfer, constitutes the extension of the Central Dogma in the context of mammalian cells. It was shown to occur in cellular circumstances requiring exceptionally high levels of production of specific polypeptides, e.g. globin chains during erythroid differentiation or defined secreted proteins in the context of extracellular matrix deposition. Its potency is reflected in the observed cellular levels of the resulting amplified mRNA product: At the peak of the erythroid differentiation, for example, the amount of globin mRNA produced in the amplification pathway is about 1500-fold higher than the amount of its conventionally generated counterpart in the same cells. The cellular enzymatic machinery at the core of this process, RNA-dependent RNA polymerase activity (RdRp), albeit in a non-conventional form, was shown to be constitutively and ubiquitously present, and RNA-dependent RNA synthesis (RdRs) appeared to regularly occur, in mammalian cells. Under most circumstances, the mammalian RdRp activity produces only short antisense RNA transcripts. Generation of complete antisense RNA transcripts and amplification of mRNA molecules require the activation of inducible components of the mammalian RdRp complex. The mechanism of such activation is not clear. The present article suggests that it is triggered by a variety of cellular stresses and occurs in the context of stress responses in general and within the framework of the integrated stress response (ISR) in particular. In this process, various cellular stresses activate, in a stress type-specific manner, defined members of the mammalian translation initiation factor 2 $\alpha$ , eIF2 $\alpha$ , kinase family: PKR, GCN2, PERK and HRI. Any of these kinases, in an activated form, phosphorylates eIF2 $\alpha$ . This results in suppression of global cellular protein synthesis but also in activation of expression of select group of transcription factors including ATF4, ATF5 and CHOP. These transcription factors either function as inducible components of the RdRp complex or enable their expression. The assembly of the competent RdRp complex activates mammalian RNA-dependent mRNA amplification, which appears to be a two-tier process. Tier One is a “chimeric” pathway, named so because it results in an amplified chimeric mRNA molecule containing a fragment of the antisense RNA strand at its 5' terminus. Tier Two further amplifies one of the two RNA end products of the chimeric pathway and constitutes the physiologically occurring intracellular polymerase chain reaction, iPCR. Depending on the structure of the initial mRNA amplification progenitor, the chimeric pathway, Tier One, may result in multiple outcomes including chimeric mRNA that produces either a polypeptide identical to the original, conventional mRNA progenitor-encoded protein or only its C-terminal fragment, CTF. The chimeric RNA end product of Tier One may also produce a polypeptide that is non-contiguously encoded in the genome, activate translation from an open reading frame, which is “silent” in a conventionally transcribed mRNA, or initiate an abortive translation. In sharp contrast, regardless of the outcome of Tier One, the mRNA end product of Tier Two of mammalian mRNA amplification, the iPCR pathway, always produces a polypeptide identical to a conventional mRNA progenitor-encoded protein. This discordance is referred to as the Two-Tier Paradox and discussed in detail in the present article. On the other hand, both Tiers are similar

**Citation:** Volloch V, Rits-Volloch S. News from Mars: Two-Tier Paradox, Intracellular PCR, Chimeric Junction Shift, Dark Matter mRNA and Other Remarkable Features of Mammalian RNA-Dependent mRNA Amplification. Implications for Alzheimer's Disease, RNA-Based Vaccines and mRNA Therapeutics. *Ann Integr Mol Med.* 2021; 2(1): 1009.

**Copyright:** © 2021 Vladimir Volloch

**Publisher Name:** Medtext Publications LLC

**Manuscript compiled:** Feb 4<sup>th</sup>, 2021

**\*Corresponding author:** Vladimir Volloch, Department of Developmental Biology, Harvard School of Dental Medicine, USA, E-mail: vladimir.volloch@gmail.com; vladimir\_volloch@hms.harvard.edu

in that they result in heavily modified mRNA molecules resistant to reverse transcription, undetectable by reverse transcription-based methods of sequencing and therefore constituting a proverbial “Dark Matter” mRNA, despite being highly ubiquitous. It appears that in addition to their other functions, the modifications of the amplified mRNA render it compatible, unlike the bulk of cellular mRNA, with phosphorylated eIF2 $\alpha$  in translation, implying that in addition to being extraordinarily abundant due to the method of its generation, amplified mRNA is also preferentially translated under the ISR conditions, thus augmenting the efficiency of the amplification process. The vital importance of powerful mechanisms of amplification of protein-encoding genomic information in normal physiology is self-evident. Their malfunctions or misuse appear to be associated with two types of abnormalities, the deficiency of a protein normally produced by these mechanisms and the mRNA amplification-mediated overproduction of a protein normally not generated by such a process. Certain classes of beta-thalassemia exemplify the first type, whereas the second type is represented by overproduction of beta-

amyloid in Alzheimer's disease. Both examples are discussed in detail in the present article, which summarizes and systematizes our current understanding of the field and describes two categories of reporter constructs, one for the chimeric Tier of mRNA amplification, another for the iPCR pathway; both reporter types are essential for elucidating underlying molecular mechanisms. It also suggests, in light of the recently demonstrated feasibility of RNA-based vaccines, that the targeted intracellular amplification of exogenously introduced amplification-eligible antigen-encoding mRNAs via the induced or naturally occurring RNA-dependent mRNA amplification pathway could be of substantial benefit in triggering a fast and potent immune response and instrumental in the development of future vaccines. Similar approaches can also be effective in achieving efficient and sustained expression of exogenous mRNA in mRNA therapeutics.

**Keywords:** Mammalian RNA-dependent mRNA amplification; The chimeric mRNA amplification pathway; Intracellular PCR mRNA amplification pathway; Two-Tier paradox; Integrated stress response; Asymmetric mRNA amplification; Alzheimer's disease; Beta-thalassemia; Reporters for mammalian mRNA amplification; Exogenous RNA-based vaccines; mRNA therapeutics

## Introduction

The initial indications of mammalian RNA-dependent mRNA amplification were obtained in studies of the kinetics of globin mRNA synthesis in differentiating erythroid cells, which strongly suggested the occurrence of cytoplasmic *de novo* production of globin mRNA [1]. In these experiments, unexpectedly, in very short labeling pulses, the bulk of radioactivity incorporated into globin mRNA and anticipated to be localized solely in the nuclei, was accounted for in the cytoplasm [1]. The labeled molecules appeared to be of a genuine cytoplasmic origin. Indeed, in the nuclei, globin RNA is synthesized in the form of a 15S precursor, which is processed to its mature 9S size with a half-life of about 10 minutes. Following the seconds-long pulse, only mature size radioactive globin RNA was seen in the cytoplasm and only precursor-size molecules in the nuclei [1]. In such a pulse-labeling, cytoplasmic globin RNA was labeled uniformly and not by the end addition to preexisting molecules, consistent with *de novo* synthesis utilizing mature cytoplasmic globin mRNA as the initial template. Moreover, a high concentration of actinomycin D completely inhibited incorporation into nuclear globin RNA but had little effect on globin RNA labeling in the cytoplasm suggesting the involvement of a distinct enzymatic activity in cytoplasmic RNA synthesis [1]. Two additional lines of evidence suggested a possible involvement of mammalian RNA-dependent RNA polymerase, RdRp, an enzymatic activity previously detected in and isolated from rabbit reticulocytes [2]. One was the detection of globin antisense RNA, of a size comparable with globin mRNA and apparently containing a poly(U) segment, presumably a complement of the 3'-terminal poly(A) of globin mRNA. Another was the detection of the RdRp enzymatic activity in erythroid cells at the levels greatly exceeding those seen in control undifferentiated cells [1].

Findings described above were substantiated more directly in studies with cytoplasts, i.e. erythroid cells enucleated *in vitro* by mechanical manipulations [3]. Utilization of cytoplast preparations, shown to be free of contaminating nucleated cells, allowed definitive corroboration of the occurrence of cytoplasmic synthesis of both positive- and negative-strand globin RNA. Cytoplasts were further selectively permeabilized enabling the use of substances that do not readily cross the plasma membrane. By using mercury-substituted CTP as a substrate, it was possible to physically separate newly synthesized globin RNA by adsorption to thiol-agarose and show that the cytoplasmic incorporation of radioactive precursors into both sense and antisense globin RNA constituted a *de novo* synthesis. The system also allowed experiments establishing that synthesis of cytoplasmic globin RNA required the presence of  $Mg^{++}$ , was inhibited by  $Mn^{++}$  and showed no response to  $Zn^{++}$  in contrast to viral RNA replicases. Cytoplasmic globin RNA synthesis was resistant to actinomycin D, alpha-amanitin and rifampicin but inhibited by the rifampicin derivative AF/ABDP. The results also established that a *de novo* DNA synthesis was not involved in the observed phenomenon: Synthesis of RNA occurred in permeabilized cytoplasts not only without the addition of dNTPs but also when ddNTP was added.

The apparent presence in antisense globin RNA of a poly(U) region, presumably a transcript of the 3'-terminal poly(A) segment of mRNA, indicated that synthesis of the antisense strand initiates at the poly(A) of its mRNA template. The precedence for initiation of antisense RNA synthesis within the poly(A) is well established for viral systems where it is primed by a uridylated protein [4]; in mammalian systems, different mechanisms for initiation of antisense RNA synthesis could be involved. A mechanism for the second stage of RNA-dependent globin mRNA synthesis, the initiation of synthesis of sense globin RNA on the antisense RNA strand template, was suggested in a study of the generation of a largely double-stranded murine beta-globin cDNA by RNA-dependent DNA polymerase (RdDp), a process that involves formation of covalently linked antisense and sense strands in a hairpin configuration [5]. This study examined a sequence of molecular events, particularly the self-priming, that initiates synthesis of the sense strand. It was shown that upon completion of reverse transcription of globin mRNA and the removal of the RNA template by RNase H activity associated with RdDp, the 3' terminus of the antisense strand snaps back to form a stable double-stranded self-priming structure which is extended by RdDp to generate the sense strand. The self-priming event is enabled by strong complementarity of 14-nucleotide-long 3'-terminal segment of the antisense strand (3'-Terminal Complementary Element, TCE) with an internal segment of the same molecule (Internal Complementary Element, ICE) corresponding to a portion of the 5'untranslated region, 5'UTR, of mRNA located just upstream of the translation start site. Surprisingly and informatively, the strong complementarity between the terminal and the internal complementary elements occurs within the antisense but not the sense strand. This is because A:C mismatches on the sense strand correspond to stable T/U:G base pairs on the antisense strand. The self-priming-enabling complementarity within the beta-globin antisense RNA strand was shown to be highly preserved throughout a vast evolutionary distance from marsupials to humans [5], with conservation of not only the occurrence of complementary elements but also their positions within the antisense RNA segment corresponding to the 5'UTR of mRNA; although nucleotide sequences diverged substantially during evolution, the complementary relationship of the 3' terminal and the internal elements of the antisense RNA, as well as the position of the internal element within a segment corresponding to the 5'UTR of mRNA, remained preserved, strongly suggesting that their functionality is physiologically relevant.

If cellular RdRp activity utilizes the same self-priming arrangement as RdDp to extend the antisense globin RNA strand into a sense-orientation molecule, the resulting hairpin-structured RNA would have to be cleaved to separate the globin RNA strands. Where the cleavage may occur within the hairpin structure was indicated by a detailed characterization of the antisense globin RNA in mouse erythroid tissues [6]. This study made use of a multistep procedure in which a molecular tag is attached to cellular RNA by ligation with a defined ribooligonucleotide. The act of ligation preserves the termini

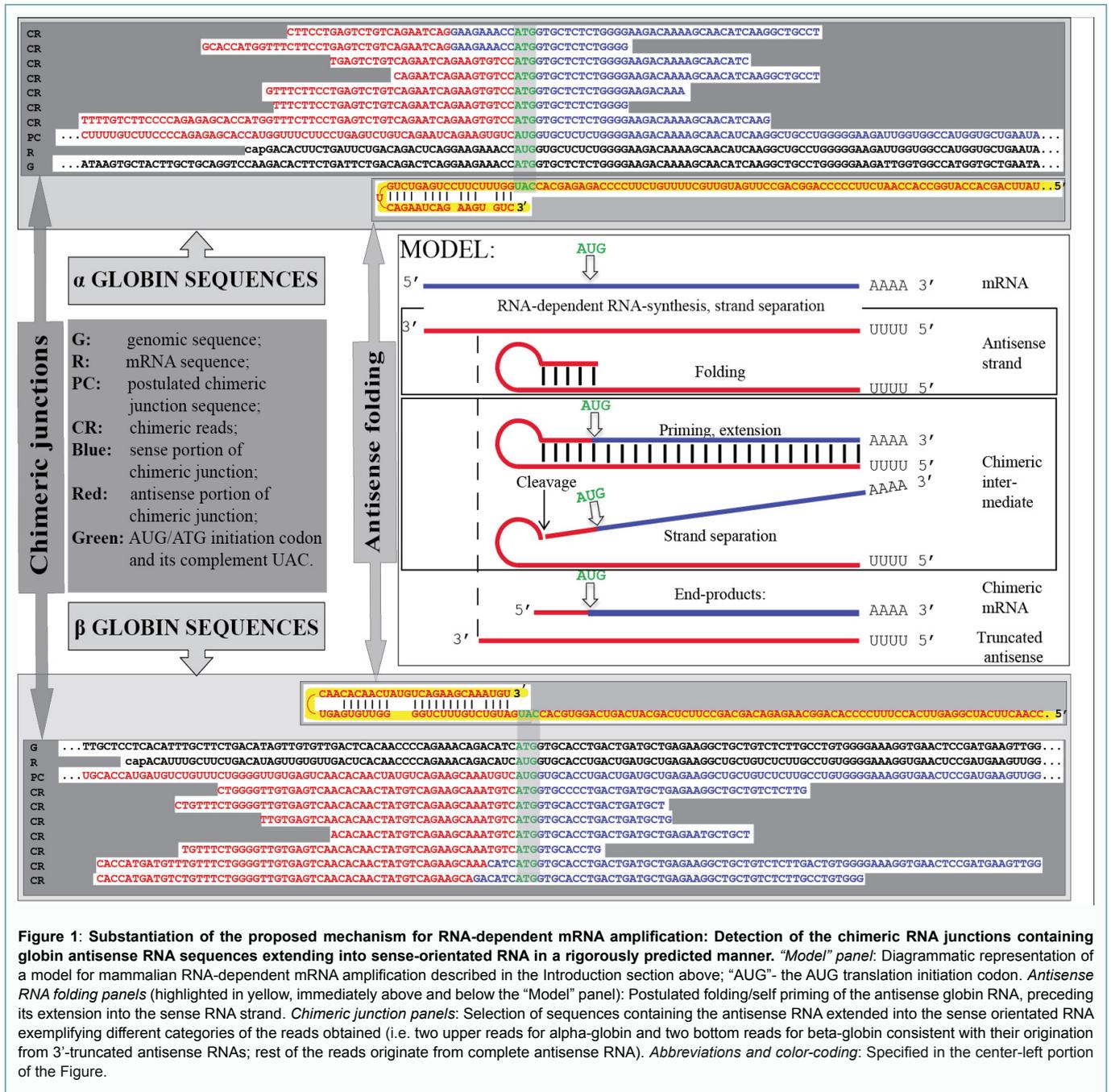
of RNA molecules, which become the junctions between cellular RNA and ligated ribooligonucleotide. It also unambiguously preserves the identity of cellular RNA as a sense or antisense molecule through all subsequent manipulations culminating with nucleotide sequencing of tag-containing globin clones [6,7]. This approach resulted in identification and characterization of antisense beta-globin RNA molecules in mouse erythroid tissues [6]. The antisense RNA was shown to be fully complementary to spliced globin mRNA, indicative of the template/transcript relationship. At the 5' end, antisense globin RNA terminates with a uridylyte stretch, reflecting the presence of the poly(A) at the 3' end of the sense globin RNA and indicating that synthesis of the former initiates at the 3'-terminal poly(A) region of the latter. With respect to the structure of their 3' termini, the detected antisense globin RNA could be divided into two classes of interest. One class represented a minor population and consisted of full-size antisense molecules corresponding precisely to globin mRNA. The other class represented the major globin antisense RNA population and consisted of 3'-truncated molecules. The truncation was not random; 3'-truncated antisense RNA molecules lacked predominantly fourteen 3'-terminal nucleotides, the same 14-nucleotide segment that constitutes the 3'-terminal complementary element required for the formation of globin antisense RNA self-priming structure.

A compendium of observations and considerations described above suggested a model for RNA-dependent amplification of mammalian mRNA [6] whose depiction is embedded in Figure 1 ("model" panel) below. In this postulated process, RdRp transcribes mature spliced cytoplasmic mRNA. Such transcription initiates at the 3' poly(A) region of mRNA and produces the antisense RNA containing poly(U) at the 5' end and terminating at the 3' end with a complement of the 5' terminus of the mRNA molecule. The subsequent transcription of the antisense RNA was postulated to occur via self-priming and extension of its 3' terminus. Cleavage within a single-stranded loop of the resulting hairpin-structured molecule and the unwinding of a double-stranded structure would then produce a 5'-truncated sense strand RNA terminating with the poly(A) at the 3' end and containing an antisense segment at the 5' terminus, and a 3'-truncated antisense RNA strand. If self-priming occurs within a segment corresponding to the 5'UTR of mRNA, the resulting sense RNA component would contain the entire protein-coding information of the original mRNA and could be translated into a polypeptide indistinguishable from the conventionally produced protein. In the case of globin mRNA amplification, the observed predominant 14-nucleotide 3'-terminal truncation of the antisense globin RNA strongly suggested that the cleavage occurs mainly at the 3' end of the single-stranded loop.

At the core of the mechanism proposed in [6] and described above is a distinctive extension of the self-primed antisense RNA strand. Therefore, the detection of the major recognizable attribute of this mechanism, an antisense RNA molecule extended into a sense strand in a rigorously predicted and uniquely defined manner, would provide conclusive evidence for the proposed mechanism of the mammalian mRNA amplification process. Two types of postulated RNA molecules would, due to their intrinsic structure, contain such evidential information. One is the largely double-stranded/hairpin-structured chimeric intermediate consisting of a sense and an antisense strands covalently joined in a precisely defined way, a direct precursor of the amplified mRNA end product. The other is the chimeric RNA end product containing a defined antisense RNA fragment joined to a predictably 5'-truncated sense RNA at its

5' terminus. Due to their transient nature, chimeric intermediates could be a rarity, but the chimeric end product could be expected to represent the majority of amplified mRNA sequences. In both cases, the region of interest, an "identifier" of the amplification process, is the junction between the sense and the antisense components. The occurrence of such chimeric sense/antisense junction sequences involving both alpha- and beta-globin mRNAs was, therefore, investigated in murine erythroid tissues using next generation RNA sequencing [8,9]. A selection of chimeric fragments detected for both alpha- and beta-globin RNA is presented in Figure 1, together with alpha- and beta-globin genomic reference sequences, conventional globin mRNA sequences, postulated chimeric sequences, and projected folded antisense sequences whose 3' extension would generate predicted chimeric sequences. Two Types of chimeric reads were obtained for both alpha- and beta-globin RNA sequences. Type One reads represented the predicted chimeric junction resulting from the extension of the full-size self-primed antisense RNA. The other Type of reads (two upper reads for alpha-globin and two bottom reads for beta-globin in "Chimeric junctions" panels of Figure 1) represented antisense/sense junction sequences consistent with their origination from a 3'-truncated antisense RNA retaining a portion of its 3'-terminal complementary element apparently sufficient for the self-priming and extension. Whereas the occurrence of the Type One of chimeric junctions was anticipated, the detection of the second Type was unexpected. Equally unexpected was the observation that the antisense portions of the Type One alpha-globin chimeric junctions uniformly contained at their 3' termini the additional "C" evidently not encoded in the genome. Most startling was the observation that all detected chimeric junctions appeared to originate from chimeric RNA intermediates; not a single chimeric RNA end product sequence read was identified despite its anticipated domination among globin RNA sequences. As the research progressed, it became apparent that the observed unanticipated features reflect various mechanistic aspects of the mRNA amplification process, which are addressed below.

The observations described above raised a question: Is mammalian RNA-dependent mRNA amplification a special occurrence limited to extreme circumstances of terminal erythroid differentiation or is it a general physiological phenomenon? This question was addressed [10] by testing for the occurrence of RNA-dependent amplification of mRNA encoding extracellular matrix proteins that are abundantly produced throughout the development and homeostasis, arguably the best indicator of the potential range and scope of this phenomenon. As shown in Figure 2, the major identifiers of RNA-dependent amplification of mRNA encoding  $\alpha 1$ ,  $\beta 1$ , and  $\gamma 1$  chains of laminin were indeed detected in mouse tissues producing large quantities of extracellular matrix proteins [10]. Multiple types of chimeric junctions observed for each of three laminin chains-encoding RNAs are consistent with their origin from conventional mRNAs transcribed from normally occurring multiple Transcription Start Sites (TSS) of TATA-less laminin genes. Some of these variations could also be due to the re-use of the antisense components of processed chimeric intermediates and the consequent "chimeric junction shift" discussed below. All unanticipated features of the mRNA amplification process seen with globin RNAs were also observed with laminin RNAs. The results obtained, therefore, established RNA-dependent mRNA amplification as a general physiological phenomenon in mammalian cells. This process, its features, as well as its mechanistic and regulatory aspects, are discussed in details in the following sections below.



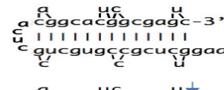
## Tier One of Mammalian RNA-Dependent mRNA Amplification: The Chimeric Pathway

The chimeric pathway of mammalian RNA-dependent mRNA amplification, as currently understood, is presented diagrammatically in Figure 3 and can be briefly (details in “Mechanistic and Regulatory Aspects” sections below) summarized as follows. The amplification process starts with synthesis of the antisense complement of a conventional, gene-transcribed mRNA by the RdRp complex. It commences with the 5'-terminal poly(U) transcribed from the 3'-terminal poly(A) of an mRNA and produces a complete antisense RNA molecule. A helicase complex then separates strands of the double-stranded sense/antisense RNA structure. Upon separation, the sense RNA molecule becomes available to serve again as a template for the antisense RNA synthesis, whereas the disengaged antisense RNA

strand forms a self-priming structure and its 3' terminus is extended by RdRp into the sense strand RNA, terminating with the poly(A) segment at the 3' end. The resulting hairpin structure is resolved by a helicase complex invoked above, starting at the linear end of the double-stranded region; upon reaching the single-stranded loop of the hairpin structure, the helicase, or an associated activity, cleaves the RNA molecule at the 3' end of the loop. This generates two end products of the chimeric mRNA amplification pathway: The antisense RNA truncated at the 3' end and the chimeric RNA molecule consisting of a 5'-truncated sense RNA and a covalently attached, in a 5' to 3' orientation, segment of the antisense RNA (its cleaved-off 3'-terminal portion). Thus, the manner of generation and processing of the chimeric intermediate leads to both the loss and the acquisition in the resulting chimeric RNA end product: The loss of a portion of conventionally produced mRNA and the acquisition, via transfer

A. 5'ACUCGCGCAGUGCCUGAGCGAGCAGCGGGCAGACCUUCCCCAGGAGCGCAGGGAGCGGCGGCGACAAC **AUG**CGCGGCAGCGGCACG...3'

B. 3'ugagcgcgucacggacacucgcucgugcccgucuggaaggggucuccgcgucuccucgcccgcgcguguu **uac**gcgccgucgcccugc...5'

C. 

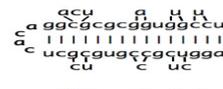
D. 

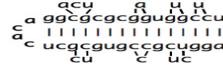
E. 5'...cgcccgtgctcgctcacaggcactgcgcgagtcCTTCCCCAGGAGCGCAGGGAGCGGCGGCGACAAC **ATG**CGCGGCAGCGGCACG...3'

F. 5'-cgcccgtgctcgctcacaggcactgcgcgagtcCTTCCCCAGGAGCGCAGGGAGCGGCGGCGACAAC **ATG**CGCGGCAGCGGCACG...3'

A. 5'GGAAGACACACUCGCGCAGUGCCUGAGCGAGCAGCGGGCAGACCUUCCCCAGGAGCGCAGGGAGCGGCGGCGACAAC **AUG**CGCGG...3'

B. 3'ccuucucgugugagcgcgucacggacacucgcucgugcccgucuggaaggggucuccgcgucuccucgcccgcgcguguu **uac**gcgcc...5'

C. 

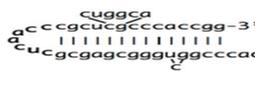
D. 

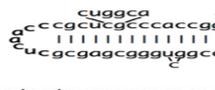
E. 5'...gaggtctcgcccgtgctcgctcacaggcactgcgcgagtgctcttccCGAGGAGCGCAGGGAGCGGCGGCGACAAC **AUG**CGCGG...3'

F. 5'...gaggtctcgcccgtgctcgctcacaggcactgcgcgagtgctcttccCGAGGAGCGCAGGGAGCGGCGGCGACAAC **AUG**CGCGG...3'

A. 5'CCGGUGGGCUGCCAGGAGCGGGUUGAGCGCUCGCCAGCCGGGUGAGGAGAAAGUAGUUAAGCGACUUGACCCCUUCC... **AUG**...3'

B. 3'ggccaccggacggucuccgcccacucgcgagcgggucggccacuccucuuuucaucauuucgcugaacugggggaagg... **uac**...5'

C. 

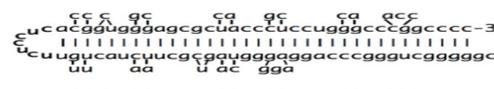
D. 

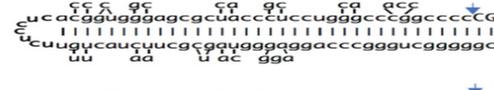
E. 5'...ttctcctcaccggcgtggggcgagcgctcaaccggctcctggcagcccaccggGTGAGGAGAAAGTAGTTAAGCGAC... **AUG**...3'

F. 5'...ttctcctcaccggcgtggggcgagcgctcaaccggctcctggcagcccaccggGTGAGGAGAAAGTAGTTAAGCGAC...3'

A. 5'GGGCCGGUGGGUCAGGAGCGGGUUGAGCGCUCGCCAGCCGGGUGAGGAGAAAGUAGUUAAGCGACUUGACCCCU... **AUG**...3'

B. 3'cccggccaccggacggucuccgcccacucgcgagcgggucggccacuccucuuuucaucauuucgcugaacuggggga... **uac**...5'

C. 

D. 

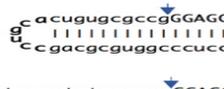
E. 5'...gctggggcgagcgctcaaccggctcctggcagcccaccggccccCGTTCCTCGTGGGAGCGGCAGGAA **ATG**GAAGGGCCCTCTCCT...3'

F. 5'...gctggggcgagcgctcaaccggctcctggcagcccaccggccccCGTTCCTCGTGGGAGCGGCAGGAA **ATG**GAAGGGCCCTCTCCT...3'

A. 5'CGGCGCACAGUGCAGGUCGCGCACCCGGGAGGUAGCGGAGGCAGCGCAUCUUGGUCGCGACGCCACCCAUUGGCUCUGCGU... **AUG**...3'

B. 3'gcccgcgugucacgucgcgagcgcguggccuccaucgcucccgucgcgcuagaaccgagccugcggguggguagccgagacgcga... **uac**...5'

C. 

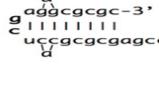
D. 

E. 5'...cggtgcgagcctgcaactgtgcccggGAGGTAGCGGAGGAGCGGATCTTGGCTCGGACGCCACCCATCGGCTCTGCGTCCG...3'

F. 5'-cggtgcgagcctgcaactgtgcccggGAGGTAGCGGAGGAGCGGATCTTGGCTCGGACGCCACCCATCGGCTCTGCGTCCG...3'

A. 5'GCGCGCUCUCGAGUGCGCGCUCGGAAGUCGGGGUCGGCGCACAGUGCAGGUCGCGACCCGGGAGGUAGCGGAGGCAGCGCG... **AUG**...3'

B. 3'cgcgcgagagcucacgcgagccuucagccccagcccgugucacgucgcgagcggccuccaucgcucccgucgccc... **uac**...5'

C. 

D. 

E. 5'...ccgacttccgagcgcgcaactcgagagcgcgcTCGGAAAGTCGGGGGTCGGCGCGCAGTGCAGGCTGCGCACCCGGGAGGTAGCGGAG...3'

F. 5'-ccgacttccgagcgcgcaactcgagagcgcgcTCGGAAAGTCGGGGGTCGGCGCGCAGTGCAGGCTGCGCACCCGGGAGGTAGCGGAG...3'

**Figure 2: Sequences of chimeric junctions containing antisense and sense segments of laminin  $\alpha 1$  (top panel),  $\beta 1$  (middle panel) and  $\gamma 1$  (bottom panel) mRNAs and the projected pathways of their generation.** *Uppercase letters:* Nucleotide sequences of the sense RNA strands; *lowercase letters:* Nucleotide sequences of the antisense RNA strands. *Highlighted in green:* "AUG" translation initiation codons on the sense RNA strands; *highlighted in blue:* "uac" complements of translation initiation codons on the antisense RNA strands. *In italics and highlighted in grey:* Detected chimeric fragments. *Blue arrows:* Positions of antisense/sense RNA junction. The *top and bottom subpanels* of each panel depict amplification of mRNA molecules presumably transcribed from different TSSs; *note* that self-priming positions and, consequently, chimeric junction sequences shown in two subpanels of each panel are different. **A:** 5'-terminal regions of conventional mRNA. **B:** Antisense complements of the 5'-terminal regions of conventional mRNA. **C:** Folding of the antisense RNA strands into self-priming configurations. **D:** Extension of self-primed antisense RNA strands into sense-oriented RNA sequences. **E:** Projected chimeric junction sequences. **F:** Detected chimeric junction sequences. Note that self-priming events occur within segments of the antisense RNA strands corresponding to the 5'UTR of mRNAs, thus preserving the coding capacity of the amplified mRNAs.

of a segment of the antisense RNA to its sense RNA counterpart (rendering the former 3'-truncated and the latter "chimeric"), of a component not present in a conventional gene-transcribed mRNA progenitor of the amplification process. The future fate of the 3'-truncated antisense RNA end product can be remarkable and is discussed in the "Tier Two" sections below. The chimeric RNA end product, on the other hand, constitutes the "functional" outcome of the chimeric pathway. Its observed cellular levels reflect the potency of the amplification process: At the peak of the erythroid differentiation, for example, the amount of globin mRNA produced in the chimeric amplification pathway is about 1500-fold higher than the amount of conventionally generated globin mRNA in the same cells [8,9]. Potentially, the chimeric RNA end product can be translated into a polypeptide [8,9]. The possible translational outcomes of the chimeric pathway of mammalian mRNA amplification, however, are neither uniform nor necessarily identical to that attained with a conventional mRNA progenitor; they are rather multiple, conceivably drastically different, and can be grouped into the four Types as follows.

#### **Type I of the chimeric pathway of mammalian mRNA amplification preserves complete protein-coding information of conventional gene-transcribed mRNA progenitor**

Depending on the structure and nucleotide sequence of the initial mRNA amplification progenitor (or, more precisely, of its antisense RNA complement), translational outcomes of the chimeric pathway of mammalian mRNA amplification can vary drastically and cover the entire spectrum from a polypeptide identical to one encoded by a corresponding conventionally produced mRNA progenitor, to only a C-terminal fragment of the latter, to a polypeptide non-contiguously encoded in the genome, to a completely unrelated polypeptide and even to an abortive translation or translational incompetence. Which variant of the chimeric pathway will occur with a particular mRNA species is determined primarily by the site of the self-primed extension of the 3' terminus of the antisense RNA (Figure 3, Step 3). This, in turn, is defined by the position of the internal complementary element, ICE, within the antisense RNA molecule. This is because the 3' terminal complementary element, TCE is, by definition, always 3'-terminal, whereas the ICE can be potentially located anywhere within the antisense RNA. In the chimeric pathway Type I, depicted in Figure 3, the ICE is located within a segment of the antisense RNA corresponding to the 5'UTR of conventionally produced mRNA molecule. Therefore, the sense-oriented RNA, produced by the extension of the TCE, contains a portion of the 5'UTR and the entire coding region of the conventionally produced mRNA utilized as the initial amplification template (Figure 3, Step 4). The translational outcome in this case is a polypeptide identical to one translated from a conventional gene-transcribed mRNA progenitor. As seen in Figures 1 and 2, this Type of chimeric pathway is characteristic for amplification of mRNAs encoding alpha- and beta-globin chains and all three chains (alpha1, beta1 and gamma1) of laminin. Moreover, in the globin case,

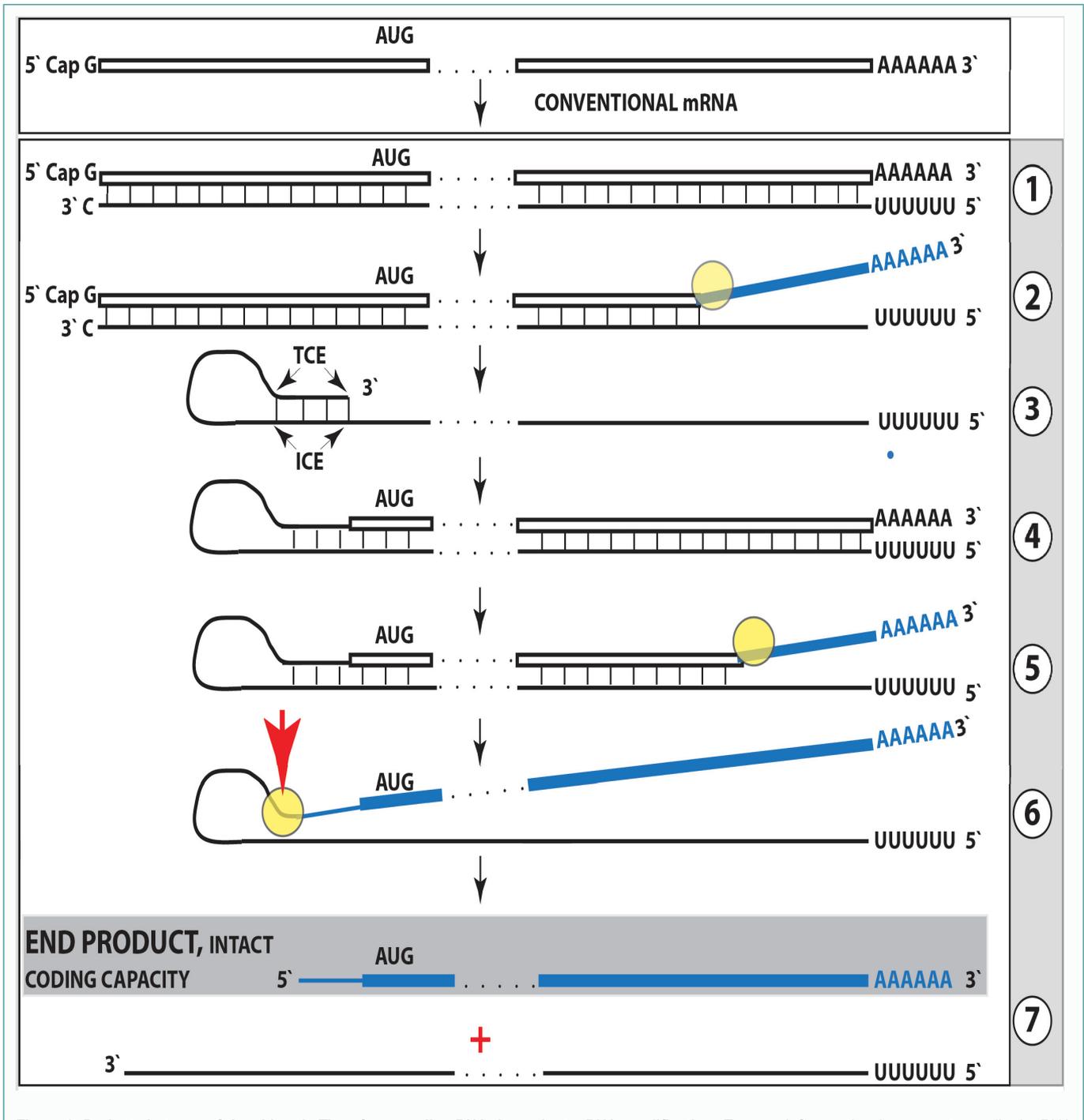
the translational product of the amplified globin mRNA was shown to be indeed indistinguishable from conventionally produced globin polypeptides [8,9]. This Type of the chimeric pathway is expected to take place in most physiologically occurring cases of mammalian RNA-dependent mRNA amplification.

#### **Type II of the chimeric pathway of mammalian mRNA amplification reduces or changes the protein-encoding information content of conventional mRNA progenitor**

In the chimeric mRNA amplification pathway, the position of the TCE element within the antisense RNA molecule is fixed: It is always strictly 3'-terminal. In contrast, the intramolecular location of the internal ICE element is variable, and potentially it can be positioned within a segment of the antisense RNA strand corresponding to the coding portion of an mRNA, a scenario diagrammed in steps 3' through 7' of the bottom panel of Figure 4. In this scenario, the chimeric RNA end product consists of a 3'-terminal segment of the antisense strand (the TCE or its fraction) attached, in a 5' to 3' orientation, to a 3' portion of a conventional mRNA amplification progenitor with a 5'-truncated coding region. In such a case, the translational outcome would be decided by the position of the first functional (capable of initiation of translation) AUG or another translation initiation-competent codon. If it were in-frame with the protein-encoding information content of conventional mRNA, translation would result in the C-terminal fragment, CTF, of a conventional mRNA-encoded polypeptide. This variant of the chimeric RNA-dependent mRNA amplification pathway would be asymmetric. Indeed, only one end, a 3'-terminal portion of the coding region continued into the 3'UTR, of conventional mRNA would be amplified, and its translation would potentially produce only one end of a corresponding conventional mRNA-encoded polypeptide, its C-terminal fragment, CTF. Such asymmetric amplification of Beta-Amyloid Precursor Protein ( $\beta$ APP) mRNA appears to be the source of beta-amyloid overproduction in Alzheimer's disease discussed in the "mRNA Amplification and Disease, Type II" section below. A CTF of a conventionally encoded protein is not the only possible outcome of this Type of chimeric pathway. If the first functional translation initiation codon is not in-frame with the conventional protein-coding information content but precedes in-frame an open reading frame, ORF, an unrelated polypeptide would result. If there were no ORF, an abortive translation would ensue and if a functional translation initiation codon were lacking, the result would be the translational incompetence.

#### **Type III of the chimeric pathway of mammalian mRNA amplification results in RNA containing protein-coding information non-contiguously encoded in the genome**

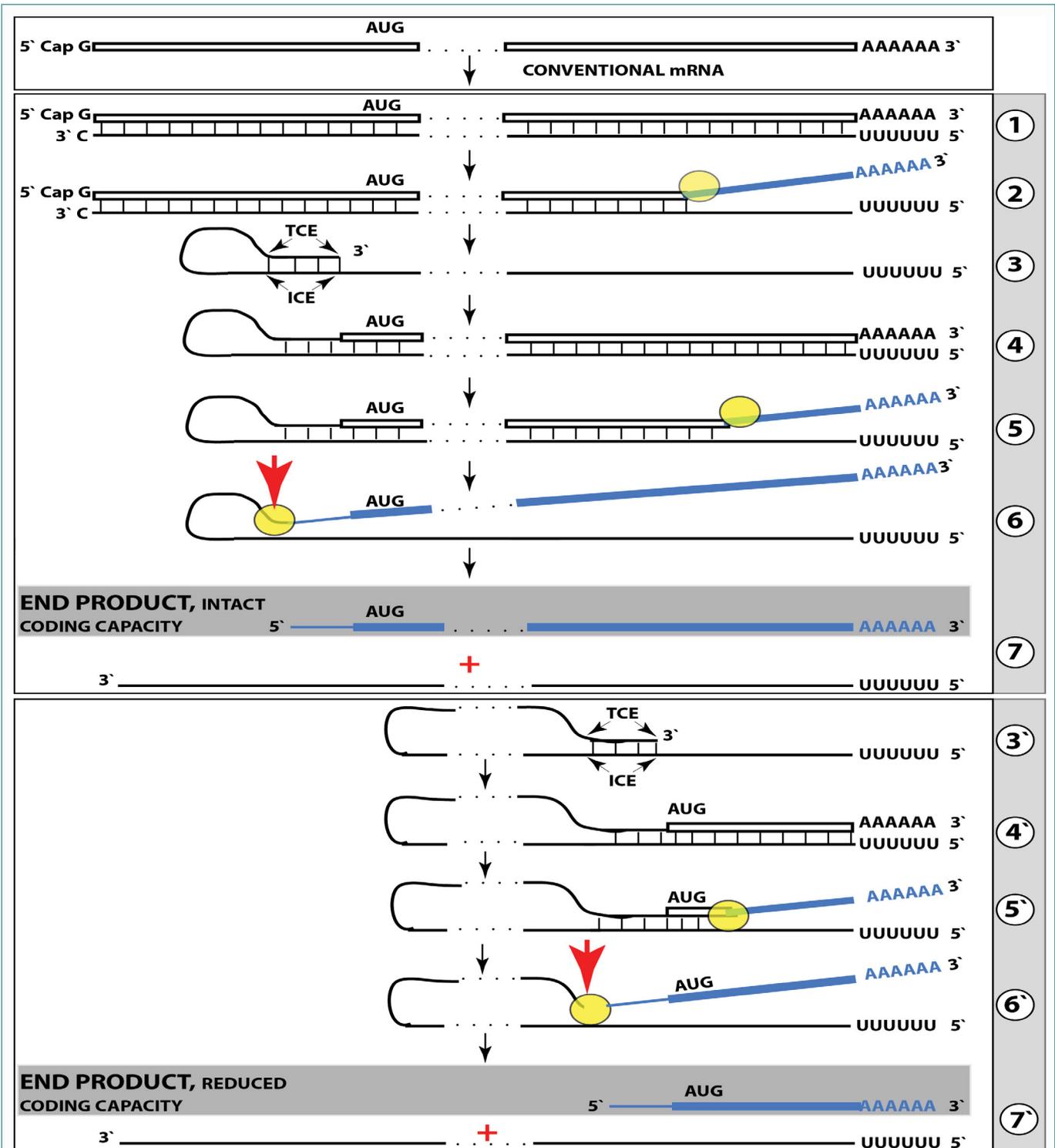
The preceding sections considered RNA-dependent mRNA amplification scenarios where protein-encoding information is contained only in the sense RNA strand. In these scenarios, the chimeric RNA end product either retains the entire coding information of a conventional mRNA and is translated from the



**Figure 3: Projected stages of the chimeric Tier of mammalian RNA-dependent mRNA amplification.** *Top panel:* Conventional, genome-transcribed mRNA molecule. *Bottom panel:* Projected stages of antisense RNA-mediated mRNA amplification. *Boxed line:* Sense RNA. *Single line:* Antisense RNA. "AUG": Functional translation initiation codon (could be other than "AUG"). "TCE": 3'-terminal complementary element; "ICE": Internal complementary element, both on the antisense RNA strand. *Yellow circle:* Helicase/modifying activity complex. *Blue lines* (both single and boxed): RNA strand, modified and separated from its complement by a helicase complex. *Red arrowhead:* Position of the cleavage of the chimeric intermediate. Step 1: Synthesis of the antisense strand. Note that, apparently, it is initiated within the poly(A) segment rather than at its end; step 2: Strand separation. Note that it can commence while the poly(U)-containing antisense RNA is still nascent; step 3: Folding of antisense RNA strand into self-priming configuration; step 4: Extension of self-primed antisense RNA; step 5: Strand separation; step 6: Cleavage of the chimeric intermediate; stage 7: End-products of amplification, the 3'-truncated antisense RNA and the chimeric RNA containing a 3'-terminal portion of mRNA and covalently joined cleaved-off 3'-terminal segment of the antisense RNA. Details in "Mechanistic Aspects" sections below.

original translation initiation codon or lacks the 5'-portion of the coding region of a conventionally produced mRNA. In case of the latter, its translation depends either on the availability of a functional in-frame translation initiation codon within the remaining portion of the coding region or, if it is out-of-frame, on the presence of a

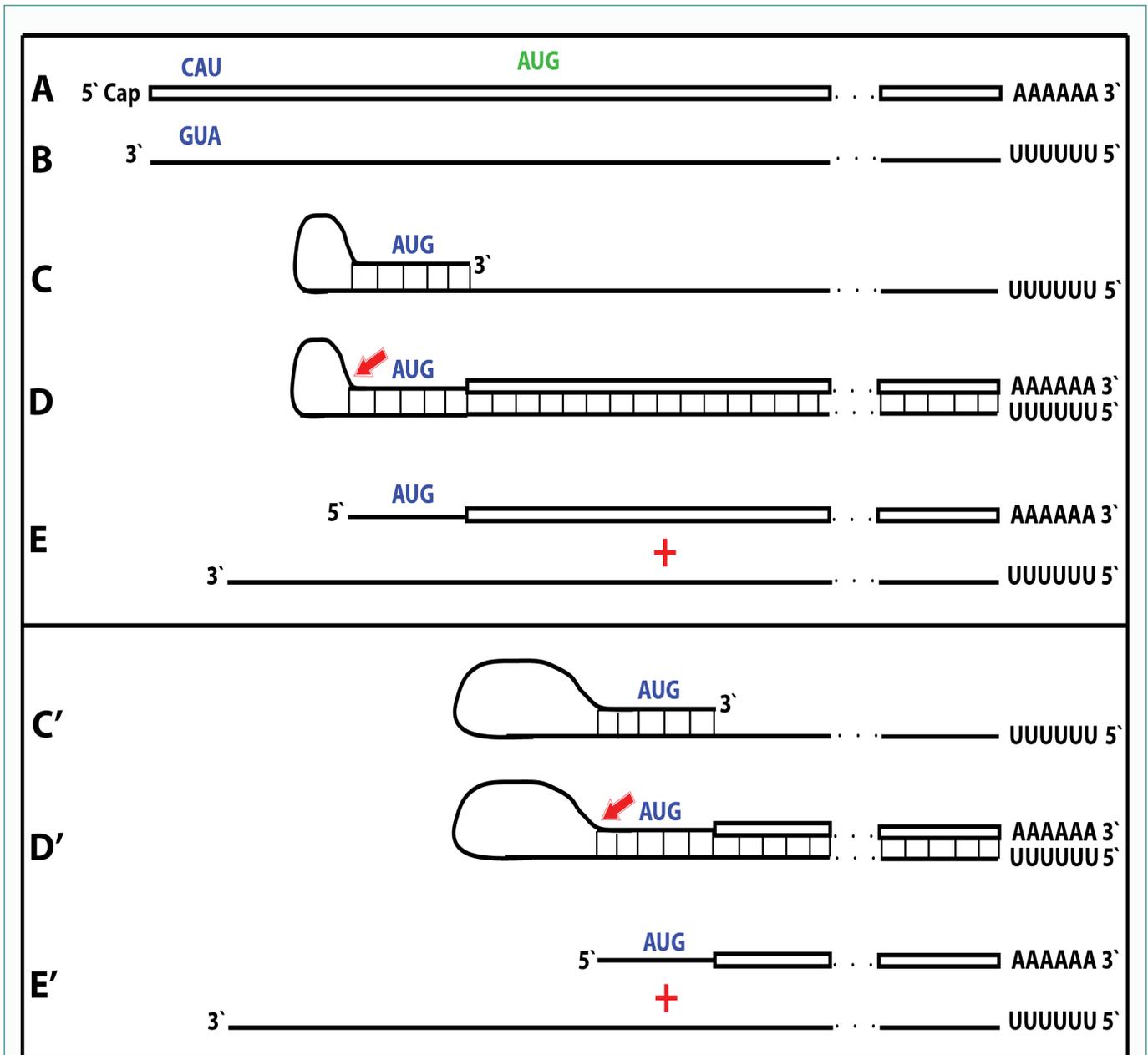
downstream ORF unrelated to the original information content. However, there is also another possibility. If the TCE of the antisense RNA contains a functional translation initiation codon in the 5' to 3' orientation, and if this translation initiation sequence was retained in the chimeric RNA end product after the cleavage of the chimeric



**Figure 4: Mammalian RNA-dependent mRNA amplification can result in a chimeric RNA end product encoding C-terminal fragment of a corresponding conventional polypeptide.** *Boxed line:* Sense strand RNA. *Single line:* Antisense strand RNA. "AUG": Functional translation initiation codon (could be other than AUG). "TCE": 3-terminal complementary element; "ICE": internal complementary element, both on the antisense RNA strand. *Yellow circle:* Helicase/modifying activity complex. *Blue lines* (both single and boxed): RNA strands, modified and separated from their complements by a helicase complex. *Red arrows:* Positions of the cleavage of the chimeric intermediate. Step 1: Synthesis of the antisense strand, note that, apparently, it is initiated within the poly(A) segment rather than at its end; step 2: Strand separation, note that it can commence while the poly(U)-containing antisense RNA is still nascent; step 3: Folding of antisense strand into self-priming configuration; step 4: Extension of self-primed antisense RNA; step 5: Strand separation; step 6: Cleavage of the chimeric intermediate; step 7: End-products of RNA amplification. Steps 3'-7' correspond to steps 3-7. *Top panel:* Conventional, genome-transcribed mRNA molecule. *Middle panel:* Projected stages of RNA-dependent mRNA amplification. "ICE" is located within a segment of antisense RNA corresponding to the 5'UTR of conventional mRNA; the chimeric RNA end product contains the entire coding content of conventional mRNA. *Bottom panel:* "ICE" is located within a segment of antisense RNA corresponding to the coding region of conventional mRNA. The amplified chimeric end product contains a 5'-truncated coding region of conventional mRNA. The translational outcome is decided by the position of the first functional translation initiation codon; if in-frame, with the protein-encoding information content of conventional mRNA, translation would result in the C-terminal fragment, CTF, of a conventionally encoded polypeptide; if not-in-frame, multiple potential outcomes, described in text, ensue.

intermediate, translation would be initiated from the initiation codon within the antisense portion of the chimeric RNA end product. In such a case, if the folding/self-priming of the antisense RNA occurs within its segment corresponding to the 5'UTR of conventional mRNA and if the initiation codon within the antisense portion of the chimeric RNA end product is in-frame with the mRNA coding sequence and there are no intervening in-frame stop codons, translation would result in a chimeric protein. It will contain, in its C-portion, the entire

polypeptide encoded by a conventional mRNA progenitor, enhanced at its N-end by additional amino acid residues encoded by the antisense portion of the chimeric RNA end product and by a segment of the 5'UTR of conventional mRNA. Thus, interestingly, the protein-encoding information content of the chimeric RNA end product of amplification would be non-contiguously encoded in the genome, as shown in the top panel of Figure 5.



**Figure 5: Mammalian RNA-dependent mRNA amplification may enhance protein-encoding information content of a conventional mRNA and generate polypeptides non-contiguously encoded in the genome.** *Boxed line:* Sense strand RNA. *Single line:* Antisense strand RNA. "AUG" (in green): Translation initiation codon of conventional mRNA progenitor. "AUG" (in blue): Antisense RNA-originated functional translation initiation codon (could be other than AUG). 5'-CAU-3' on the sense RNA; Complement of the 5'-AUG-3' (both in blue) on the antisense RNA. Red arrow: Position of the cleavage of the chimeric intermediate. **A:** Conventional mRNA; the "AUG" (in green) denotes translation initiation codon. **B:** Antisense complement of conventional mRNA; 3'-GUA-5' (in blue) is the 5'-AUG-3'. **C:** Folding of the antisense strand into self-priming configuration. **D:** Extension of self-primed antisense strand into sense-oriented sequence followed by strand separation and cleavage (red arrow) of the chimeric intermediate. **E:** The chimeric RNA and 3'-truncated antisense RNA end products of RNA-dependent mRNA amplification. *Note that translation of the chimeric end product starts from the antisense RNA-supplied initiation codon ("AUG", in blue) and produces a chimeric polypeptide non-contiguously encoded in the genome. Steps C', D', and E' correspond to C, D, and E. Top panel:* Antisense RNA folding/self priming occurs within its segment corresponding to the 5'UTR of conventional mRNA. *Bottom panel:* Antisense RNA folding/self priming occurs within its segment corresponding to the coding region of conventional mRNA. Translational outcomes are discussed in the text.

If, however, the folding/self priming of the antisense RNA strand occurs within its segment corresponding to the coding region of conventional mRNA, and if the initiation codon within the antisense portion of the chimeric RNA end product is in-frame with the mRNA coding sequence, the translational outcome would be a CTF of the conventional polypeptide, enhanced at its N-end by additional amino acid residues encoded by the antisense portion of the chimeric RNA end product, also a chimeric protein non-contiguously encoded in the genome. This variant of the chimeric pathway Type III is illustrated diagrammatically in the lower panel of Figure 5. In both variants shown in Figure 5, if the antisense RNA-derived translation initiation codon were not in-frame with the conventional protein-encoding content, the relevant alternatives described in the preceding section would apply.

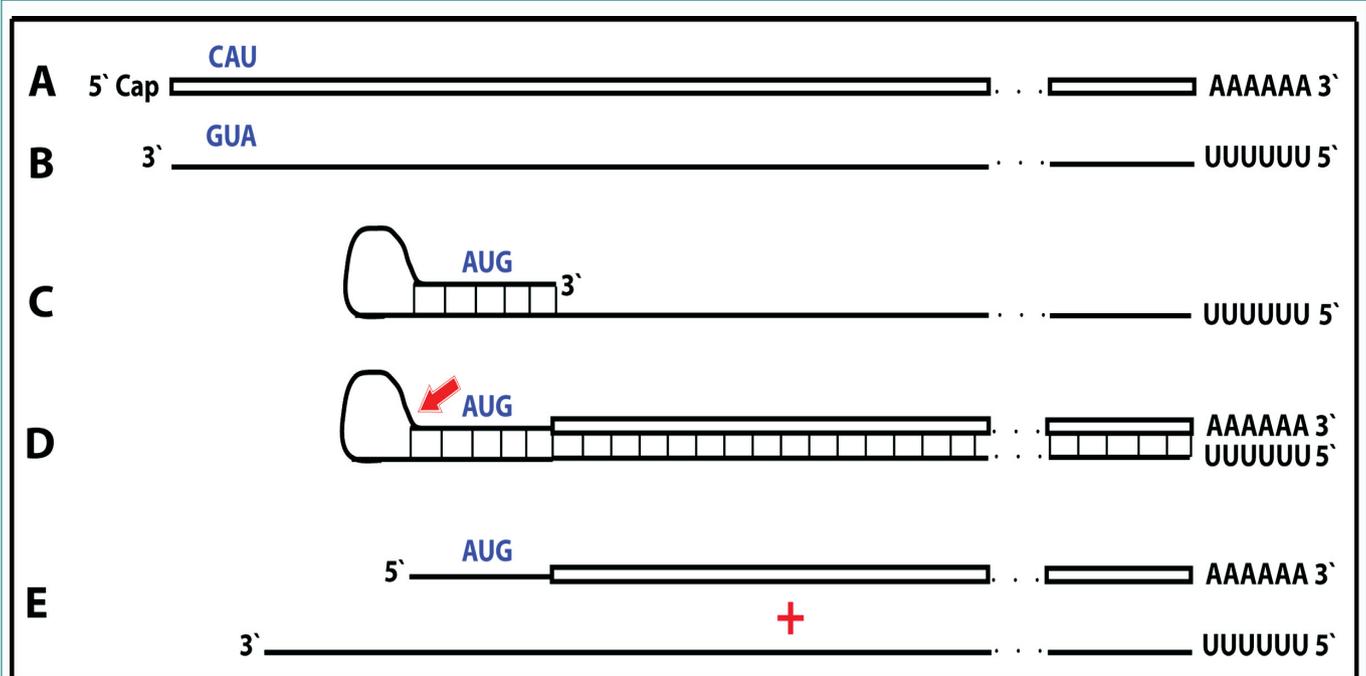
#### Type IV of the chimeric pathway of mammalian mRNA amplification activates dormant protein-encoding information

A scenario considered in this section is one where a conventional gene-transcribed mRNA contains untranslatable protein-encoding information that is expressed as a polypeptide only under specific conditions, for example a particular stress, that activate the RNA-dependent mRNA amplification process, which results, due to the manner of generation and processing of the chimeric intermediate of mRNA amplification, in a translatable chimeric RNA end product. It should be emphasized that regardless of the extent of its physiological occurrence, the understanding of such a scenario is important because it could, and actually does, suggest very useful experimental and bioengineering designs, as described in "Reporter Constructs for the Chimeric Pathway" section below. In this scenario,

a conventionally produced RNA encodes a polypeptide but lacks a functional translation initiation codon upstream from and in-frame with the encoded information. Since its potential to produce a protein is present but unrealized under regular circumstances, it can be considered a silent, or "dormant" mRNA. The potential to direct synthesis of a polypeptide can be realized, and a dormant mRNA activated, by the mRNA amplification process. In the chimeric pathway of mRNA amplification, the chimeric RNA end product contains a 5'-terminal segment contributed by the antisense RNA and not present in the conventional genome-transcribed RNA molecule. If this additional segment contains a functional translation initiation codon, and if this codon is in-frame with the encoded information, the protein-encoding content of a conventional "dormant" mRNA would be activated in the chimeric RNA end product of amplification. The resulting translational outcome would be a chimeric protein consisting of a conventionally encoded polypeptide with the N-end encoded by the antisense RNA. This chimeric protein would be, in similarity to the preceding scenario, also non-contiguously encoded in the genome. The above scenario is diagrammed in Figure 6.

#### Mechanistic and regulatory aspects of the chimeric pathway. Activation of mammalian RNA-dependent mRNA amplification: Operational definition

The present and several following sections discuss our current understanding of mechanistic and regulatory aspects of the chimeric pathway of mammalian RNA-dependent mRNA amplification. The first stage of this process is its activation. What does the "activation" mean in the context of mammalian mRNA amplification? To formulate its operational definition, it is instructive to review the example of hepatitis delta virus, HDV, an RNA virus that encodes



**Figure 6: Mammalian RNA-dependent mRNA amplification may activate a dormant protein-encoding information and generate polypeptides non-contiguously encoded in the genome.** Boxed line: Sense strand RNA. Single line: Antisense strand RNA. "AUG" (in blue): Functional translation initiation codon (could be other than "AUG"). 5'-CAU-3' (in blue) on the sense RNA: Complement of 5'-AUG-3' on the antisense RNA. Red arrow: Position of the cleavage of the chimeric intermediate. **A:** Genome-transcribed "dormant" mRNA containing protein-encoding information but lacking in-frame functional translation initiation codon. **B:** Antisense complement of conventional mRNA. Note that 3'-GUA-5' (in blue) is the 5'-AUG-3'. **C:** Folding of the antisense RNA strand into self-priming configuration. **D:** Extension of self-primed antisense strand into sense-oriented sequence followed by strand separation and cleavage (red arrow) of the chimeric intermediate. **E:** End products of RNA-dependent mRNA amplification. Note that translation of the chimeric RNA end product starts from the antisense RNA-contributed translation initiation codon ("AUG", in blue) and produces a chimeric polypeptide non-contiguously encoded in the genome.

a single protein, hepatitis delta antigen (HDAG). HDAG is neither RdRp nor any other polymerase, yet the virus is capable of vigorous replication in host mammalian cells [11,12]. This ability attests to the presence and functionality of RNA-dependent RNA polymerase activity in mammalian cells. Indeed, the RdRp activity was documented in a large variety of mammalian cells and appears to be non-conventional; two possible candidates suggested for this role are the RNA polymerase II complex or its components [13,14] and RdRp activity of the TERT complex [15], both ubiquitously present in all mammalian cells. Under regular circumstances, the RdRp activity in mammalian cells produces only short antisense RNA transcripts. For example, a widespread synthesis of diverse short antisense RNA transcripts initiating at the 3'poly(A) of mRNAs was observed in human cells [16]. It follows that in order to replicate employing cellular RdRp activity, HDV contributes a factor enabling the production of long transcripts. This factor can only be the HDAG. HDAG was indeed shown to be essential both for the production of long transcripts by cellular RdRp, and for the viral replication [17-19]. In its absence only short transcripts are generated [17-19]. These observations imply that HDAG constitutes a processivity co-factor of cellular RdRp activity. They also establish a general principle of the requirement for the presence of "inducible" components of the mammalian RdRp complex for its full functionality. Therefore, the activation of mammalian RNA-dependent mRNA amplification can be operationally defined as the assembly of the "competent" RdRp complex capable of transcribing complete complements of the legitimate RNA templates. This, in turn, necessitates either the activation or *de novo* production of "inducible" components that are absent or inactive in the cell under regular circumstances. It can be reasoned that, physiologically, the emergence of inducible components of the mammalian RdRp complex is triggered by a variety of cellular stresses and occurs in the context of stress response in general and within the framework of the integrated stress response, ISR, in particular.

### **Mechanistic and regulatory aspects of the chimeric pathway: Role of the integrated stress response in activation of the mammalian mRNA amplification process**

The "integrated stress response", ISR, is a complex signaling pathway operating in eukaryotic cells, which is activated in response to a wide range of cellular stresses [20-29]. The "integrating" event in this pathway is the convergence of all stimuli that activate the ISR to the one common point: Phosphorylation of the alpha subunit of eukaryotic translation initiation factor 2, eIF2 $\alpha$ , at serine 51. In mammalian cells, this phosphorylation is catalyzed by members of the family of eIF2 $\alpha$  kinases. This family is comprised of four members: PKR-like ER kinase, PERK; double-stranded RNA-dependent kinase, PKR; general control none-derepressible-2 kinase, GCN2; and heme-regulated kinase, HRI. All four kinases have extensive homology in their catalytic domains but exhibit distinctly different regulatory domains and each requires autophosphorylation and dimerization for full activation [30-35]. As a reflection of their unique regulatory mechanisms, each of these kinases responds to distinct stress stimuli [36]. For example, PERK is located at the endoplasmic reticulum, ER, membrane and is usually activated by ER stress [37-41]. It was also shown to be activated by glucose deprivation, ATP depletion, and oncogene activation [42-45]. GCN2 binds to deacylated tRNAs to become active in response to amino acid deprivation [46] and was also shown to be activated by prolonged glucose deprivation and by UV light [47-49]. PKR is activated mainly by double-stranded RNA, usually during viral infection [50,51]. Other types of stress, such

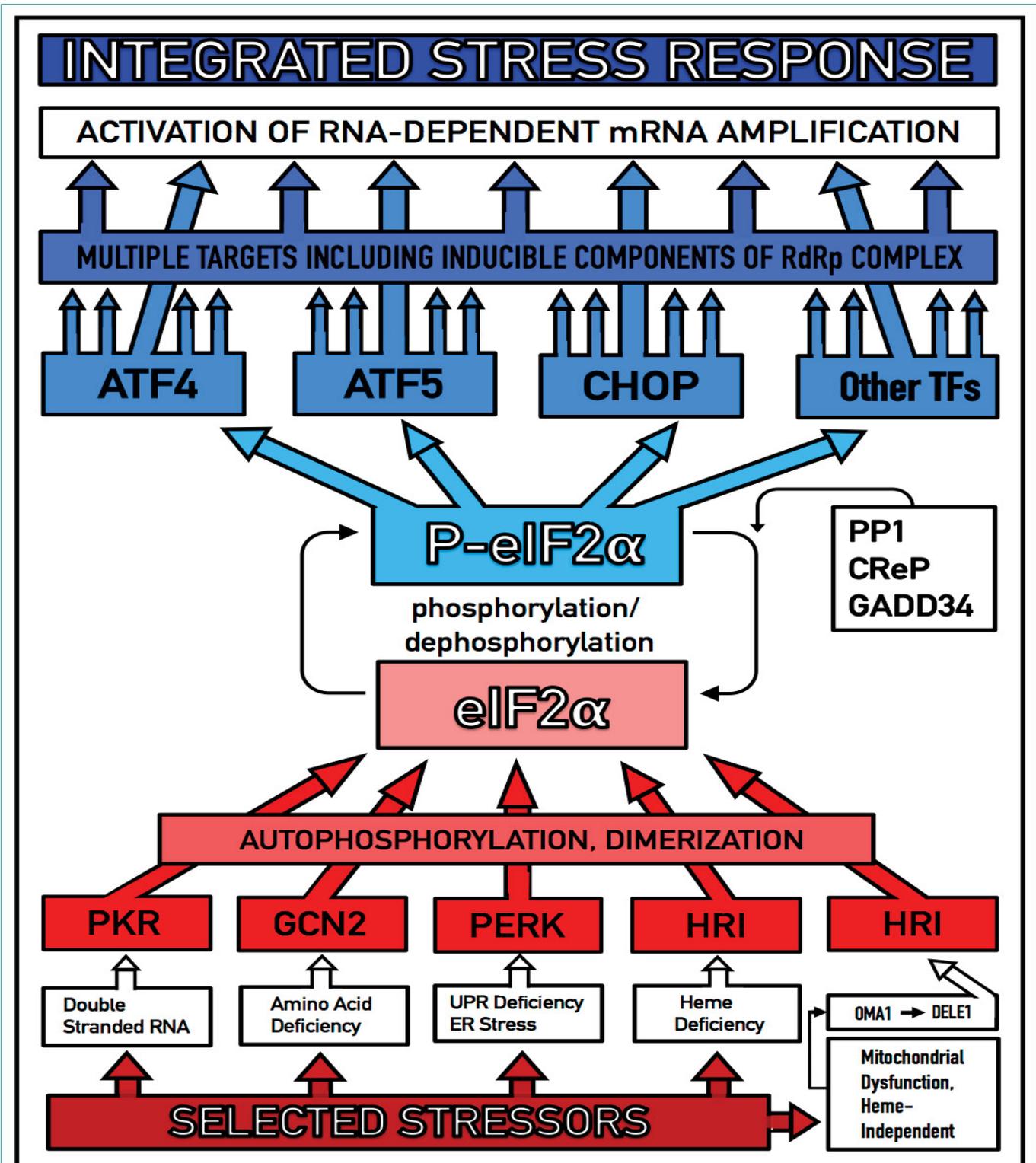
as oxidative and ER stress, growth factor deprivation, cytokines, ribotoxic stress, stress granules, and caspase activity have also been shown to activate PKR in dsRNA-independent manner [52-59]. HRI, unlike other family members that have a broad tissue distribution, is mainly expressed in erythroid cells where it is activated by heme deficiency and is involved in erythroid differentiation [60,61]. It can also be activated, independently of heme, by a variety of other types of stress, including arsenite-induced oxidative stress, heat shock, osmotic stress, proteasome inhibition and nitric oxide [62-65]. Recently, HRI activation and, consequently, the occurrence of the ISR were shown to be triggered, via OMA1 to DELE1 to HRI signaling pathway, by mitochondrial distress [66,67].

Phosphorylation of eIF2 $\alpha$  at serine 51 causes a severe reduction in global cellular protein synthesis, primarily through inhibition of the cap-dependent initiation, while at the same time facilitating cap-independent translation of selected mRNAs, including those encoding specific transcription factors. By the virtue of their ability to interact with nucleic acids, transcription factors are prime candidates for the role of inducible components of the RdRp complex. Alternatively the inducible RdRp co-factors are among products of genes activated by transcription factors produced under the ISR conditions. In turn, the availability of inducible components of the RdRp complex allows the activation of the RNA-dependent mRNA amplification process. This sequence of events is depicted diagrammatically in Figure 7.

The logic of the process illustrated in Figure 7 can be used to explain two documented and one postulated cases of RNA-dependent mRNA amplification in mammalian cells. The first one is amplification of mRNAs encoding globin chains in cells undergoing erythroid differentiation [8,9]. In this case, it can be envisioned that differentiation-related increase in conventional production of globin polypeptide chains results in sequestration of heme in the hemoglobin tetramere structure thus forming the hemoglobin. This leads to heme depletion, which triggers the activation of the HRI kinase. Activated HRI kinase phosphorylates eIF2 $\alpha$  at serine 51, thus suppressing total protein production and activating expression of select transcription factors. Either transcription factor(s) themselves or products of some genes activated by them serve as inducible components of the RdRp complex. RNA-dependent amplification of globin mRNAs is activated, yet more globin chains are produced via the mRNA amplification pathway (see relevant discussion at the conclusion of the present section), more heme is sequestered in hemoglobin structures, heme deficiency stress is sustained, the ISR is maintained, and the cycle is perpetuated.

The second documented example is that of amplification of mRNA encoding alpha1, beta1 and gamma1 chains of laminin [10] in cells producing and secreting extraordinary quantities of ECM components (in fact, matrigel-producing cells). In this case, physiologically increased conventional expression of genes encoding secreted ECM components "clogs" endoplasmic reticulum and triggers ER stress, which activates PERK. Activated PERK phosphorylates eIF2 $\alpha$ , which, in turn, suppresses the bulk of cellular protein synthesis and activates the expression of select transcription factors. Consequently, inducible co-factors of mammalian RdRp are produced, RNA-dependent amplification of laminin-encoding mRNAs is activated, even more laminin polypeptides (and other ECM components) are generated in the mRNA amplification pathway, ER stress is sustained, the ISR is supported, and the whole cycle is self-propagated.

The third example is that of postulated asymmetric amplification



**Figure 7: Integrated stress response-mediated activation of mammalian RNA-dependent mRNA amplification.** Signals from various cellular stresses, some listed in the Figure, activate, in a stress type-specific manner, one or more members of the family of eIF2 $\alpha$  kinases: PKR, GCN2, PERK, and HRI. In a number of cases, signaling pathways leading to kinase activation have been elucidated and are described in text. The detailed pathway for heme-independent HRI activation shown in the Figure depicts mitochondrial stress-induced depolarization activating mitochondrial protease OMA1, which cleaves DELE1 inside of the organelle. One of DELE1 fragments is then released to the cytosol where it binds to and activates HRI. Full activation requires autophosphorylation and dimerization of the kinases. Activated kinases converge on phosphorylation of eIF2 $\alpha$  at serine 51; this is the “integrating” event and the core of the integrated stress response pathway. eIF2 $\alpha$  phosphorylation results in severe global attenuation of cap-dependent translation and also in concomitant initiation of the preferential translation of ISR-specific mRNAs, many encoding transcription factors. Some of these transcription factors, and/or products of genes activated by them, act as inducible components of the RdRp complex; their availability allows assembly of the complex and thus the activation of the mammalian RNA-dependent mRNA amplification process. Termination of the ISR is regulated by the constitutively expressed CReP and stress-inducible phosphatase GADD34 that, in conjunction with PP1 phosphatase, dephosphorylate eIF2 $\alpha$ .

of  $\beta$ APP mRNA in Alzheimer's disease [68-75], described in detail in the "mRNA Amplification and Disease, Type II" section below. In this case, a lifelong intraneuronal accumulation of conventionally produced beta-amyloid reaches critical levels and triggers mitochondrial dysfunction. This, in turn, activates mitochondrial protease OMA1, which cleaves, still within the mitochondria, protein DELE1. A newly generated fragment of DELE1 is released to the cytosol where it activates the HRI kinase [66,67]. Activated HRI phosphorylates eIF2 $\alpha$ , the bulk of cellular protein synthesis is suppressed, select transcription factors are expressed, and the mRNA amplification pathway is activated in neuronal cells. Within this pathway, human beta-amyloid precursor protein-encoding mRNA undergoes asymmetric amplification of Type II described above. It results in a chimeric RNA molecule encoding a C-terminal fragment of  $\beta$ APP containing beta-amyloid amino acid sequence at its N-end, the immediate precursor of beta-amyloid peptide, which is generated by gamma cleavage of the CTF in a non-secretory pathway. As a result, more intraneuronal beta-amyloid is produced via the mRNA amplification pathway, mitochondrial distress persists, the ISR endures and the cycle is perpetuated.

Two important notions should be emphasized in the context of the above discussion. First, the ISR could be not the sole activator of mammalian RNA-dependent mRNA amplification. Some types of stress, for example ER stress, activate certain transcription factors by mechanisms distinct from the ISR [76,77]. These transcription factors or products of genes activated by them could potentially serve as inducible co-factors of mammalian RdRp. The second important notion is that, by the intrinsic logic of the process of ISR-activated mRNA amplification, the cap-less amplified mRNA has to be compatible with phosphorylated eIF2 $\alpha$  and freely translatable under the ISR conditions. Such compatibility could be the consequence of extensive nucleotide modifications of amplified mRNAs and/or their shortened 5'UTRs (on both points see more below). In such a case, amplified mRNA, in addition to being extraordinarily abundant due to the method of its generation, it is also preferentially translated under the ISR conditions, thus augmenting the efficiency of the amplification process.

### **Mechanistic and regulatory aspects of the chimeric pathway: Generation of chimeric intermediates**

Following the activation of the mammalian RNA-dependent mRNA amplification pathway, presumably via the integrated stress response-mediated expression of inducible components of the RdRp complex, the amplification process commences with generation of the complete antisense transcript of a conventionally produced, mature, i.e. spliced, cytoplasmic, amplification-eligible mRNA molecule, as shown in Figure 3, Step 1. The only major prerequisite for an mRNA to serve as the initial amplification template appears to be the presence of the poly(A) segment at its 3' terminus [8,9]. The vast majority of mammalian mRNA species contains 3'-terminal poly(A) segments. The notion that many, or possibly most, of them could be eligible templates for RdRp was suggested in earlier studies [6]. Subsequent observations by Kapranov and co-investigators showed a widespread synthesis of antisense RNA initiating, apparently indiscriminately, at the 3'-terminal poly(A) of diverse mRNA species in human cells [16]. This, evidently undiscerning, RdRp template eligibility of the bulk of mammalian mRNA species raises questions with regard to mechanisms underlying the manifestly stringent specificity of the mRNA amplification process as seen, for example, in erythropoietic

differentiation [6,8,9]. The specificity of the amplification process appears to be determined, as discussed below, at the 3' terminus of an antisense RNA transcript by its ability or inability to support production of a complementary sense strand RNA molecule, i.e. amplified mRNA [6,8,9].

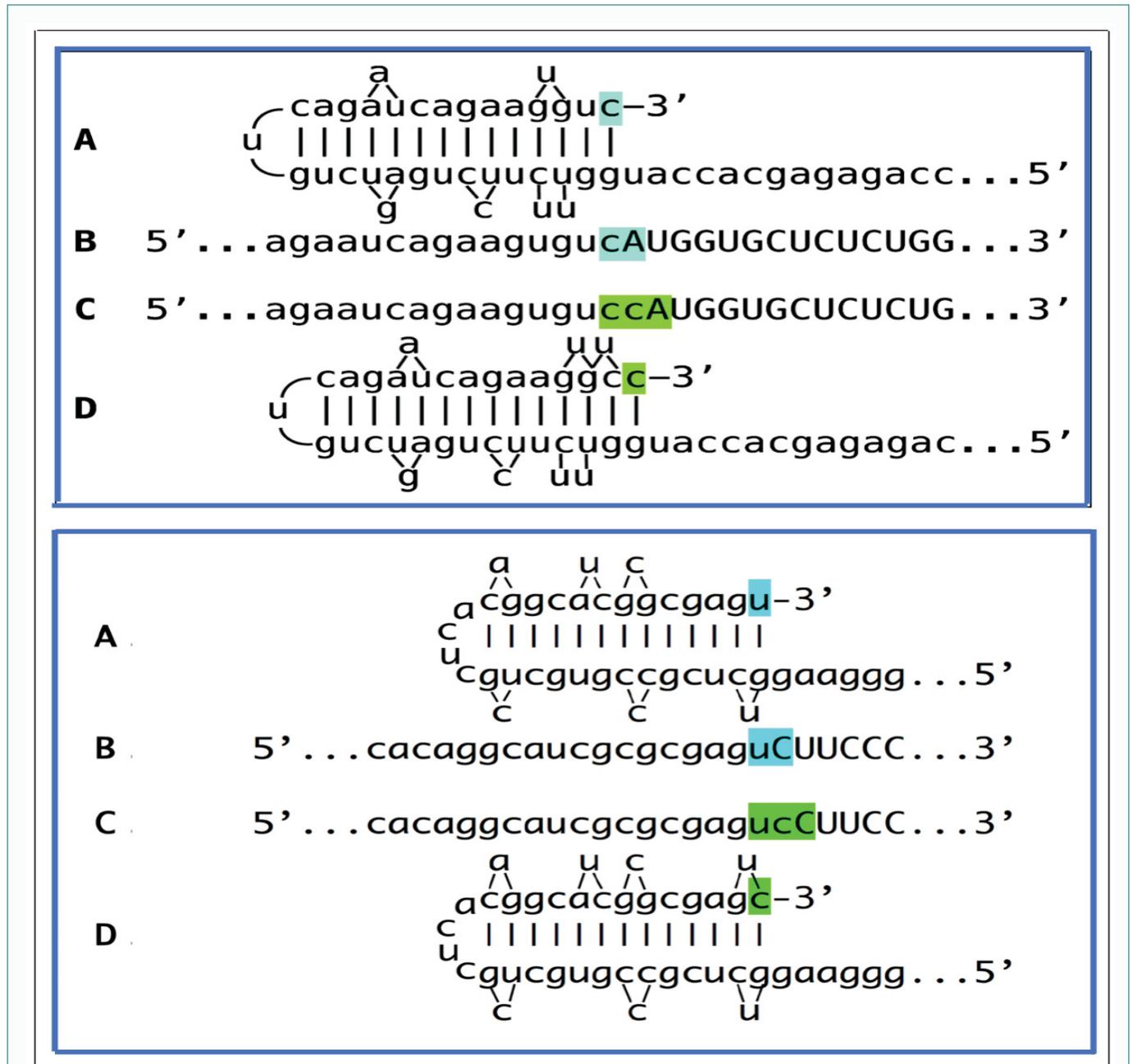
The observed uniformly short 3'poly(A) region of the chimeric RNA end product of the amplification process [9], apparently transcribed in a run-off manner from the 5'poly(U) segment of the antisense RNA strand, indicates that synthesis of the antisense RNA starts not at the 3' end of mRNA molecules, characterized by a wide size-range of their 3'-terminal poly(A) segments [9], but within the 3' poly (A) segment of the sense RNA strand, in a relatively narrow distance range from the encoded 3'UTR of a conventional mRNA. The manner of antisense RNA initiation remains to be determined; it may involve priming with a uridylylated protein, as seen in viral systems [4], or occur by a different mechanism. Once transcription is initiated, the RdRp complex proceeds along an mRNA template until a complete antisense RNA complement is produced. Interestingly, at the 3'terminus of the antisense strand, RdRp appears to be capable of transcribing the cap "G" of conventional mRNA despite its inverted orientation [9,10]. The rationale for this conclusion is based on the detected Type One antisense/sense junction sequences of chimeric alpha-globin RNA amplification intermediates [8,9], uniformly containing at the 3' termini of their complete antisense portions the "C" not encoded in the genome, as seen in Figure 1; it is summarized in the top panel of Figure 8. If the cap "G" is not transcribed, the antisense strand would terminate with the 3'-terminal "c" corresponding to the transcription start site of mRNA (highlighted in blue in line A). In such a case, antisense RNA folding/self-priming configuration would be as shown in line A, and, following the extension of self-primed antisense RNA, the antisense/sense junction would consist of the "c/A" (highlighted in blue) as depicted in line B. The experimental results (presented in Figure 1) are different. They show that the sequence of the antisense/sense junction is, in fact, the "cc/A" (line C; highlighted in green). Since the genomic sequence upstream of the transcription start site cannot account for the additional 3'-terminal "c" in the antisense strand, the remaining possibility is that the "c" in question (line D; highlighted in green) is a transcript of the cap "G" of the sense strand and that the antisense folding into a self-priming configuration occurs as shown in line D.

Similar conclusion was reached in the case of the observed amplification of laminin mRNA [10]. It is based on the detected antisense/sense junction sequences of chimeric  $\alpha$ 1 laminin RNA amplification intermediates shown in the upper subpanel of the top panel of Figure 2; this conclusion is illustrated in the bottom panel of Figure 8. If the cap "G" is not transcribed, the antisense RNA strand would terminate with the 3'-terminal "u" (highlighted in blue in line A) and its folding configuration would be as shown in line A. After the extension, the sense/antisense junction would consist of the "uC" highlighted in blue and depicted in line B. The experimental results, however, show that the sequence of the sense/antisense RNA junction is, in fact, the "ucC" highlighted in green in line C. Since the genomic sequence upstream of the TSS cannot account for the additional 3'-terminal "C" in the antisense strand [78], the only remaining possibility is that the "C" in question (highlighted in green in line D) is a transcript of the cap "G" of the sense RNA strand and that the antisense RNA folding into a self-priming configuration occurs as shown in line D. It should be emphasized that the above two examples of the capG transcription by RdRp are presented because they are unambiguous.

In other cases, like that of beta-globin mRNA amplification, either a “G” occurs in the genome immediately upstream from the TSS, or the antisense component of the sense/antisense RNA junction sequence could be truncated, and the capG-transcribed “C” removed, in the process of the “chimeric junction shift”, a phenomenon discussed

below; in these cases, transcription of the capG during generation of the antisense RNA from an mRNA template could be obscured. On the other hand, the two unequivocal examples presented above are convincing enough to warrant generalization.

Following the completion of the antisense RNA complement of a



**Figure 8: Mammalian RNA-dependent RNA polymerase can transcribe the cap “G” of mRNA.** *Uppercase letters:* Nucleotide sequence of the sense strand; *lowercase letters:* Nucleotide sequence of the antisense strand. **A:** Projected self-priming configuration of the antisense RNA strand in the absence of the cap “G” transcription. **B:** Projected nucleotide sequence of the antisense/sense RNA junction in the absence of the cap “G” transcription; “c” highlighted in blue: 3’-terminal nucleotide (if the capG is not transcribed) of the antisense strand, corresponding to the transcription start site of mRNA; “cA” highlighted in blue: The projected antisense/sense RNA junction structure in the absence of the cap “G” transcription; “c” highlighted in green: Transcript of the cap “G”; “cA” highlighted in green: The resulting antisense/sense RNA junction structure when the cap “G” is transcribed. **Bottom panel:** Transcription of the capG of murine alpha1-laminin mRNA; data shown is adopted from the top subpanel of the top panel of Figure 2. “u” highlighted in blue: 3’-terminal nucleotide (if the capG is not transcribed) of the antisense strand, corresponding to the transcription start site of mRNA; “uC” highlighted in blue: The projected antisense/sense RNA junction structure in the absence of the cap “G” transcription; “c” highlighted in green: Transcript of the cap “G”; “ucC” highlighted in green: The resulting antisense/sense RNA junction structure when the cap “G” is transcribed. *Note that in both cases the genomic sequence upstream of the TSS cannot account for the additional 3’-terminal “C” in the antisense strand.*

conventional mRNA, strands of the double-stranded RNA structure are separated (Figure 3, Step 2). This is effected by a helicase activity, which mounts the 3'-terminal poly(A)-containing RNA strand of the double-stranded structure at its 3' terminus and proceeds along this strand toward its 5' terminus [8,9]; strand separation, apparently, can commence while the poly(U)-containing antisense RNA strand is still nascent. As the helicase progresses, it, or an associated activity, modifies on average every fifth nucleotide of the poly(A)-containing RNA strand [8,9]. Only two species of nucleotides are modified during strand separation; it appears that both are purines, the A and the G [8,9]. Moreover, it appears that the same modifying group with mass number of 37 is appended to both the A and the G [8,9]. The conclusion that both modified bases are purines implies that the modification likely occurs at a position(s) not available in pyrimidine nucleotides. One likely candidate is purine position 3 in the proximity to the Watson-Crick interphase, a position used in the pyrimidines for sugar attachment. The modified chimeric RNA end product appears to be "sticky" in that it forms extensive secondary and tertiary structures [9]. Therefore, the chimeric RNA end product of amplification, containing modified nucleotides, cannot be hybridized to long probes under stringent condition [8,9]. Moreover, it appears that the modified nucleotides are either impassable or facilitate the creation of RNA structures that are impassable for a reverse transcriptase. Experimentally, denaturation of the modified chimeric RNA end product of amplification is challenging. The only known method to accomplish this is treatment with methyl mercury hydroxide [8,9]. Considering that methyl mercury hydroxide is known to interact with the nucleotide's NH group during denaturation, and allowing for the mass number requirement of 37, one plausible candidate for the modifying group in question is  $\text{LiCH}_2\text{NH}_2$  [8,9].

At the conclusion of the strand separation stage of the amplification process, a separated conventional mRNA progenitor molecule emerges heavily modified but otherwise remains intact and can be re-utilized repeatedly as an RdRp template [8,9], whereas its complete antisense RNA complement contains no modified nucleotides [8]. As described below, this outcome, the lack of nucleotide modifications in the antisense RNA transcript, is crucial for the progression of the amplification process.

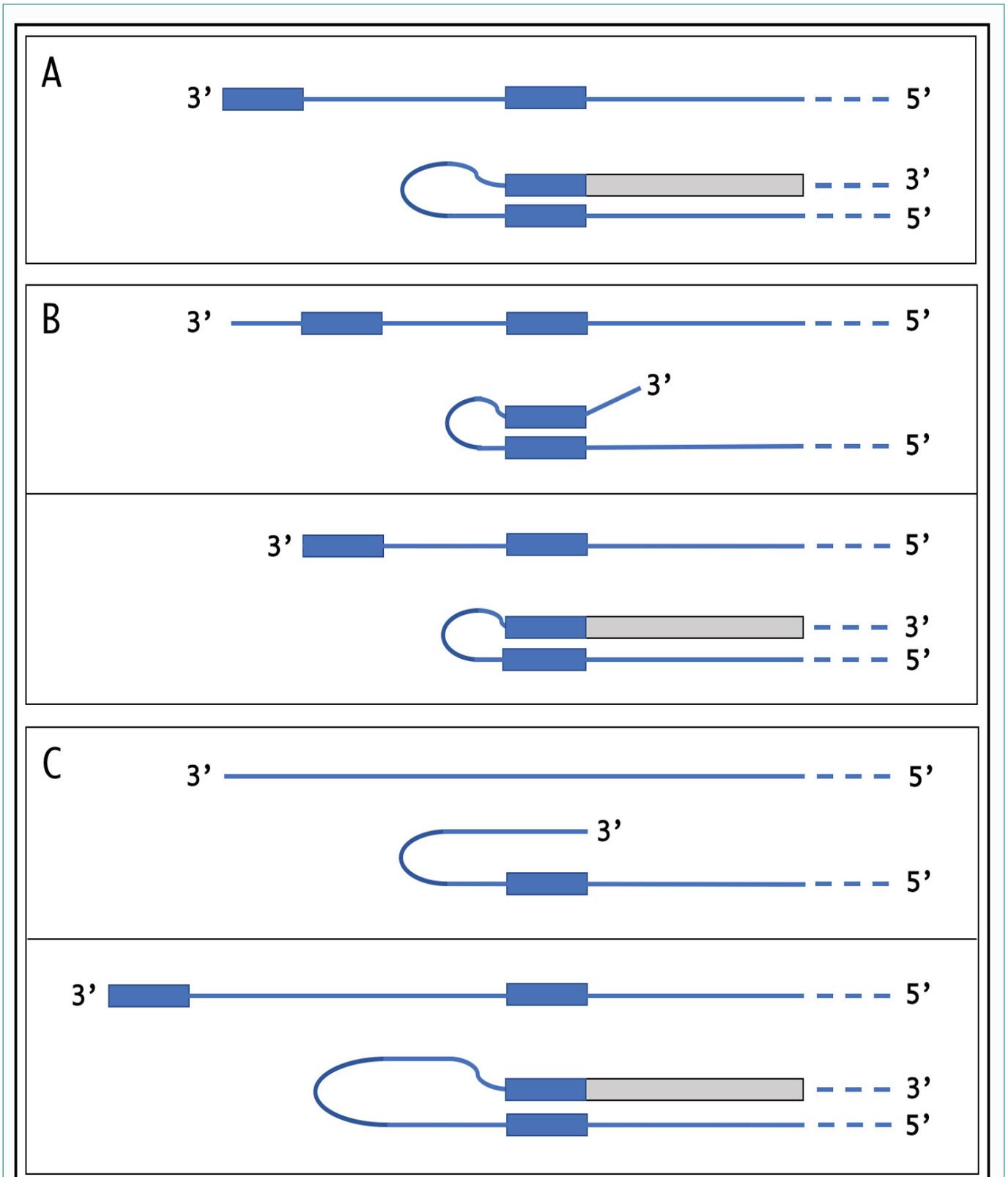
The generation of the sense RNA strand on the antisense RNA template occurs via the extension of the 3' terminus of a self-primed antisense template and requires the presence within the antisense RNA transcript of two spatially independent complementary elements whose occurrence in antisense globin RNA molecules was found to be evolutionary conserved across mammalian species [5]. One of these is the strictly 3'-Terminal Complementary Element (TCE), the other is the Internal Complementary Element (ICE). These elements (Figure 3, Step 3) must be complementary to an extent sufficient to form a stable priming structure but may contain both mismatches and unconventional G/U pairings. Even if the TCE and the ICE nucleotide sequences are present in the antisense RNA, their proper interaction resulting in a stable self-priming structure is not assured. This is because extensive secondary and tertiary structures, characteristic for any RNA molecule, can obscure and/or interfere with the TCE and the ICE elements and prevent their association. It appears, therefore, that the generation of the sense RNA on the antisense RNA template requires the thermodynamic feasibility, enhanced/enabled by the occurrence of these two complementary and topologically compatible elements, of the antisense RNA strand folding into a stable self-priming configuration. The requirement for

terminal localization of the TCE appears to be stringent; an overhang of even a single nucleotide significantly diminishes self-priming efficiency [5]. The crucial importance of the antisense RNA strand remaining unmodified following the separation from its conventional mRNA template (Figure 3, Step 2) plays out at this stage by enabling complementary interactions and the formation of a self-priming structure that otherwise would be impeded. Provided that a self-priming structure is formed, the 3' end (the TCE/primer element) of the folded antisense strand is extended by the RdRp complex into a sense-orientated RNA molecule terminating with the poly(A) at the 3' end (Figure 3, Step 4), thus generating a hairpin-structured chimeric intermediate consisting of covalently joined sense and antisense RNA strands.

The observed examples of mammalian RNA-dependent mRNA amplification [9,10] described above involve mRNA species transcribed from two types of genes. One type, exemplified by globin chains-encoding genes, contains the "TATA" regulatory element and a rigidly fixed Transcription Start Site (TSS). The other type, exemplified by laminin chains-encoding genes, is TATA-less. With the latter, transcription can be initiated from multiple positions and the eligibility for the RdRp-mediated amplification process appears to be regulated, at least in part, by a positional shift of transcription start site. For example, most of the known multiple TSS positions [79,80] in genes encoding all three studied chains of laminin [10] are inconsistent with the eventual generation of antisense RNA molecules capable of self-priming within their segments corresponding to the 5'UTRs of mRNA because positional requirements for the TCE and ICE elements within these molecules are not satisfied. The same is also true for another TATA-less gene, which encodes human beta-amyloid precursor protein and is discussed below. However, a shift in the TSS position may result in compliance with the requirements for the TCE and the ICE and thus can confer on an mRNA the eligibility for the mRNA amplification process. The concepts of such regulation are summarized in Figure 9. If the 3'-distant complementary element of the antisense RNA strand is terminal (Figure 9, panel A), it can form a self-priming structure. If, however, both elements are internal (Figure 9, panel B), the downstream shift of the TSS position can make one of them a 3'-terminal and thus enable self-priming. When, on the other hand, the antisense RNA strand has no suitable topologically compatible complementary sequences, an upstream shift of the TSS position (Figure 9, panel C) can generate such elements, one of which is 3'-terminal, and make an mRNA transcript eligible for the amplification process. The events diagrammed in panels B and C of Figure 9 can also occur in reverse, making previously eligible mRNAs ineligible for the mRNA amplification process.

### **Mechanistic and regulatory aspects of the chimeric pathway: Processing of chimeric intermediates and the "Chimeric Junction Shift"**

The hairpin-structured chimeric intermediate of mRNA amplification consists of covalently joined sense and antisense RNA strands and therefore contains an extensive double-stranded region (Figure 3, Step 4). The double-stranded portion of such a structure is separated by a helicase activity invoked above, which mounts the 3'-terminal poly(A) of a newly synthesized sense strand RNA component of the chimeric intermediate and proceeds along this RNA strand in the 5' direction modifying the molecule as it advances (Figure 3, Step 5). When the helicase complex reaches a single stranded portion of the hairpin structure, it, or an associated activity, cleaves the molecule either within the TCE, at a TCE/ICE



**Figure 9: TSS shift as potential regulator of the eligibility of an mRNA for the amplification process.** *Single line:* 3'terminus of the antisense RNA strand. *Filled grey boxes:* Sense RNA strand. *Filled blue boxes:* Topologically compatible complementary elements on the antisense RNA strand. **A:** One of the complementary elements is 3'-terminal; folding results in a self-priming structure that is extended into the sense strand. **B:** Both complementary elements are internal, no self-priming is possible (*top subpanel*); TSS shift in the downstream direction makes one of the elements 3'-terminal and allows self-priming of the antisense RNA and its extension into the sense strand (*bottom subpanel*). **C:** There are no complementary elements, no self-priming structure is formed, no extension occurs (*top subpanel*); TSS shift in the upstream direction generates 3'-terminal complementary element and thus enables self-priming and extension (*bottom subpanel*). *Note* that processes depicted in panels **B** and **C** can occur in reverse, resulting in a loss, rather than the acquisition, of the eligibility.

mismatch or, if there are no mismatches, immediately upstream of the TCE; the cleavage was shown to occur between the 5' hydroxyl group and the 3' phosphate [8,9] (red arrow, Figure 3, Step 6). The cleavage releases the chimeric RNA end product with shortened 5'UTR and the 3'-truncated antisense RNA (Figure 3, Step 7). If the cleavage were to occur at a TCE/ICE mismatch close to and upstream of the 3' end of the TCE, i.e. close to the 3' end of the antisense RNA component of the chimeric intermediate, and a substantial portion of the TCE was retained in the antisense RNA, the self-priming configuration of the antisense RNA could remain stable and be re-used, apparently after 3'-P to 3'-OH phosphatase-mediated conversion, to generate another chimeric intermediate. In such a case, the site of the antisense/sense RNA junction would shift upstream from the previous one by the size of the antisense RNA truncation resulting from the initial cleavage, a phenomenon we refer to as the "chimeric junction shift". If such a process occurs more than once, multiple antisense RNA truncations are generated, and every time the position of the antisense/sense RNA junction is shifted upstream. In this context, multiple versions of chimeric intermediate, differing by the sizes of the 3' truncation of their antisense RNA components, would co-exist concurrently. Eventually, the substantial portion of or the entire TCE element would be cleaved off the antisense RNA and it could not be re-utilized to generate a chimeric intermediate. In fact, such a process of sequential cleavages at different TCE positions was observed with chimeric intermediates of globin mRNAs amplification. As can be seen in Figure 1, one shift of the antisense/sense junction was observed with alpha-globin RNA chimeric intermediates and two shifts were seen with beta-globin RNA chimeric intermediates, thus explaining the origin of the observed chimeric junction sequences with 3'-truncated antisense RNA components. The observed "chimeric junction shift" phenomenon also explains the evident absence of the capG-transcribed 3'-terminal "C" in the antisense components of some chimeric intermediates of the mRNA amplification process.

### **Mechanistic and regulatory aspects of the chimeric pathway: The chimeric RNA end product of mammalian RNA-dependent mRNA amplification is a proverbial Dark Matter mRNA**

If the "visualization" of a nucleic acid species is defined as its detection by hybridization to a labeled complementary probe or the determination of its nucleotide sequence, the modified chimeric RNA end products of the chimeric mRNA amplification pathway can be categorized as a proverbial "Dark Matter" mRNA because it is invisible to the above methods of detection, i.e. hybridization under stringent conditions or nucleotide sequencing based on the reverse transcription of RNA of interest. This is because, as was mentioned above, nucleotide modifications of the chimeric RNA end product of amplification interfere with and prevent both, hybridization under stringent conditions and reverse transcription-based nucleotide sequencing. In this context, therefore, it is important to lucidly define which constituents of the chimeric mRNA amplification pathway containing its distinguishing feature, the "identifier", i.e. sense/antisense RNA junction sequence, can and which cannot be detected by reverse-transcription-based nucleotide sequencing. In the chimeric pathway of mRNA amplification, the cleavage of the chimeric intermediate following strand separation and the associated modifications of its poly(A)-containing chimeric component (sense-oriented RNA and either a portion of or the entire TCE element of the antisense RNA) is the ultimate act in the generation of the chimeric RNA end product. Consequently, it is formed and released

already modified [8,9] and, unlike its hairpin-structured chimeric intermediate precursor, is never present in the unmodified or even partially unmodified form and, therefore, cannot be detected by sequencing methods dependent on the lack of modifications. On the other hand, once the extension of the 3' terminus of unmodified self-primed antisense RNA strand commenced, the antisense/sense RNA junction, the "identifier", is generated and it remains unmodified, and therefore detectable, until the extension process is completed. Moreover, only when the extension of the self-primed antisense RNA strand into the sense-orientation RNA molecule is concluded by the synthesis of the 3' poly(A), transcribed from the 5' poly(U) of the antisense strand, thus forming a double-stranded poly(A)/poly(U) structure, can the strand separation and accompanying nucleotide modification begin. Accordingly, following the completion of the extension and the commencement of strand separation, the sense/antisense RNA chimeric junction remains further not modified for the duration of the strand separation process downstream from the junction region. Thus, once the junction is created, the time required for the completion of the extension process, cumulatively with that required for the strand separation and associated nucleotide modifications downstream from the junction region, constitute a temporal window of opportunity enabling the occurrence of the yet unmodified sense/antisense RNA junction sequences at a certain steady-state level and their detection by conventional hybridization/reverse transcription-based methods.

### **Possible functions of nucleotide modifications in the chimeric RNA end product of Tier One**

The functions of nucleotide modifications occurring during mRNA amplification remain to be elucidated. They could include the facilitation of double-stranded RNA strand separation, stimulation of the cap-independent translation of the cap-less chimeric mRNA end product of amplification, and regulation of its stability. The latter is suggested by a drastic decrease in levels of the chimeric RNA end product of globin mRNA amplification, but not of conventionally produced globin mRNA, during the transition from the erythroblast stage, characterized by rapid hemoglobin accumulation, to the reticulocyte stage where the levels of hemoglobin are only maintained [9] and conventionally produced globin mRNA is apparently sufficient. Arguably, the most important function of nucleotide modifications within the chimeric RNA end product of the amplification process could be that they render the modified mRNA compatible, unlike the bulk of conventional cellular mRNA, with phosphorylated eIF2 $\alpha$ . This implies that in addition to being extraordinarily abundant due to the method of its generation, the amplified, and thus modified, mRNA is also preferentially translated under the ISR conditions, thus augmenting the efficiency of the amplification process. The compatibility with phosphorylated eIF2 $\alpha$  also implies a different mode of translation than that of conventional mRNA. Indeed, the modified RNA end product of the amplification process showed strong affinity to ribosomal RNA [8,9]. It appears that it forms complexes with ribosomal RNA where modifying groups are sequestered, enabling a codon/anticodon interaction [9]. Shortened 5'UTRs, characteristic for the chimeric RNA end product of amplification [8,9], could also play a role in preferential translation of amplified mRNA under the ISR conditions.

### **Tier One summary: Principal players of the chimeric mRNA amplification pathway and their functions**

Beside the competent RdRp complex that includes its inducible

components, generates complete complements of conventional mRNA, and is capable of extending self-primed antisense RNA, the principal players in the chimeric pathway of mammalian RNA-dependent mRNA amplification are the initial amplification template, a conventional gene-transcribed, amplification-eligible mRNA progenitor molecule, and two distinct RNA end products of the amplification process: The chimeric RNA end product and the antisense RNA end product, a 3'-truncated antisense RNA complement of the initial mRNA amplification template. Of these three types of RNA molecules, only the first, a conventional mRNA progenitor, can be used repeatedly in the amplification process. This is because only this molecule can be recurrently transcribed, acquiring nucleotide modifications when utilized as a template and separated from its transcript the first time but otherwise retaining its complete integrity, into the antisense RNA containing both elements, the TCE and the ICE, required for the progression of the amplification process. The resulting chimeric RNA is the "functional" end product, the amplified mRNA, i.e. in most cases it can be translated into a polypeptide. It cannot, however be amplified further in the chimeric pathway because its possible antisense transcript would lack the TCE element and would be, therefore, incapable of self-priming/extension. This leaves the 3'-truncated antisense RNA end product. It cannot be amplified further in the chimeric pathway because it lacks either a substantial portion of or the entire TCE element (transferred to the chimeric RNA end product; the source of the chimerism). Moreover, it has neither role nor function in Tier One of the mRNA amplification process. Is it, then, an inconsequential "by-product" of the chimeric mRNA amplification pathway? The answer, communicated below, is resoundingly negative.

## Tier Two of Mammalian RNA-Dependent mRNA Amplification: Physiologically Occurring Intracellular Polymerase Chain Reaction, iPCR

### Detection of uniformly 5'-truncated and polyuridylylated globin mRNA sequences

In the study of RNA-dependent amplification of globin chains-encoding mRNAs [9], in addition to testing for the presence of chimeric RNA junctions containing both sense and antisense components, a second general search was conducted for 5'-truncated, sense-orientated globin mRNA sequences containing poly(U) segments appended to their truncated 5' termini. The primary reason for this search was the observation by Kapranov and co-investigators [16] of a class of such mRNA molecules in human cancer cells where a robust synthesis of antisense RNA, initiating at the 3'-terminal poly(A) segments of their sense-oriented mRNA counterparts, was also detected. It was reasoned [8,9] that a connection might exist between these two observations and, as discussed below, it appears that it indeed does. In the search for 5'poly(U)-containing globin mRNA molecules [9], cytoplasmic RNA from murine differentiating erythroid tissues was analyzed by directional RNA sequencing.

A selection of 5'-terminal poly(U)-containing fragments detected for both alpha- and beta-globin mRNA is presented in Figure 10 (lines G, H, and I highlighted in gray in both top, alpha-globin, and bottom, beta-globin, panels) together with projected chimeric pathways for alpha- and beta-globin mRNA amplification. In all detected sequences, globin genome-encoded portions are 5'-truncated. The truncations are not random. Interestingly, they are rather uniformly clustered around the nucleotide positions corresponding to the projected

predominant truncations of the antisense RNA "by-product" of the chimeric amplification pathways for both alpha- and beta-globin mRNA. The implications and significance of these observations, and their relevance to the processes of mammalian RNA-dependent mRNA amplification are addressed below.

### The iPCR pathway of mammalian mRNA amplification: Physiologically occurring intracellular polymerase chain reaction

What could be the origin of sense-oriented mRNA sequences with non-conventionally templated 5'-terminal poly(U) appended to the 5'UTRs of mRNA truncated at positions corresponding to the apparent cleavage sites within the antisense components of the chimeric intermediates? As discussed above, the accumulated data strongly suggest that the only major requirement for an RNA molecule to serve as mammalian RdRp template is the presence of the poly(A) segment at its 3' terminus. Provided the presence of the 3'-terminal poly(A) in an RNA molecule is necessary and sufficient for initiation of RNA-dependent RNA synthesis in mammalian cells, in addition to the chimeric pathway of mammalian RNA-dependent mRNA amplification, another mRNA amplification paradigm, constituting Tier Two of the amplification process, may be considered. If an antisense transcript is polyadenylated at the 3' terminus by a known or a novel cellular poly(A) polymerase, it would become a valid template for mammalian RdRp. Since the antisense RNA strand has, by virtue of its initiation within the 3' poly(A) region of a conventional mRNA, the poly(U) stretch at the 5' terminus, its transcription by the RdRp complex would result in a sense strand with the poly(U) at the 5' end and the poly(A) at the 3' end, also a legitimate mammalian RdRp template. Since strand separation mechanisms (helicase complex and associated activities discussed above) are in place and the described sequence of events can occur repeatedly, the process will amount to an intracellular polymerase chain reaction, iPCR. If, in such a process, the initial 3' polyadenylation of the antisense RNA is coupled and occurs in conjunction with, and at the site of, the cleavage of the chimeric intermediate, the 5' truncation of the gene-encoded portion of the sense-oriented mRNA transcript from such an antisense RNA template would correspond precisely to the 3' truncation of the antisense RNA end product ("by-product") of Tier One of mRNA amplification, exactly so as was observed experimentally (Figure 10).

Such sequence of events is presented diagrammatically in Figure 11. Up to the Step 6, all processes occur exactly as are shown in Figure 3 and described above. In steps 6/7, the cleavage at a TCE/ICE mismatch or at the 5' end of the TCE (red arrow) is coupled with polyadenylation of a newly generated 3' terminus of the truncated antisense RNA strand. This event marks a conclusion of the chimeric cycle of amplification. One of its end products, the chimeric RNA marked in Figure 11 as "END PRODUCT, TIER ONE", is identical to the chimeric end product of Figure 3. The other end product of the chimeric pathway of mRNA amplification is different from its equivalent in Figure 3. It is the antisense RNA strand, 3'-truncated in the same position as depicted in Figure 3, but in addition to the 5'-terminal poly(U), it also contains the 3'-terminal poly(A) segment (Figure 11, Step 7). This molecule constitutes the initial template of the intracellular polymerase chain reaction, its "progenitor". Indeed, just as in Step 1, RdRp activity initiates transcription within the 3'poly(A) of the antisense RNA template and generates the sense RNA strand containing both the 5'poly(U) segment and the 3'poly(A) segment (Figure 11, Step 8). After strand separation (shown in Figure 11, Step 9 in a simplified form, since strand separation probably commences



while the poly(U)-containing RNA strand is still nascent), there are now two templates, each containing the 3'-terminal poly(A) as well as the 5'-terminal poly(U) segments (Figure 11, Step 10; note that only one of possibly multiple iPCR rounds is shown), and the iPCR is under way. Strand separation and associated nucleotide modifications probably commence at the 3'-terminal poly(A) as soon as it becomes double-stranded, i.e. when the poly(U)-containing RNA strand is still nascent or even when only a sufficiently long 5'-terminal poly(U) segment has been generated; in such a case, the helicase complex would be trailing closely the progression of the RdRp complex. Therefore, as detailed further in the "Mechanistic Aspects" section below, at steady-state there would be many more not yet modified, and thus detectable, 5'-truncated, -terminal poly(U)-containing sense RNA strand sequences than 3'-truncated, -terminal poly(A)-containing antisense RNA strand sequences; both types of sequences are uniquely specific for the iPCR pathway and constitute its "identifiers". The above considerations apply equally to the opposite termini of the same iPCR-amplified molecules: At steady state there would be many more unmodified 5'-terminal poly(U)-containing antisense RNA sequences than 3'-terminal poly(A)-containing sense RNA sequences; these types of sequences, however, would be indistinguishable from corresponding segments of either conventional mRNA or intermediates and products of the chimeric pathway.

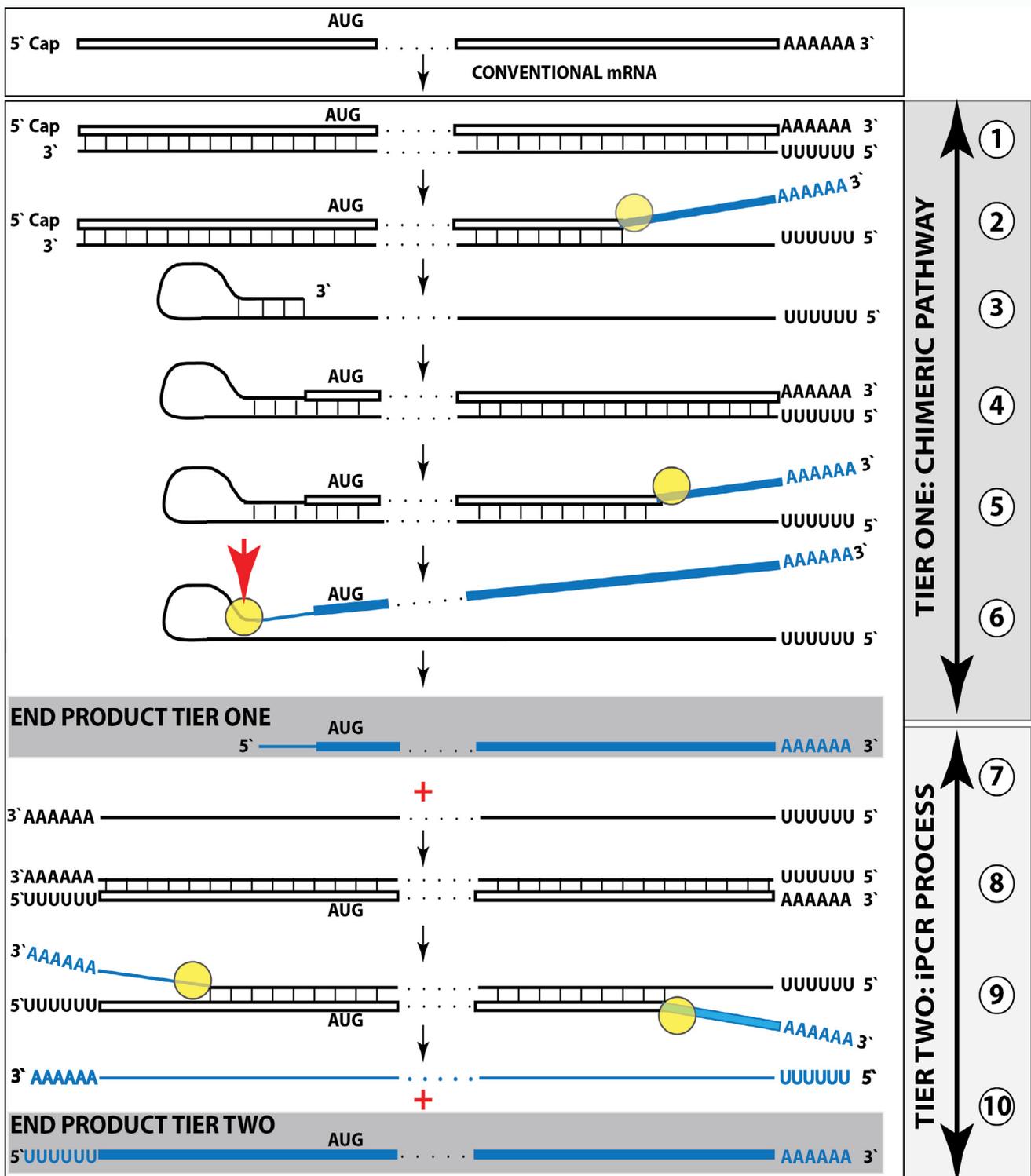
### **Mechanistic and regulatory aspects of Tier Two of mammalian RNA-dependent mRNA amplification, the iPCR pathway**

Conceptually, amplification of a nucleic acid molecule by a polymerase chain reaction necessitates, beside the presence of the building blocks, the occurrence of a template, the priming arrangement for the initial nucleic acid strand and for its complement, a suitable polymerase, and the procedure for strand separation that doubles the number of template molecules in each cycle. In a conventional PCR reaction, a single-stranded DNA molecule serves as a template, oligonucleotides complementary to the initial DNA strand and to its complement in desired and appropriate positions function as primers (the "forward" and the "return" primers), DNA polymerase of choice extends the 3' end of a primer, generating a double-stranded molecule, and a thermal treatment separates the strands to enable the next cycle of the chain reaction where both complementary DNA strands serve as templates; with each cycle the number of molecules, and thus of templates, in the reaction doubles.

In the intracellular PCR process, iPCR, the eligible initial template is a single-stranded RNA molecule, the 3'-truncated antisense RNA end product of the chimeric mRNA amplification pathway, containing the poly(A) segment at its 3' terminus and the poly(U) segment at its 5' terminus. The process is driven by the mammalian RdRp complex, which initiates transcription within the poly(A) segment and generates a complete complement of the initial amplification template. The priming arrangements for both the initial template and its complement are reflected in the template's RdRp eligibility requirements that are satisfied by the occurrence of the poly(A) segment at the 3' termini of both the initial RNA template as well as of its transcript/complement, and strand separation is carried out by the helicase/modifying activity described above in the discussion of Tier One of mRNA amplification. In the chimeric pathway, when a full-length 5' poly(U)-containing antisense RNA strand is generated (transcribed by the RdRp complex from a conventional mRNA progenitor) and separated from its template (steps 1 and 2, Tier One, Figure 11), it is not an iPCR-eligible template because it lacks the 3'-terminal poly(A) segment. Instead,

provided that it contains the topologically compatible 3'-terminal and internal complementary elements, it self-primers its extension into a sense RNA strand; thus generating the intermediate in the chimeric pathway of mRNA amplification. It is the processing of this intermediate that has the potential to produce, in addition to the chimeric RNA end product, the iPCR-eligible template. This requires chimeric intermediate cleavage-coupled 3' polyadenylation of the cleavage-truncated antisense RNA strand, which already contains the 5'-terminal poly(U) segment transcribed from the 3'-terminal poly(A) of a conventional mRNA progenitor molecule. The presence of the 3'-terminal poly(A) segment on the antisense RNA would allow the RdRp complex to initiate its transcription and to proceed with the synthesis of the sense RNA strand complement that commences with the 5'-terminal poly(U) and concludes with the 3'-terminal poly(A) segments, also an iPCR-eligible template. The following separation of strands by the helicase/modifying activity would double the number of templates and enable the next cycle of a polymerase chain reaction. Thus, the key feature underlying the feasibility of iPCR is that both the initial cleavage/polyadenylation-released antisense RNA molecule and its transcript are eligible RdRp templates. Potentially and purely hypothetically, if an antisense RNA transcribed from a conventional mRNA were polyadenylated at the 3' end upon conclusion of its transcription, it would, after strand separation, become an eligible iPCR template. This, however, is highly unlikely physiologically, because such a process would completely lack the specificity. In the Two-Tier amplification process, because the generation of the initial iPCR progenitor template is coupled with and enabled by the concluding step of the previous Tier, the specificity of the iPCR pathway is precisely equal in its stringency to that of the preceding chimeric pathway, which, in turn, is defined by the occurrence within an antisense RNA transcribed from a conventional mRNA of the topologically compatible TCE and the ICE features and by the thermodynamic feasibility allowing the initial antisense transcript to properly fold and self-prime its extension. On the other hand, importantly, whether or not Tier Two of mammalian mRNA amplification, iPCR, occurs does not affect in any way Tier One, the chimeric pathway.

The translationally functional end product of the iPCR amplification pathway is a non-chimeric sense strand RNA identical to a conventional genome-encoded mRNA in informational content and in all other attributes except three – it contains modified nucleotides, it is 5'-polyuridylylated (and apparently not capped), and it is 5'-truncated in its genome-encoded portion. The truncations are no larger than the TCE of a corresponding antisense RNA, and since TCEs appear to be relatively short, significantly shorter than a typical mammalian 5'UTR, so would the 5' truncations be. Mechanisms underlying the iPCR Tier of amplification could explain the not yet elucidated observations in human cancer cells of a class of unconventional mRNA sequences that differ from their conventional counterparts only in two aspects: (a) their genome-encoded portions are truncated at the 5' end, typically by 14 to 18 nucleotides, not affecting their protein-encoding capacity, and (b) they contain poly(U) segments appended to their truncated 5' termini [16]. In terms of the Two-Tier mRNA amplification process, such molecules could be products of the second Tier of amplification, the iPCR pathway, and the extent of 5' truncations in their genome-encoded portions could reflect the average size of the antisense 3' truncation at the conclusion of the preceding chimeric pathway Tier, i.e. the distance between the cleavage site within the chimeric intermediate and the 3' end of the antisense strand, or, in other words, the average size of their TCE elements.



**Figure 11: Mammalian RNA-dependent mRNA amplification is a Two-Tier process.** *Top panel:* Conventional, genome-originated mRNA molecule. *Bottom panel:* Tier One, the chimeric pathway, progressing into Tier Two, the iPCR pathway. *Boxed line:* Sense strand RNA. *Single line:* Antisense strand RNA. "AUG": Functional translation initiation codon (could be other than "AUG"). *Yellow circle:* Helicase/modifying activity complex. *Blue lines* (both single and boxed): RNA strands, both sense and antisense, modified and separated from their complements by a helicase complex. *Red arrowhead:* Position of cleavage of the chimeric intermediate. **Step 1:** Synthesis of antisense RNA strand; **step 2:** Strand separation; **step 3:** Folding of the antisense RNA strand into self-priming configuration; **step 4:** Extension of self-primed antisense RNA into sense RNA; **step 5:** Strand separation; **step 6:** Cleavage of the chimeric intermediate coupled with 3' polyadenylation of the truncated antisense RNA; **step 7:** End-products of the chimeric pathway of the mRNA amplification process; **step 8:** RdRp-mediated synthesis of the sense strand initiated at the 3' poly(A) of antisense RNA; **step 9:** Strand separation; **step 10:** iPCR products. Each RNA strand constitutes an iPCR template. The antisense RNA can be further amplified whereas the sense RNA strand can be used either for amplification or for translation. *Note that the Figure is simplified:* (a) Initiation of transcription at the 3'-terminal poly(A) apparently occurs within the poly(A) segment rather than at its end; (b) Strand separation shown in steps 2 and 9 apparently commences while the poly(U)-containing RNA strand is still nascent, as described in text.

Whereas the occurrence and regulation of Tier One of mammalian mRNA amplification, the chimeric pathway, is entirely independent from Tier Two, the iPCR process, the occurrence of the latter depends on and is impossible without the completion of the former. Moreover, the occurrence of the iPCR also requires that the last step of the chimeric pathway, the cleavage of the chimeric intermediate, were coupled with the 3' polyadenylation of the cleavage-truncated antisense RNA strand. Arguably, the efficiency of the iPCR pathway may vary in a regulated manner in accordance with the needs for the production of a protein encoded by an mRNA species being amplified. Considering that the presence of the 3'-terminal poly(A) segment is presumably a single major template eligibility requirement for the initiation of RdRp-mediated RNA-dependent RNA synthesis [6,8,9], and allowing for the operational presence of the RdRp complex, a key enzymatic component of both the chimeric pathway and the iPCR process, it is reasonable to suggest that the principal regulatory checkpoint for the entry into Tier Two of mRNA amplification could be the rate of the chimeric intermediate cleavage-coupled 3' polyadenylation of the cleavage-truncated antisense RNA strand. In the extreme case, if the polyadenylation (but not the cleavage) were completely suppressed, so would Tier Two be, but neither the outcome nor the efficiency of the Tier One process would be affected in any way.

In the chimeric pathway of mammalian mRNA amplification, the RdRp complex appears to initiate transcription within the 3'-terminal poly(A) segment rather than at its 3' end. Possibly, the initiation "within" the poly(A) segment is not a stringent requirement, and only the distance from the encoded 3'UTR of mRNA is important; then RdRp can also initiate at the end of the 3'poly(A) of the right size range. Otherwise, it can be assumed that the initiation within the 3'poly(A) is true for the iPCR process as well as for the chimeric pathway. In such a case, successive iPCR rounds would result in the shortening of both the 5'-terminal poly(U) segments and, consequently, the 3'-terminal poly(A) segments. This, in turn, would eventually lead to the loss of template's RdRp eligibility unless the 3'-terminal poly(A) elongation, known to occur in the cell, takes place in parallel with its shortening. This is in contrast to the chimeric mRNA amplification pathway where repeated utilization of conventional mRNA as a template for the antisense RNA synthesis apparently has no impact on the length of its 3'-terminal poly(A). Additional augmentation of the efficiency of the iPCR pathway could occur at the level of the iPCR templates availability, which is comprised of two components: stability of both sense and antisense strands produced in the reaction and the utilization of the mRNA end product (sense-oriented strand) of the second Tier. The iPCR is, as any chain reaction, potentially an exponential process provided that the entire output of every cycle is utilized as templates in the next cycle of a chain reaction. However, if all sense strand RNA products in every cycle were withdrawn due to the recruitment for protein synthesis, this Tier of amplification would be linear, akin to the chimeric pathway. In all probability, some sense-oriented iPCR-produced molecules are removed for translation and some are utilized as templates for further amplification, making the process faster than linear but slower than exponential. As for the stability of the iPCR templates, it could be regulated, at least in part, by nucleotide modifications, present in both RNA strands, as was discussed in the "chimeric pathway" sections above.

Considering that the RNA-dependent mRNA amplification process proceeds through the double-stranded RNA stages, an obvious question is how it evades (if it does) the dsRNA response? One possibility is that the amplification process is compartmentalized

and the components of the dsRNA response have no access to it. Another possibility is the involvement of factors protecting from the dsRNA response. What is known experimentally, is that unmodified double-stranded RNA components associated with RNA-dependent mRNA amplification appear to be extremely short-lived as reflected in their very low steady-state levels [8,9]. This means that strand separation and associated nucleotide modifications occur very rapidly and double-stranded molecules may have very limited exposure. Moreover, provided that one of functions of modifications is strand separation, when already modified RNA is used as a template for RdRp, as in the re-use of conventional mRNA in Tier One, or as occurs routinely in Tier Two, the iPCR pathway, strand separation would take place concurrently with the progression of the RdRp complex and no double-stranded RNA exposure would result. An interesting alternative possibility is that the occurrence of both Tiers of mammalian RNA-dependent mRNA amplification generates a certain steady-state level of double-stranded RNA, which, in turn, supports a steady-state population of activated double-stranded RNA-dependent kinase PKR, a member of the family of eIF2 $\alpha$  kinases. Consequently, eIF2 $\alpha$  is phosphorylated, the integrated stress response elicited, the production of inducible components of the RdRp complex sustained, and the mRNA amplification process perpetuated.

Whereas in Tier One only the chimeric RNA end product is modified because the hairpin-structured chimeric intermediate contains a single linear double-stranded 3'poly(A)/5'poly(U) region, a requirement and the starting point for commencement of the strand separation/modification (initiating on and proceeding along and modifying the poly(A)-containing strand), in the iPCR pathway such regions are present at both ends of a linear double-stranded structure. Accordingly, strand separation and associated nucleotide modifications commence on the poly(A) segments of both, sense and antisense, RNA strands and, therefore, both strands are eventually modified. The distinguishing features of the iPCR pathway Tier, its "identifiers", are the 3'-terminal region of the cleavage-truncated (or transcribed from the sense 5'poly(U)/3'poly(A)-containing RNA in Tier Two) antisense strand RNA containing the 3'-terminal poly(A) segment, and the 5'-terminal region of its corresponding sense strand RNA transcript, with a 5' truncation mirroring a 3' truncation of the antisense strand and containing the 5'-terminal poly(U) segment; both "identifiers" are, in fact, another kind of "junctions". Unlike the chimeric pathway, where the chimeric RNA end product is never present even in partially unmodified form, and its "identifier", a sense/antisense RNA junction, cannot be detected by reverse-transcription based methods, the iPCR pathway presents possibilities to detect unmodified "identifiers" of Tier Two in both of its RNA end products. The detection opportunities, however, are dissimilar for the sense and the antisense RNA-derived iPCR "identifiers". The 3'-truncated and polyadenylated antisense RNA, produced either by the cleavage/polyadenylation of the chimeric intermediate or transcribed and separated from the sense RNA strand containing both 5'poly(U) and 3'poly(A), is unmodified but not for a long duration. It is, probably, very rapidly utilized as RdRp template and as rapidly separated (and modified, thus becoming undetectable, in the process) from still nascent 5'poly(U)-containing sense strand being transcribed from it. Alternatively, this duration could be even shorter, comprised only of the time period between the completion of the 3' polyadenylated antisense RNA synthesis and the commencement of separation/modification starting at its newly added poly(A) as soon as a double-stranded 3'poly(A)/5'poly(U) structure, an initial helicase substrate, is

generated. Thus, a temporal window of opportunity for the detection of an unmodified, and thus detectable, 3'poly(A)-containing antisense RNA-derived "identifier" of the iPCR pathway, and, correspondingly, its steady-state level, are slight.

A temporal window of opportunity is much greater for the detection of an unmodified 5'poly(U)-containing sense RNA-derived "identifier" of Tier Two of mammalian mRNA amplification. After the commencement of the sense RNA synthesis starting with the 5'poly(U), transcribed from the 3'poly(A) of the antisense RNA, the nascent sense RNA strand remains unmodified until its completion by the synthesis of the 3'poly(A) segment transcribed from the 5'poly(U) segment of the antisense RNA template, which is being modified and separated (unless it is being re-used as a template and therefore is already modified) by a helicase/modifying activity trailing the RdRp complex. Only when a nascent antisense RNA strand is being synthesized starting at the 3'poly(A) of the newly produced sense strand RNA, and the double-stranded 3'poly(A)/5'poly(U) structure is present, can the strand separation of and associated nucleotide modifications in the 5'poly(U)-containing sense RNA strand begin at its 3', the poly(A)-containing, end. And even then, the 5'poly(U)-containing "identifier" remains unmodified for the duration of the antisense RNA synthesis and strand separation/nucleotide modification downstream from it. Thus, once transcription of the sense RNA is initiated, starting with the 5'poly(U) and continuing into a truncated 5'UTR, the "identifier" is generated, and the time required for the completion of transcription, cumulatively with that needed for strand separation and associated nucleotide modifications downstream from the "identifier", provide a relatively large temporal window of detection opportunity. Consequently, at steady-state, there are many more unmodified, and therefore detectable by reverse-transcription based methods, 5'poly(U)-containing sense strand RNA-derived "identifiers" of the iPCR process than their 3'poly(A)-containing antisense strand RNA-derived counterparts. It should be remembered, however, that the detectable, not yet modified, "identifiers" of the iPCR pathway represent only a very minor subpopulation of the products of Tier Two of mammalian mRNA amplification, with the bulk of its fully modified and thus undetectable products remaining for now, alongside with highly abundant chimeric RNA end product of Tier One, a "Dark Matter" RNA.

### The Two-Tier paradox: Asymmetry of outcomes

Of the four Types of the chimeric pathway of mammalian RNA-dependent mRNA amplification described above, each constitutes a variant of Tier One of the mRNA amplification process, is defined by its translational outcome, and potentially can progress to Tier Two. Thus, the Type I of the chimeric pathway is characterized by the location of the ICE element of the antisense RNA strand within its segment corresponding to the 5'UTR of the conventionally produced mRNA. Consequently, the chimeric RNA end product, resulting from this Type of the chimeric Tier of amplification, contains the intact protein-encoding information of a conventional mRNA progenitor, and the translational outcome would be a polypeptide identical to the conventionally produced protein. If this Type of the chimeric pathway were to progress to Tier Two, the iPCR pathway, the functional outcome would be the same as in Tier One, namely the amplified sense strand RNA that retains a complete coding information content of a conventional mRNA and produces a polypeptide identical to the conventional mRNA-encoded protein. In the Type II of the chimeric pathway, the ICE is located within a segment of the antisense RNA

strand corresponding to the coding region of a conventional mRNA and one of the possible outcomes is the chimeric RNA end product encoding a CTF of a conventional polypeptide. If this Type of the chimeric pathway were to progress to Tier Two of the mRNA amplification process, one may intuitively assume that the iPCR-amplified sense RNA strand would also encode the same CTF as the chimeric RNA end product of Tier One and that, in general, any particular translational outcome of Tier One would carry on into Tier Two. This, however, is not the case because of the phenomenon of "Two-Tier Paradox".

The Two-Tier Paradox is defined as follows: Regardless of the translational outcome of the chimeric pathway Tier of mRNA amplification – a conventional polypeptide, a CTF, an unrelated polypeptide, and even an abortive translation or translational incompetence – the sense RNA strand produced in Tier Two, the iPCR pathway, would always retain the entire coding information of a genome-transcribed mRNA progenitor and would always direct translation of a polypeptide identical to its conventional mRNA-encoded counterpart.

As was mentioned above, whether Tier Two of mammalian RNA-dependent mRNA amplification, the iPCR process, does or does not occur is completely inconsequential for Tier One, the chimeric pathway. In turn, any particular translational outcome of Tier One does not affect in any way the translational outcome of Tier Two. This is because, in sharp contrast to and irrespective of any of the potentially multiple translational outcomes of the chimeric pathway, there is only one possible translational outcome of the iPCR process of mammalian mRNA amplification: A polypeptide identical to that produced by a conventional, genome-transcribed mRNA progenitor of the amplification process. This discordance reflects the asymmetry in the truncation of sense and antisense constituent components of the chimeric intermediate during its cleavage at the conclusion of Tier One. The reason for this asymmetry is that the truncations of molecules in question are determined by two different, spatially independent, elements. The truncation of the sense strand component of the chimeric RNA end product is defined by the position of the ICE element whereas the TCE position determines the truncation of the antisense strand. Because the ICE can be anywhere in the antisense RNA molecule, so can the sense RNA strand's 5' truncation be, and because the TCE, by definition, can be only 3'-terminal, the antisense RNA always loses only a short segment at its 3' end. This phenomenon is displayed in Figure 4. Whereas the sizes of the chimeric RNA end products are drastically different in steps 7 and 7', the sizes of the antisense RNA end products in the same steps are identical (the same is true for steps E and E' of Figure 5). The 3'-truncated antisense RNA, provided it is 3'-polyadenylated at the conclusion of Tier One, gives rise in Tier Two to the 5'poly(U)-containing sense RNA strand 5'-truncated by exactly the same extent as seen in the 3' antisense RNA truncation, which in most, if not all, instances is substantially shorter than a typical 5'UTR of a conventional mammalian mRNA [81]. Consequently, the 5' poly(U)-containing sense RNA molecule, derived from the 3' poly(A)-containing antisense progenitor (the initial iPCR template) via the iPCR pathway, retains the bulk of its 5'UTR and the intact protein-encoding portion of a corresponding conventional mRNA. Therefore, regardless of the translational outcome of Tier One, even if it is an abortive translation or translational incompetence, in Tier Two the translational outcome would always be a polypeptide identical to that attained with a conventional mRNA amplification progenitor. The translational outcome of Tier

Two would match that of Tier One only when the ICE is positioned within the antisense RNA segment corresponding to the 5'UTR of conventional mRNA progenitor. In all other cases, the outcomes can be quite remarkable. It is conceivable, for example, that the outcome of Tier One is an abortive translation or translational incompetence, yet this "failure" could be "rescued" by the Tier Two process, which would produce an intact conventional polypeptide encoded by an mRNA progenitor. No less remarkable could be cases where Tier One results in a CTF of a conventional protein whereas Tier Two would produce, in parallel, the intact conventionally encoded polypeptide. The possible beta-APP-independent generation of beta-amyloid in Alzheimer's disease (discussed below) provides a good illustration of such a phenomenon. If the antisense beta-APP RNA strand, resulting from the cleavage of the chimeric intermediate and lacking only its TCE (or a fraction thereof), were 3'-polyadenylated in a cleavage-coupled manner, and utilized as the initial iPCR template in the second Tier of the mRNA amplification process, it would give rise to a 5'-truncated and poly(U)-containing sense strand retaining the bulk of the 5'UTR and the complete coding region of the conventional beta-APP mRNA. Following its translation, the intact beta-APP polypeptide would be generated alongside its beta-amyloid-carrying CTF, produced from the chimeric RNA with 5'-truncated coding region in the beta-APP-independent manner, in the first, chimeric pathway, Tier of the mRNA amplification process. Even more striking is the possibility of a polypeptide unrelated to a corresponding conventionally produced protein and a translation product identical to it being generated concurrently by the chimeric pathway and the iPCR process respectively.

### Reporter Constructs for the Chimeric Pathway of Mammalian RNA-Dependent mRNA Amplification

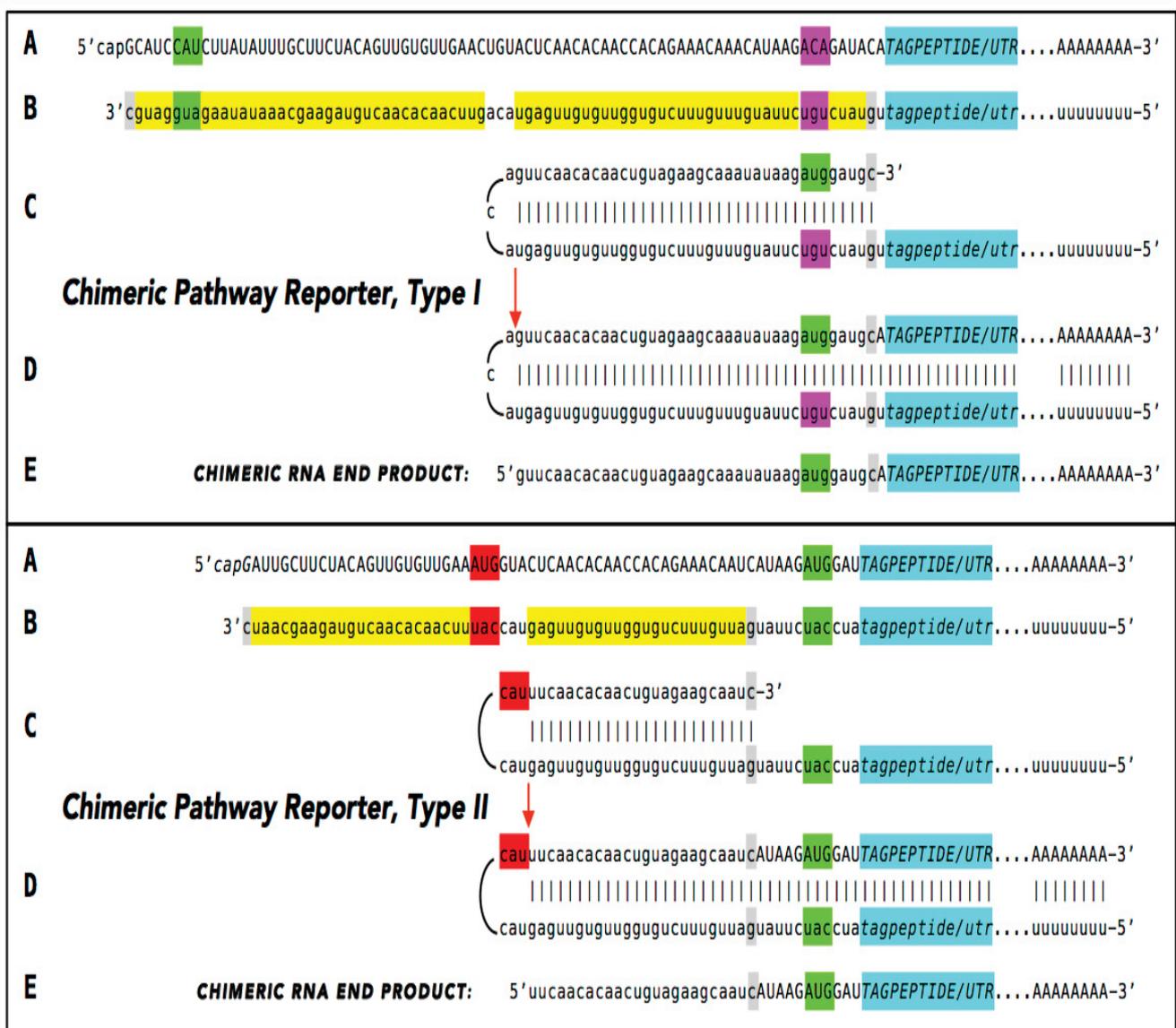
The availability of reporters for both Tiers of mammalian RNA-dependent mRNA amplification is essential for the further development of the field. To be effective, these reporter constructs should satisfy quite stringent requirements, namely to generate no "reporting" protein product (presuming that it is a protein, which does the "reporting") in the conventional gene expression process but to produce it only in the pathway the construct is designed to "report" on. To meet these requirements is not straightforward. Fortunately, a certain feature of the amplification process affords the opportunity to fulfill them. The feature in question is the manner of generation and processing of the chimeric intermediate that leads to both the loss of a portion of conventionally produced mRNA progenitor in the protein-encoding end products of both Tiers of mRNA amplification, and to the acquisition in the chimeric RNA end product of a component not present in the initial mRNA progenitor. This feature allows to implant in and/or to remove from the mRNA end products of amplification elements regulating the production of a "reporting" polypeptide.

Accordingly, a reporter construct for the chimeric pathway of mRNA amplification should be designed in such a way as to allow the expression of a "reporting" polypeptide only in this pathway and to preclude it when the pathway is not in operation. Conceptually, there are two ways to do this. One approach (Type I) is to generate an RNA transcript eligible for the RNA-dependent amplification process and encoding a "reporting" polypeptide, say a tag peptide, but lacking an element required for its translation. Such transcript is "silent" and cannot be translated into a "reporting" protein product under conventional circumstances. However, if this RNA transcript is undergoing RNA-dependent mRNA amplification, an

element needed for its translation can be provided by the antisense component of the chimeric RNA end product of amplification. Such design is based on the concept of "Type IV of the Chimeric mRNA Amplification Pathway" described above. Second approach (Type II) is to generate an amplification-eligible RNA transcript encoding a tag peptide and containing all necessary elements for its expression but also incorporating an upstream element preventing translation of the desired polypeptide. This inhibitory upstream element is then eliminated in the 5'-truncated sense RNA component of the chimeric RNA end product of Tier One of mRNA amplification. In both Types of construct design, the TCE and ICE elements are placed in close proximity to ascertain their anticipated interaction.

An example of the Type I of such reporter constructs, encoding a tag peptide that is produced only in the RNA-dependent mRNA amplification chimeric pathway, is shown in the upper panel of Figure 12. In the amplification-eligible RNA transcript, *TAGPEPTIDE/UTR* highlighted in blue denotes nucleotide sequence encoding a tag peptide followed by the 3'UTR and lacking the preceding AUG translation initiation codon, which is replaced with the ACA codon (highlighted in pink, as is its complement in the antisense RNA). There is also no functional translation initiation codon upstream from and in-frame with the tag peptide-encoding segment; no tag peptide can be translated from this RNA, it is translationally "silent" (line A). In the corresponding antisense RNA (line B), *tagpeptide/utr*, highlighted in blue, denotes antisense complement of the *TAGPEPTIDE/UTR* nucleotide sequence; it also contains the 3'-terminal TCE and the internal ICE sequences (both highlighted in yellow). Moreover, the TCE contains the 5'aug3' codon (shown in 3' to 5' orientation and highlighted in green, as is its complement in the sense RNA strand) in optimal translation initiation context. During the antisense RNA folding into a self-priming configuration (line C), the "aug" codon in the TCE is accommodated by the "ugu", complement of the ACA codon highlighted in pink in the sense RNA. The design of the construct ascertains that the additional 3'-terminal "C" of the antisense RNA, not encoded in the DNA but a transcript of the 5'cap "G" of the sense-strand RNA [8-10], can be accommodated in the folded self-priming configuration (both 3'-terminal "C" of the antisense RNA and matching "G" at the 5' end of the ICE are highlighted in gray). The self-primed structure is then extended, strands are separated and the cleavage occurs (red arrow) at the 5' end of the TCE (line D). In the chimeric RNA end product of amplification (line E), the TCE-derived "aug" codon, highlighted in green, is positioned, in optimal translation initiation context, upstream from and in-frame with the tag peptide-coding sequence and can initiate its translation. If the amplification process progresses to Tier Two, no tag peptide will be produced in the iPCR pathway, consistent with the Two-Tier Paradox principles (Tier Two translational outcome is identical to that attained with conventional mRNA). With this construct, a tag peptide is produced only if and when the chimeric mRNA amplification pathway is operational; its occurrence reports the activity of the pathway.

An example of the Type II of a reporter construct for the chimeric pathway of mammalian mRNA amplification is presented in the lower panel of Figure 12. mRNA transcribed from this construct (line A) is eligible for RNA-dependent amplification and encodes a tag peptide (*TAGPEPTIDE/UTR* highlighted in blue denotes nucleotide sequence encoding tag peptide followed by the 3'UTR). The tag peptide-encoding segment is preceded by the in-frame AUG codon (highlighted in green, as is its complement in the antisense RNA) in optimal translation initiation context. Yet this tag peptide-



**Figure 12: Design of RNA transcripts from reporter constructs producing a “reporting” polypeptide solely through Tier One, the chimeric mRNA amplification pathway.** *Uppercase letters:* Nucleotide sequence of the sense RNA. *Lowercase letters:* Nucleotide sequence of the antisense RNA. *TAGPEPTIDE/UTR*, highlighted in blue, denotes nucleotide sequence encoding a tag peptide followed by 3'UTR; *tagpeptide/utr*, highlighted in blue, denotes complement of the *TAGPEPTIDE/UTR* nucleotide sequence in the antisense RNA. *Highlighted in green:* 5'AUG3' in optimal translation initiation context or its complement on the opposite-oriented RNA strand. *Highlighted in red:* 5'AUG3' in optimal translation initiation context or its complement on the antisense RNA strand. *Highlighted in yellow:* the 3'-terminal complementary element, TCE, and internal complementary element, ICE, both on the antisense RNA. *Highlighted in gray:* 3'-terminal “c” of the antisense RNA not encoded in the DNA, a transcript of the cap“G” of mRNA, and matching “g” at the 5' end of the ICE on the antisense RNA. *Red arrow:* Position of cleavage of chimeric intermediate following strand separation. *Top panel:* **A** – Translationally silent RNA transcript encoding tag peptide and lacking the AUG translation initiation codon that was replaced with the ACA codon highlighted in pink, as is its complement in the antisense RNA (on the choice of the “ACA” see below); there are also no functional translation initiation codons upstream from and in-frame with the tag peptide-encoding segment. **B** - Antisense RNA; note that it contains the TCE and ICE elements and the 5'aug3' in optimal translation initiation context (shown as 3'gua5' and highlighted in green). **C** - Antisense RNA folded into self-priming configuration; note that the 5'aug3' is accommodated by the 3'ugu5' highlighted in pink, a complement of the “ACA” in the sense strand (the reason for its choice), and that the 3'-terminal “c”, a transcript of the capG, is accommodated by the “g” (both in gray). **D** - Extension of self-primed antisense RNA generates the chimeric intermediate containing covalently bound sense and antisense RNA strands; position of cleavage following strand separation is indicated by red arrow. **E** - Chimeric mRNA end product of mRNA amplification; note that the “aug” translation initiation codon in optimal translation initiation context (contributed by the antisense portion of the chimeric RNA) is in-frame with the tag peptide-encoding nucleotide sequence in the sense portion of the chimeric RNA end product and can initiate its translation. *Note* that no tag peptide can be produced via the iPCR process due to the lack of functional translation initiation codon in its sense RNA end product. *Bottom panel:* **A** – Tag peptide-encoding segment in sense RNA strand is preceded in-frame by the AUG codon, highlighted in green, in optimal translation initiation context. However, tag peptide cannot be produced since this RNA contains an additional upstream “AUG” codon, highlighted in red (as is its complement in the antisense RNA), also in optimal translation initiation context, but out-of-frame with tag peptide-encoding nucleotide sequence. **B** - Antisense RNA containing the TCE and ICE elements. **C** - Antisense RNA folded into self-priming configuration; note that the 3'-terminal “c”, a transcript of the capG, is accommodated by the “g” (both in gray). **D** - Extension of self-primed antisense RNA generates the chimeric intermediate; position of cleavage following strand separation is indicated by red arrow. **E** - Chimeric RNA end product of RNA-dependent mRNA amplification; the upstream AUG has been removed and the chimeric RNA contains only one “AUG” translation initiation codon, in optimal translation initiation context, preceding in-frame the tag peptide-encoding segment and capable of initiating its translation.

coding information cannot be translated because upstream from it is located another AUG (highlighted in red, as is its complement in the antisense RNA), also in optimal translation initiation context but out-of-frame with the tag-peptide-encoding nucleotide sequence, and it is this upstream AUG that will initiate translation and thus preclude production of the tag peptide. In the corresponding antisense RNA (line B), *tagpeptide/utr*, highlighted in blue, denotes antisense complement of the *TAGPEPTIDE/UTR* nucleotide sequence. It contains the 3'-terminal TCE and the internal ICE sequences (both highlighted in yellow). Its folded configuration (line C) can accommodate the additional 3'-terminal "C" of the antisense RNA, not encoded in the DNA but a transcript of the 5'cap "G" of the sense-strand RNA (both 3'-terminal "C" of the antisense RNA and matching "G" at the 5' end of the ICE are highlighted in gray). The 3' terminus of the self-priming structure is then extended, strands are separated and the cleavage occurs (red arrow) at the 5' end of the TCE (line D). If the upstream "inhibitory" AUG (highlighted in red) in the initial RNA transcript is placed within a segment of mRNA corresponding either to the TCE, or to the loop region of folded antisense RNA (as shown in the Figure), it would be lost in the chimeric RNA end product of amplification (line E). The same result could be achieved by placing the upstream AUG within a segment of mRNA corresponding to the ICE element of the antisense RNA, but this could lead to complications stemming from a potential "chimeric junction shift" and generation of an additional upstream AUG codon in the chimeric RNA end product, and may require additional precautions. In this process, the resulting chimeric RNA end product retains only the AUG preceding, in-frame and in optimal translation initiation context, the tag peptide-encoding sequence. With such construct, a reporter peptide would be generated only when the RNA-dependent mRNA amplification process has been activated.

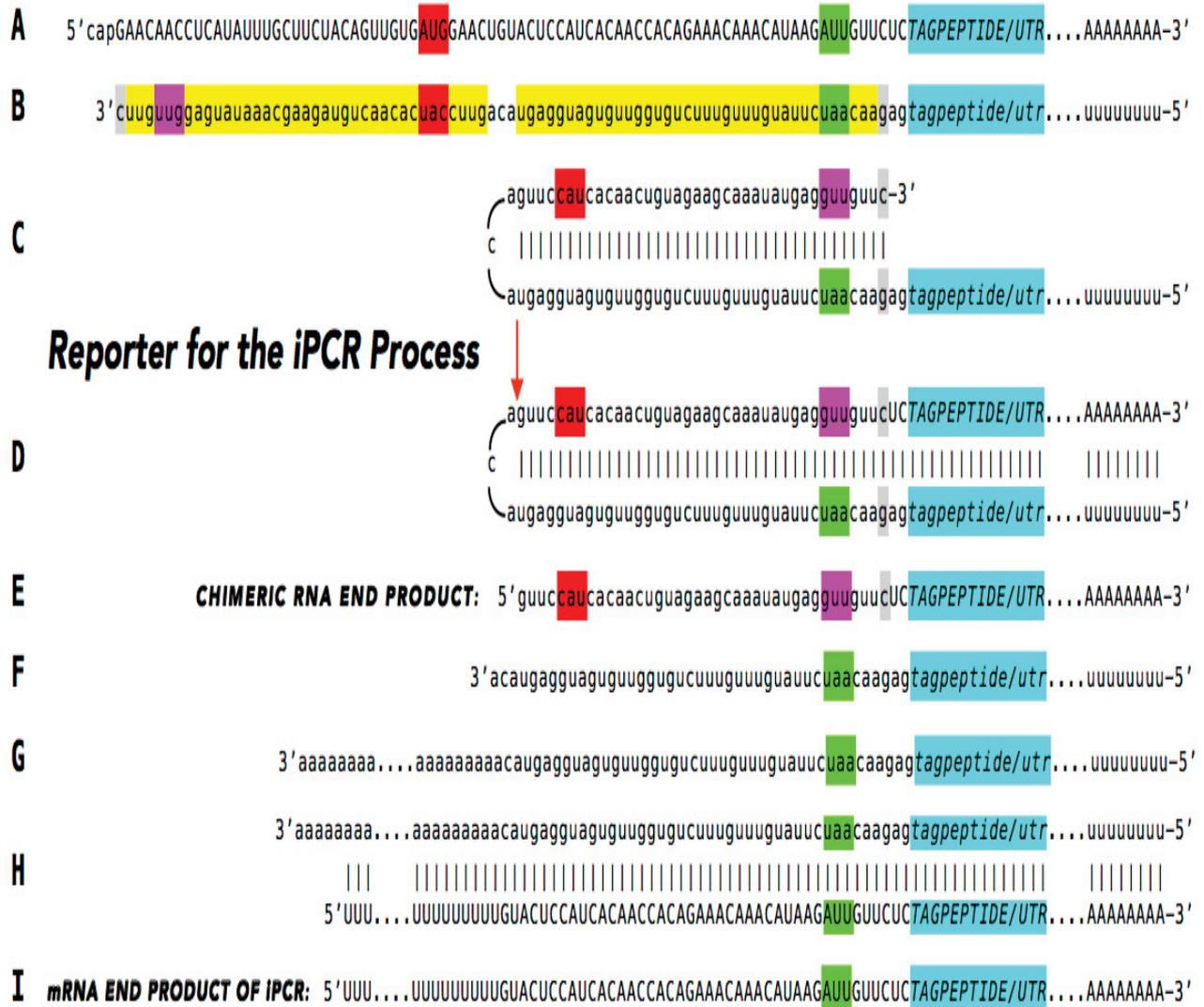
## Reporter Construct for Tier Two of Mammalian mRNA Amplification, The iPCR Pathway

The design of a reporter construct for Tier Two of mammalian mRNA amplification, the iPCR pathway, is significantly more demanding and challenging. Not only should the tag peptide-encoding information in the initial RNA transcript be untranslatable, it should also remain untranslatable in the chimeric RNA end product of Tier One of the amplification process and yet it should be competent of translation into the tag peptide in the mRNA end product of the intracellular PCR pathway. Moreover, in addition to their evident complexity, these demands appear to fly in the face of the principle of Two-Tier Paradox, which posits that regardless of the translational outcome of the chimeric pathway Tier of mRNA amplification, the sense RNA strand produced in the iPCR pathway Tier would always retain the entire coding information of a genome-encoded mRNA and would direct the translation of a polypeptide identical to its conventionally encoded counterpart (described in the "Two-Tier Paradox" section above and further discussed below). Yet, however complex, these difficulties can be circumvented and the above demands satisfied in the following design.

In its initial part, this design exploits a concept utilized in the Type II of a reporter construct for the chimeric pathway of mRNA amplification, which is to generate an amplification-eligible RNA transcript that encodes a reporting polypeptide and contains all necessary elements for its expression but also includes an upstream element preventing translation of the desired polypeptide. This inhibitory upstream element is then removed in the 5'-truncated

sense RNA component of the chimeric RNA end product, and an additional measure, described below, is used to prevent its translation into a tag peptide. As shown in Figure 13, mRNA transcribed from this construct (line A) is eligible for RNA-dependent amplification and encodes a tag peptide (*TAGPEPTIDE/UTR* highlighted in blue denotes nucleotide sequence encoding a tag peptide followed by the 3'UTR). The tag peptide-encoding segment is preceded by the in-frame AUU codon (highlighted in green, as is its complement in the antisense RNA; the choice of the "AUU" rather than "AUG" is explained below) that was shown to retain up to 67% efficiency of the AUG codon when in optimal translation initiation context [138], which is the case in this design. Yet this protein-coding information cannot be translated into a tag peptide because upstream from it is located another AUG (highlighted in red, as is its complement in the antisense RNA) in optimal translation initiation context and out-of-frame with the tag-peptide-encoding segment, and it is this upstream AUG that will initiate translation and prevent production of the tag peptide. In the corresponding antisense RNA (line B), *tagpeptide/utr*, highlighted in blue, denotes antisense complement of the *TAGPEPTIDE/UTR* nucleotide sequence. It contains the 3'-terminal TCE and the internal ICE sequences (both highlighted in yellow). Its folded configuration (line C) can accommodate the additional 3'-terminal "C" of the antisense RNA, not encoded in the DNA but a transcript of the 5'cap "G" of the sense-strand RNA (both 3'-terminal "C" of the antisense RNA and matching "G" at the 5' end of the ICE are highlighted in gray). The 3' terminus of the self-priming structure is then extended, strands are separated and the cleavage occurs (red arrow) at the 5' end of the TCE (line D). The upstream AUG in the initial RNA transcript is located within a segment of mRNA corresponding to the TCE element of the antisense RNA. Consequently, it is removed in the chimeric RNA end product of amplification (line E), which, unless additional measures are taken, would retain the AUU preceding in-frame the tag peptide-coding sequence that could be then translated.

As an "additional measure" to preclude translation of the chimeric RNA end product of Tier One, the 5'guu3' codon (highlighted in pink) is introduced in the antisense RNA in such position that in the folded molecule it matches the 3'uaa5', the complement of the AUU translation initiation codon on the initial RNA transcript. Accordingly, in the resulting chimeric RNA end product of Tier One, the tag peptide-encoding sequence is preceded by the "guu" codon, not known to initiate translation. There is also no functional translation initiation codon upstream from and in-frame with the tag peptide-encoding segment and therefore it remains "silent", i.e. untranslatable. In this construct, the AUU translation initiation codon in the initial transcript was chosen because its complement, the 3'uaa5' (highlighted in green), is also complementary to the 5'guu3', which is used in this design to replace the AUU codon in the chimeric RNA end product of Tier One and thus to translationally incapacitate it. Such complementarity is important because in this design, the TCE and the ICE elements should be fully complementary in order to ascertain that the cleavage of the chimeric intermediate occurs solely at the 5' end of the TCE element. If a TCE/ICE mismatch were allowed, the cleavage of the chimeric intermediate would occur at this mismatch. This would result in a sense/antisense "chimeric junction shift" if the 3-truncated antisense RNA retains a self-priming configuration and is re-used in the extension process (discussed in the "Mechanistic Aspects of the Chimeric Pathway" section above). In such a case, during the extension of the self-primed 3'-truncated antisense RNA, the 3'uaa5' on the antisense RNA (highlighted in



**Figure 13: Design of RNA transcripts from reporter constructs producing a “reporting” polypeptide solely through Tier Two, the iPCR pathway of mRNA amplification.** *Uppercase letters:* Nucleotide sequence of the sense RNA. *Lowercase letters:* Nucleotide sequence of the antisense RNA. *TAGPEPTIDE/UTR*, highlighted in blue, denotes nucleotide sequence encoding a tag peptide followed by 3'UTR; *tagpeptide/utr*, highlighted in blue, denotes complement of the *TAGPEPTIDE/UTR* nucleotide sequence in the antisense RNA. *Highlighted in green:* 5'AUU3' in optimal translation initiation context or its complement on the antisense RNA strand. *Highlighted in red:* 5'AUG3' in optimal translation initiation context or its complement on the antisense RNA strand. *Highlighted in pink:* 5'guu3' on the antisense RNA strand. *Highlighted in yellow:* the 3'-terminal complementary element, TCE, and internal complementary element, ICE, both on the antisense RNA. *Highlighted in gray:* 3'-terminal “c” of the antisense RNA not encoded in the DNA, a transcript of the cap“G” of mRNA, and matching “g” at the 5' end of the ICE on the antisense RNA. *Red arrow:* Position of cleavage of chimeric intermediate following strand separation. **A** – Tag peptide-encoding segment in sense RNA strand is preceded in-frame by the AUU codon, highlighted in green, in optimal translation initiation context and capable of initiating its translation. However, tag peptide cannot be produced since this RNA contains an additional upstream “AUG” codon, highlighted in red (as is its complement in the antisense RNA), also in optimal translation initiation context, but out-of-frame with tag peptide-encoding nucleotide sequence. **B** - Antisense RNA containing the TCE and ICE elements; note that it also contains the 5'guu3' highlighted in pink in its 3'-terminal portion. **C** - Antisense RNA folded into self-priming configuration; note that the 5'guu3' highlighted in pink is accommodated, via complementary interactions, by the 3'uua5' highlighted in green, a complement of the AUU codon on the sense RNA strand (also highlighted in green) and that 3'-terminal “c”, a transcript of the cap“G”, is accommodated by the “g” (both in gray). **D** - Extension of self-primed antisense RNA generates the chimeric intermediate; position of cleavage following strand separation is indicated by red arrow. **E** - Chimeric RNA end product of RNA-dependent mRNA amplification. Note that the upstream AUG has been removed but the “AUU” translation initiation codon preceding in-frame the tag peptide-encoding segment in the initial RNA transcript has been replaced by the “guu” incapable of initiating its translation; there are no functional translation initiation codons upstream from and in-frame with tag-encoding segment, which remains translationally “silent”. **F** - Second end product of the chimeric pathway of mRNA amplification, the 3'-truncated antisense RNA; note that it retains the conventional complement, 3'uua5' (highlighted in green), of the AUU translation initiation codon of the initial RNA transcript. **G** – The 3'-truncated antisense RNA is 3'-polyadenylated in conjunction with the cleavage of the chimeric intermediate; since it also contains the 5'-terminal poly(U) segment, it constitutes the initial template in the iPCR pathway of mRNA amplification. **H** - The sense RNA strand is transcribed from the antisense RNA template initiating within its 3'poly(A) segment. Note the 3'uua5' highlighted in green on the antisense RNA template is transcribed into the AUU codon on the sense RNA strand. **I** – The mRNA product of the iPCR pathway released by strand separation (only one of potentially multiple iPCR cycles is shown). Note that it contains the tag peptide-encoding information preceded in-frame by the AUU codon in optimal translation initiation context and can be translated into a polypeptide “reporting” the occurrence of Tier Two of mammalian mRNA amplification, the intracellular PCR process.

green) would be transcribed into the 5'AUU3' codon preceding in-frame and in optimal translation initiation context the tag peptide-encoding nucleotide sequence which would then become translatable in the chimeric RNA end product.

Meanwhile, whereas a tag peptide-encoding information remains untranslatable in both the initial RNA transcript and in the chimeric RNA end product of Tier One, the second end product of the chimeric pathway of mRNA amplification, the 3'-truncated antisense RNA, retains the conventional complement, 3'uaa5' (highlighted in green), of the AUU translation initiation codon of the initial RNA transcript (line F). If this antisense RNA is 3'-polyadenylated in conjunction with the cleavage of the chimeric intermediate (line G), the mRNA amplification process would progress into Tier Two, the iPCR pathway. In this process, the sense RNA strand would be transcribed from the antisense RNA template initiating within its 3'poly(A) segment (line H); accordingly, the 3'uaa5' on the antisense RNA template would be transcribed into the AUU codon on the sense RNA strand. Separation of strands of such double-stranded structure would release the mRNA end product of Tier Two (line I; note that only one of potentially multiple iPCR cycles is shown in the Figure). It contains the tag peptide-encoding information preceded in-frame by the AUU codon in optimal translation initiation context and can be translated into a polypeptide "reporting" the occurrence of Tier Two of mammalian mRNA amplification, the intracellular PCR process.

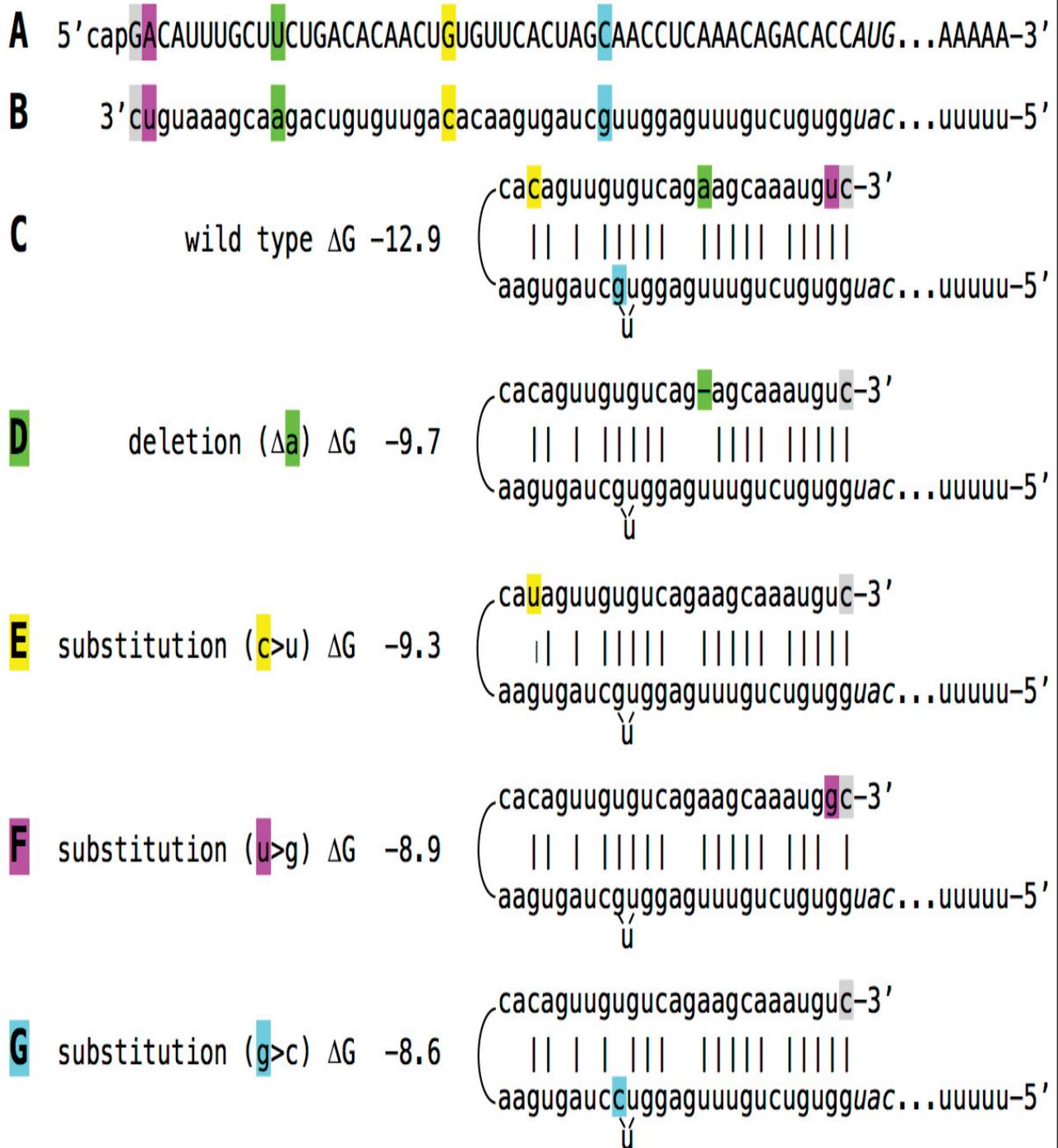
As was mentioned above, the functional reporter construct for the iPCR pathway seemingly violates the principle of the Two-Tier Paradox. However, it does not; the principle remains valid. The Two-Tier Paradox was formulated [8] for the defined system where a 5'-terminal segment of an mRNA corresponding to the 3'-terminal TCE element of the antisense RNA is always smaller than the 5'UTR of the same mRNA molecule. Considering that the average TCE could be 14-18 nucleotides long, as follows from our interpretation of the results described in [16], this assumption appears to be correct for at least the decisive bulk of mammalian mRNA species. The above design of the reporter construct for the Tier Two pathway evades the restraints of the Two-Tier Paradox by having the upstream translation initiation codon positioned within the segment of the construct's RNA transcript corresponding to the TCE of the antisense RNA. In this setting, with regard to the upstream translation initiation codon, the TCE element is larger than the 5'UTR of mRNA; therefore rules of the Two-Tier Paradox do not apply. These rules, on the other hand, remain in effect with regard to the downstream translation initiation codon, the AUU, since for this codon the preceding 5'UTR of the mRNA transcript is larger than the TCE of its antisense RNA complement. Accordingly the translational outcome of the iPCR pathway is identical to that, which would be attained with the initial mRNA transcript if its translation were initiated from the AUU codon, consistent with the Two-Tier Paradox principle; the design, therefore, succeeds in separating regulatory elements into the Two-Tier Paradox rules-relevant and -inapplicable.

## **mRNA Amplification and Disease, Type I. The Deficiency of a Protein Normally Produced via mRNA Amplification: Certain Classes of Beta-Thalassemia**

The potential significance of RNA-dependent mRNA amplification in normal physiology, with every eligible genome-transcribed mRNA molecule constituting a potential mRNA amplification progenitor, can hardly be overstated; in documented examples described above

this process appears to be at the core of extremely high levels of production of globin chains in erythroid differentiation [8,9] and laminin chains in extracellular matrix deposition [10]. Malfunctions of this process may be involved in pathologies associated both with the deficiency of a protein normally produced by this mechanism and with the overproduction of a protein encoded by an mRNA species normally not involved in such a process. One of the experimental approaches to assess the involvement of mRNA amplification in pathologies of the first kind would be to interfere *in vivo* with the extent of complementarity between the two elements, the TCE and the ICE, involved in the antisense RNA self-priming, a core requirement for the progression of the amplification process. In fact, numerous experiments of this sort have been carried out by nature. One of the best opportunities to study the results of such "natural" experimentation is afforded by the occurrence of multiple types of beta-thalassemia. Beta-thalassemia is caused by the reduced or absent synthesis of the beta-globin chains of the hemoglobin tetramer and presents three clinical conditions, the beta-thalassemia carrier state, thalassemia intermedia, and thalassemia major. The beta-thalassemia carrier state results from heterozygosity for the disorder, is clinically asymptomatic and is defined by specific hematological features. Beta-thalassemia major is a severe transfusion-dependent anemia. Beta-thalassemia intermedia comprises a large group of disorders ranging from the mild to severe and heterogeneous at the molecular level. More than two hundreds beta-thalassemia-causing mutations have been identified. The majority of these mutations are single nucleotide substitutions, deletions and insertions. Rarely, there are manipulations of an oligonucleotide, usually leading to a frame shift, even more rare are deletions of the whole gene.

Four known single nucleotide mutations of interest to the subject matter under discussion occur within beta-globin gene segment corresponding to the 5'UTR of the human beta-globin mRNA [82]. They are the following: +1 A>C, +10 -ΔT, +22 G>A, and +33 C>G (numbering from the transcription initiation site). None of these mutations interferes with translation of human beta-globin mRNA when compared with the wild type in cell-free translation assays. Each of these mutations reduces beta-globin production physiologically *in vivo*, causes mild beta-thalassemia in affected patients and appears to solely account for the disease. How can such an outcome be accomplished? Taking into account the established physiologically occurring phenomenon of globin mRNA amplification during erythroid differentiation, it is reasonable to venture that these mutations achieve their effect by interfering with the beta-globin mRNA amplification process; in such a case, the decreased production of beta-globin chains would result even if translation of beta-globin mRNA is completely unaffected. Using this assumption as the point of departure and considering that all four mutations are located within segment of the gene corresponding to the 5'UTR of mRNA, the most likely manner of interference is hindrance with the formation and function of self-priming structure on the beta-globin antisense RNA, a crucial stage in the mRNA amplification process. As diagrammed in Figure 14, this appears to be indeed the case. All four mutations evidently decrease the stability of the self-priming beta-globin antisense RNA structure. Moreover, there is more to it than the change in G. Of the four mutations, three (+10 -ΔT, +22 G>A, and +33 C>G) destabilize the entire double-stranded blocks of nucleotides that they anchor, and one (+1 A>C) weakens the interaction of the 3'-terminal "C" of beta-globin antisense RNA, essential for the commencement of the elongation process, with the complementary nucleotide within



**Figure 14: Beta-thalassemia-causing point mutations in the 5'UTR of human beta-globin mRNA: Projected mechanism of interference with the RNA-dependent mRNA amplification process.** *Uppercase letters:* Nucleotide sequence of the sense RNA; *lowercase letters:* Nucleotide sequences of the antisense RNA. The "G" and the "c" highlighted in grey: Cap nucleotide in mRNA and its complement transcript in the antisense RNA respectively. The 5'AUG3' and the 3'uac5' in italics: Translation initiation codon in mRNA and its complement in the antisense RNA respectively. Highlighted in pink: "A" to "C" substitution in position +1 (counting starts at the TSS) of mRNA and corresponding substitution "u" to "g" in the antisense RNA. Highlighted in green: deletion of "U" in position +10 of mRNA and corresponding deletion of "a" in the antisense RNA. Highlighted in yellow: "G" to "A" substitution in position +22 of mRNA and corresponding substitution "c" to "u" in the antisense RNA. Highlighted in blue: "C" to "G" substitution in position +33 of mRNA and corresponding substitution "g" to "c" in the antisense RNA. Note that all mutations are in the 5'UTR of mRNA and in the corresponding segment of the antisense RNA. Line **A**: Nucleotide sequence of human beta-globin mRNA, wild type. Line **B**: Nucleotide sequence of beta-globin antisense RNA, wild type; note the 3'-terminal "c" highlighted in grey, a transcript of the capG of mRNA. Line **C**: Wild type antisense RNA folding into self-priming configuration; note that the 3'-terminal "c", not encoded in the genome, is accommodated by the complementary "g". Lines **D**, **E**, **F**, and **G**: Folding into self-priming configuration of the antisense RNAs containing one of the mutations described above. Note that each mutation causes destabilization of self-priming structure; this destabilization, in turn, causes decrease of the efficiency of the elongation of the 3' end of the antisense RNA and of the effectiveness of the mRNA amplification process. Less beta-globin mRNA is produced in the mRNA amplification pathway, less beta-globin polypeptide chains are generated, and the disease ensues.

the self-priming structure. The destabilization of the self-priming antisense RNA structure results in decrease in the efficiency of the elongation of the 3' terminus of the antisense RNA in particular and of the mRNA amplification process in general; less beta-globin mRNA is produced, less beta-globin polypeptide chains are generated, and the disease ensues. Thus, the results of the "natural" experiments confirm the predictable effects of the interference with and the malfunction of physiologically occurring mRNA amplification. The four classes of beta-thalassemia described above exemplify, therefore, the Type of diseases due to irregularities in the normally occurring mRNA amplification process.

## mRNA Amplification and Disease, Type II. The mRNA Amplification-Mediated Overproduction of a Protein Normally not Produced via this Pathway: Alzheimer's Disease

On the other hand, Alzheimer's disease exemplifies another Type of disorders, which are due to the occurrence of the mRNA amplification process in the circumstances where it normally does not take place. Moreover, considering the possibility that one of the factors determining the occurrence of (or resistance to) the disease could be the position of the transcription start site, TSS of  $\beta$ APP mRNA, Alzheimer's disease can be regarded as another kind of "natural" experimentation of a type described in Figure 9 above. Because of its immense medical and societal burden, as well as pressing needs to develop effective therapeutic approaches, currently non-existing, this disorder and its apparent molecular underpinnings are discussed below in some detail.

### Interpreting Alzheimer's Disease: Two paradigms

The prevailing cause of dementia, Alzheimer's disease, AD, was, until recently, believed to be initiated and driven by the extracellular accumulation of beta-amyloid,  $A\beta$ , shown to be overproduced in the disease. This peptide, it was assumed, is derived, both in health and disease, solely by two proteolytic cleavages of a large beta-amyloid precursor protein,  $\beta$ APP, which occur in the secretory pathway. First cleavage, by beta-secretase (beta-site amyloid precursor protein cleaving enzyme, BACE) between residues 671 and 672 (numbering according to the  $\beta$ APP770 isoform), generates the C-terminal fragment of  $\beta$ APP (C99, reflecting the number of its amino acid residues) and forms the N-terminus of  $A\beta$ . Subsequent second cleavage of C99 by gamma-secretase (gamma-site  $\beta$ APP cleaving enzyme) forms the C-terminus of  $A\beta$  and completes its production; the resulting beta-amyloid peptide is then secreted into the extracellular space. In AD, according to this view,  $A\beta$  overproduction and associated damages begin early, accumulate slowly throughout the lifespan, and manifest symptomatically late in life in sporadic cases, SAD, or around the fifties in familial cases, FAD; in this paradigm, Alzheimer's disorder is a quintessential "slow" disease. These notions were formalized in the Amyloid Cascade Hypothesis, ACH, which became a prevailing evidence-based theory of AD, and  $A\beta$  emerged as the most compelling therapeutic target of the disease. In recent years, a large number of potential AD drugs, targeting either the generation of beta-amyloid via the  $\beta$ APP proteolytic pathway or the extracellular  $A\beta$  directly, have been developed. They include inhibitors of beta-secretase such as atabecestat, lanabecestat, AZD3839, LY2811376, LY2886721, verubecestat; inhibitors and modulators of gamma-secretase such as avagacestat, begacestat, semagacestat, tarenflurbil; inhibitors of beta-amyloid aggregation such as scillo-inositol, PBT2, tramiprosate;

$A\beta$ -based antigens such as vanutide, AD02, CAD-106, AN-1792; and both polyclonal and monoclonal  $A\beta$ -targeting antibodies such as crenezumab, gantenerumab, solanezumab, bapineuzumab and ponezumab. Many of these potential drugs achieved spectacular successes in treatment, even the reversal, of AD symptoms in animal models. They all, however, exhibited equally spectacular failures in human clinical trials. To explain this discordance, it was suggested that the drugs were administered too late in the course of the disease in human trials. Therefore, the trials, initially involving mild-to-moderate AD patients, were repeated with patients suffering from mild cognitive impairment and biomarker evidence of  $A\beta$  deposition in brain (prodromal AD). The outcomes, however, did not change; moreover, in a number of trials the administration of drugs actually worsened the cognitive or clinical condition of the subjects. This was explained away by suggesting that even the prodromal cases are way too advanced for treatment because  $A\beta$ -triggered damages have been accumulating for decades, presumably starting in the early twenties, and, to be effective, this is when therapeutic intervention should commence and continue the entire life.

Recently, a new theory of Alzheimer's disease has been advanced in studies by Volloch and co-investigators [68-75]. This theory does not dispute the seminal role of  $A\beta$  in triggering and sustaining the disease. It introduces, however, a distinct notion and an entirely different interpretation of the AD-specific source of  $A\beta$  overproduction and its mode of operation in the disease. To understand the discordance between the outcomes seen in animal models and those obtained in human clinical trials, it is instructive to address the apparent exclusivity of the disease to *Homo sapiens*. All non-human mammals tested to date accumulate  $A\beta$  as they age but they do not develop AD. This is true for small-sized, short-lived animals such as mice, and for large-sized, long-lived non-human mammals such as elephants. In contrast, humans do both. It appears, therefore, that the extent of increase in beta-amyloid levels, sufficient to trigger a cascade of molecular events culminating in AD, can be reached physiologically in Alzheimer's disease patients but not in non-human mammals or in healthy humans. This disparity could be easily explained if different pathways of beta-amyloid generation were involved in these instances. It is clear, however, that non-human mammals, healthy humans, and AD patients all share the same  $A\beta$  production pathway, namely the proteolysis of beta-amyloid precursor protein. It can be suggested, therefore, that in AD, in addition to the common beta-amyloid production pathway, another, qualitatively different pathway of  $A\beta$  generation, exclusive to Alzheimer's disease, is in operation, and it is this pathway that drives the disease [72-75]. Two key features of this pathway can be deduced from the results of multiple human clinical trials. First, since BACE inhibitors, highly efficient in treatment of AD symptoms in animal models (where  $A\beta$  was produced solely through  $\beta$ APP proteolysis) and in suppression of  $\beta$ APP proteolysis in both healthy and AD-affected human subjects [83], have no therapeutic effect whatever in AD patients, it appears that the AD-specific  $A\beta$  production pathway is  $\beta$ APP-independent, i.e. that in this pathway beta-amyloid is generated not via the proteolysis of  $\beta$ APP [72-75]. Second, because treatments directly targeting extracellular  $A\beta$ , namely beta-amyloid immunotherapy, decidedly successful in animal models, as well as a significant reduction of extracellular  $A\beta$  levels by BACE inhibitor [83], were also completely ineffective in AD human clinical trials, disease-causing beta-amyloid produced in the AD-specific,  $\beta$ APP-independent pathway appears to be not secreted but retained intraneuronally [74,75]. As described in the following

sections, this pathway appears to be the predominant process of beta-amyloid generation in Alzheimer's disease. Importantly, the disease-causing potential of the AD-specific,  $\beta$ APP-independent pathway of A $\beta$  production is due no less to the intraneuronal retention of its entire beta-amyloid output than to the volume of the resulting A $\beta$ . The above reasoning necessitated a comprehensive reassessment of our interpretation of the disease and led to a conclusion, discussed below, that Alzheimer's disorder is a "fast" disease [73-75] preventable by therapeutic intervention initiated even late in life and treatable at the early symptomatic stages.

### Potentially pivotal role of the AUG encoding Met671 in conventional human $\beta$ APP mRNA in $\beta$ APP-independent generation of A $\beta$ in Alzheimer's Disease

A mechanism of the AD-specific,  $\beta$ APP-independent overproduction of the disease-causing beta-amyloid was indicated by the elucidated chimeric pathway of mammalian RNA-dependent mRNA amplification described in the above sections. It was proposed [68-75] that in AD,  $\beta$ APP mRNA is amplified in the Type II chimeric pathway in such a way that the translational outcome of the amplification process is the C99 fragment, a CTF of  $\beta$ APP and the immediate precursor of beta-amyloid, which is then processed into A $\beta$  by the gamma-secretase cleavage in a non-secretory manner. For this outcome to occur, five distinct requirements must be satisfied. (1) Human  $\beta$ APP mRNA must be eligible for RNA-dependent amplification. (2) The amplification must occur "asymmetrically" in terms defined above. (3) The first competent translation initiation codon in the asymmetrically amplified mRNA must be in-frame with the information content of  $\beta$ APP mRNA. (4) The N-terminus of the resulting polypeptide must be precisely that of A $\beta$ . (5) Last but not the least, the RNA-dependent mRNA amplification process must be activated in the AD-affected neurons.

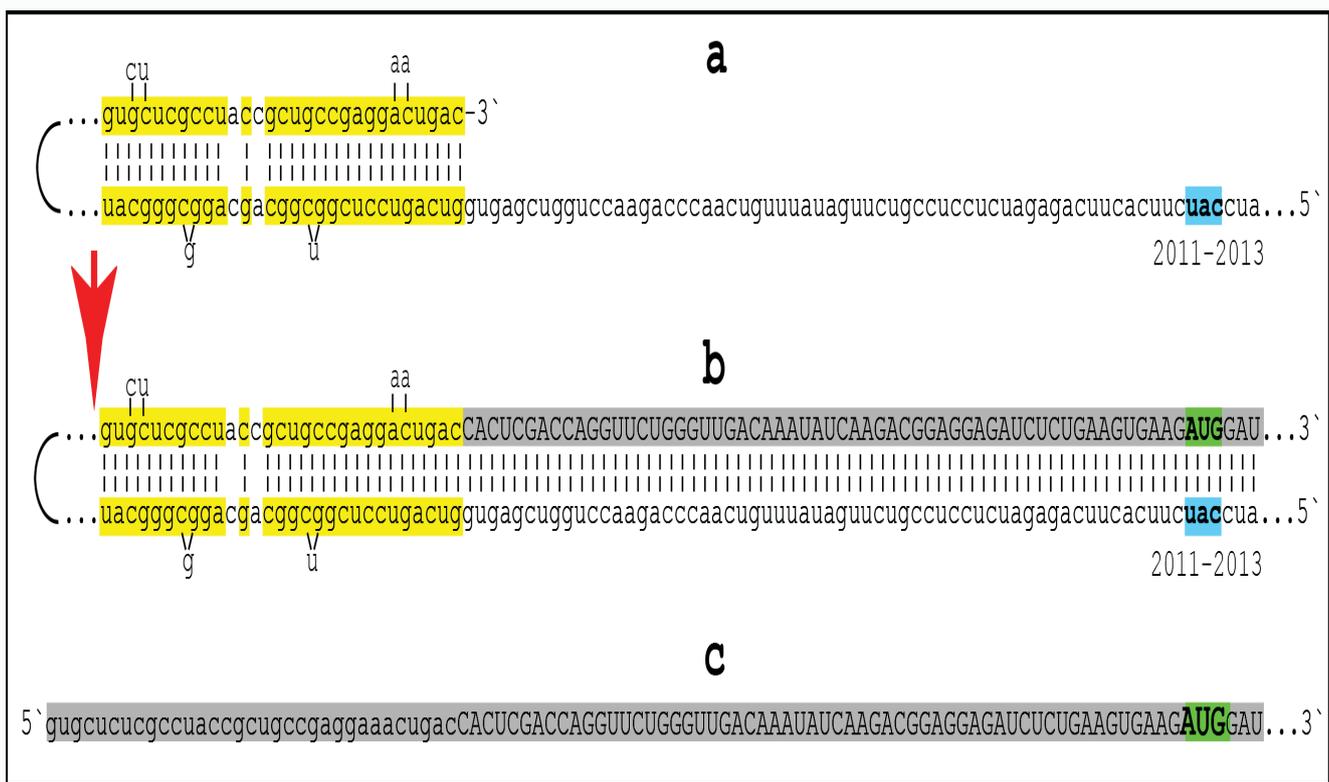
Whether the requirement (4) can potentially be satisfied in AD is relatively easy to evaluate. The answer is affirmative and is suggested by the primary structure of human  $\beta$ APP mRNA. In this molecule, the A $\beta$ -encoding segment is preceded immediately and in-frame by the AUG codon encoding methionine in position 671 of the  $\beta$ APP, the position of the BACE-mediated cleavage (671/672). If translation were initiated at this position, it would produce, after the removal of the N-terminal methionine by the N-terminal methionine aminopeptidase, the C99 fragment, the immediate precursor of A $\beta$ , independently of  $\beta$ APP. Interestingly, the AUG in question is situated within a nucleotide context optimal for the initiation of translation (an "A" in position -3 and a "G" in position +4 relative to the "A" of the AUG codon). In fact, of the twenty AUG codons encoding methionine residues in the human  $\beta$ APP mRNA, only the AUG encoding Met671 (not even the AUG encoding Met1) is located within optimal translation initiation context. Such favorable positioning of the AUG codon in question was the basis for a proposal that in Alzheimer's disease, the C99 fragment of beta-amyloid precursor protein may be generated independently from  $\beta$ APP by the internal initiation of translation at the AUG codon normally encoding Met671 in the intact  $\beta$ APP mRNA [84]. Such precursor-independent generation of C99 would be an efficient way to overproduce A $\beta$ . This is because (a) C99 is not susceptible to the alpha-secretase cleavage, and (b) cleavage by gamma-secretase was shown to be not the rate-limiting step in the production of A $\beta$  [85-87]. The possibility of internal initiation of translation, proposed in [84], has been, however, subsequently ruled out by experiments of Citron and co-investigators [88].

On the other hand, the implications of the occurrence of an AUG codon in such a position for potential  $\beta$ APP-independent generation of A $\beta$  via asymmetric  $\beta$ APP mRNA amplification are apparent: If the RNA-dependent mRNA amplification process is activated in neuronal cells, if human  $\beta$ APP mRNA is eligible for RNA-dependent amplification, if it is amplified in an asymmetric manner in AD, and if, in the resulting 5'-truncated chimeric mRNA, the first, 5'-most, functional translation initiation codon were the AUG encoding Met671 in the conventional  $\beta$ APP mRNA, the translational outcome of the  $\beta$ APP mRNA amplification pathway would be the C99 fragment of amyloid precursor protein produced independently of  $\beta$ APP and containing A $\beta$  at its N-terminus. Is such an outcome feasible?

### Projected pathway of asymmetric amplification of human $\beta$ APP mRNA resulting in chimeric mRNA encoding the C99 fragment of beta-amyloid precursor protein

With regard to the requirements (1), (2), and (3) formulated above, to determine if an mRNA species of interest can potentially be a subject of RNA-dependent mRNA amplification (provided that the cellular RdRs machinery is activated; the putative mechanism of activation of the mRNA amplification process in AD-affected neurons, the requirement (5), is discussed in detail below), one needs to assess whether its antisense complement contains the topologically compatible TCE and ICE elements and is capable of folding into a self-priming configuration. If it is, this will satisfy the requirement (1), and the position of the ICE would define structure of the chimeric RNA end product of the amplification process and indicate whether the requirements (2) and (3) are also satisfied. Such an assessment can be conducted in a model experiment where an mRNA of interest serves as a template for synthesis of cDNA, initiating at the 3'-terminal poly(A), and is subsequently removed by RNase H activity present in a preparation of reverse transcriptase used. If an mRNA is fully transcribed, if complementary elements are present within the antisense strand (cDNA), if one of them is 3'-terminal, and if they are topologically compatible, i.e. mutually accessible within the folded antisense molecule, self-priming and the extension synthesis of a segment of the sense RNA strand would occur. The junction between the antisense and sense components would define the site of self-priming and facilitate identification of the TCE and ICE elements. Just such an experiment was inadvertently carried out with human  $\beta$ APP mRNA [89]. The results of this experiment, misinterpreted and eventually dismissed by the authors as an artifact [90], indicated the occurrence of the topologically compatible TCE and ICE elements within the antisense strand of  $\beta$ APP mRNA and defined their nucleotide sequences as well as the position of self-priming. Based on these results, the TCE/ICE-guided folding and extension of the antisense strand of human  $\beta$ APP mRNA [68-75] can be depicted as shown in Figure 15.

An approximately 30 nucleotide-long 3'-terminal segment of the antisense strand of human  $\beta$ APP mRNA constitutes the TCE. Its counterpart, the ICE, is separated by nearly 2000 nucleotides, yet these elements (both the TCE (top) and the ICE (bottom) are highlighted in yellow in Figure 15) are topologically compatible and the folding of the antisense molecule results in a self-priming configuration capable of accommodating the additional 3'-terminal "C" not encoded in the genome, a transcript of the capG of  $\beta$ APP mRNA (the first nucleotide, "C", in the segment highlighted in grey in Figure 15b). The TCE serves as a primer and is extended; thus generating the sense strand as shown in Figure 15b. Strands are then separated as illustrated in Steps 5 and 6 of Figure 3, and cleavage occurs either at the mismatches within the



**Figure 15: Projected topology of RNA- dependent generation of 5'-truncated mRNA encoding the C99 fragment of human beta-amyloid precursor protein.** Lowercase letters: Nucleotide sequence of the antisense RNA. Uppercase letters: Nucleotide sequence of the sense RNA. Double-stranded portions highlighted in yellow: The TCE (top) and the ICE (bottom) elements of the antisense RNA. Note that the TCE and ICE are separated by about 2000 nucleotides. "2011-2013": Nucleotide positions on the antisense RNA (starting from the complement of the AUG encoding Met1 of the  $\beta$ APP) of the "uac" (highlighted in blue) corresponding to the "AUG" (highlighted in green) encoding Met671 in the  $\beta$ APP mRNA. Panel a: TCE/ICE-guided folding of the  $\beta$ APP antisense RNA. 3'-terminal "c" corresponds to one of multiple transcription start sites of  $\beta$ APP mRNA located 149 nucleotides upstream from its AUG initiation codon; note that such folding configuration would accommodate the additional 3'-terminal "C" (not shown), a transcript of the capG of  $\beta$ APP mRNA. Panel b: Extension of self-primed  $\beta$ APP antisense RNA into sense RNA (highlighted in gray) and cleavage (red arrow; may also occur at one of the TCE/ICE mismatches) of the chimeric intermediate following strand separation. Panel c: Chimeric RNA end product (highlighted in gray) contains 5'-terminal antisense segment extending into  $\beta$ APP mRNA truncated within its coding region. Its translation initiates from the "AUG" (highlighted in green and encoding Met671 in conventional  $\beta$ APP mRNA) immediately preceding the beta amyloid-encoding segment.

TCE or immediately upstream of it as indicated by the arrow in Figure 15b. The resulting chimeric RNA end product, shown in Figure 15c, consists of an antisense segment (the TCE or its portion) continued into a sense-orientated molecule. The translational outcome is decided by the first, 5'-most, initiation-competent translation initiation codon. As shown in Figure 15b and c, the first AUG codon (highlighted in green in Figure 15) is located 58 nucleotides downstream from the antisense portion of the chimeric RNA end product and it is, in fact, the AUG encoding Met671 in the intact  $\beta$ APP mRNA! Translation from this position would produce the C-terminal fragment of  $\beta$ APP containing beta-amyloid at its N-terminus, the C99 fragment of the beta-amyloid precursor protein, in a  $\beta$ APP-independent manner. The major prediction of such a mechanism is a complete inefficiency of beta-secretase inhibition in Alzheimer's disease. This prediction was, in fact, born out in several massive stage III clinical trials [71,83]. Currently, there are no indications that Tier Two, the iPCR pathway, is operational in  $\beta$ APP mRNA amplification in AD. But if it were employed, it could be expected that, as discussed above, alongside the C99 fragment produced in Tier One, the intact  $\beta$ APP polypeptide would be generated, in parallel, as the translational outcome of Tier Two of  $\beta$ APP mRNA amplification. Interestingly, if Tier Two of  $\beta$ APP mRNA amplification, the iPCR process, were to occur, it would have

no decisive effect on the progression of Alzheimer's disease because the bulk of its beta-amyloid output, generated in the  $\beta$ APP proteolytic/secretory pathway, would be secreted into the extracellular pool, in sharp contrast to the apparent intraneuronal retention of the entire A $\beta$  output of Tier One.

#### In humans, only a subset of $\beta$ APP mRNA transcripts is eligible for the RNA-dependent mRNA amplification process: Transcription start site usage may contribute to susceptibility or resistance to AD

As was mentioned above, the presence of a regulatory element known as the "TATA-box" is characteristic for a large class of mammalian genes. Usually, it occurs about 25 nucleotides upstream from the Transcription Start Site (TSS) and rigidly defines its position. The mammalian  $\beta$ APP gene belongs to a class of TATA-less genes that are characterized by multiple transcription start sites. There are at least five, and possibly more, positions where transcription of human  $\beta$ APP mRNA can be initiated [68,91]. Of those, only one, 149 nucleotides upstream from the AUG translation initiation codon, shown in Figure 15, results in an mRNA molecule eligible for the RNA-dependent mRNA amplification process because only for this transcript would the position of the TCE element on its antisense strand be strictly 3'-terminal [68,91], and the additional 3'-terminal "C"; a transcript

of the capG of  $\beta$ APP mRNA [8-10], would be accommodated in the antisense RNA self-priming structure. Utilization of human  $\beta$ APP transcription start sites can be, therefore, one of the factors that define susceptibility or resistance to Alzheimer's disease. The ability to regulate the usage of  $\beta$ APP transcription start site(s), or even to shift it at will, could, therefore, open powerful therapeutic applications.

### **Exclusivity of Alzheimer's Disease to humans: Animal $\beta$ APP mRNA is ineligible for amplification**

The presence of the poly(A) segment at the 3' terminus of an RNA makes the molecule an eligible RdRp template, but it does not necessarily make it eligible for amplification. For this, the antisense RNA strand should be capable of forming a stable self-priming structure in which its 3' terminus can be extended into a segment of the conventional mRNA molecule. Folding of the antisense RNA in such a self-priming configuration requires, in turn, the occurrence of two complementary and topologically compatible elements, one of which is strictly 3'-terminal. It appears that this requirement is not met in antisense RNA complements of  $\beta$ APP mRNA in non-human mammals. In studied animals,  $\beta$ APP antisense RNA segments corresponding to the TCE and ICE elements of human  $\beta$ APP antisense RNA show little, if any, complementarity; moreover, the 3'-terminal segments of various animal antisense  $\beta$ APP RNAs have no extensive complementarity with the rest of the molecule. Thus, in non-human mammals,  $\beta$ APP mRNA, although a suitable RdRp template by virtue of containing the 3'-terminal poly(A), is not eligible for the RNA-dependent mRNA amplification process. Therefore, if, as suggested above, the extent of increase in beta-amyloid levels sufficient to trigger a cascade of molecular events culminating in AD cannot be reached by the  $\beta$ APP proteolytic pathway alone and requires the activation of the  $\beta$ APP mRNA amplification pathway, Alzheimer's disease can occur in humans but not in animals, a conclusion consistent with observations in the field. It should be mentioned that in light of the above considerations, the currently used AD models are inadequate for studies of the disease because they are lacking the  $\beta$ APP mRNA amplification pathway of beta-amyloid production, the major driver of the disease. The reason for this is not the lack of enzymatic apparatus but ineligibility of both, endogenous and transgenic  $\beta$ APP mRNAs utilized in these models, for the RNA-dependent amplification process [72-75], as well as the lack of means for accelerated activation of the enzymatic machinery of the mRNA amplification pathway [74,75]. Therefore, further development of the field requires generation of conceptually new experimental models. The proposed design of such models, as well as suggested methods of their construction, are discussed elsewhere [74,75]. The, arguably, optimal induced pluripotent stem cells (iPSCs)-based class of experimental AD models is described below.

### **Activation of the $\beta$ APP mRNA amplification pathway in Alzheimer's Disease**

Above, we rationalized that in Alzheimer disease, the overproduction of beta-amyloid is effected by an AD-specific A $\beta$  production pathway and, moreover, that this pathway is independent from the  $\beta$ APP proteolytic processing and that the entire output of this pathway is retained intraneuronally. The  $\beta$ APP mRNA amplification pathway of beta-amyloid generation discussed above is evidently independent from  $\beta$ APP proteolysis. The remaining outstanding questions are: How this pathway is activated and maintained in the disease and why the resulting beta-amyloid peptide, unlike its counterpart produced in the  $\beta$ APP proteolytic pathway, is retained within neuronal cells rather than being secreted?

Recently, it was suggested [72-75] that in Alzheimer's disease, the  $\beta$ APP mRNA amplification pathway in neuronal cells is triggered by mitochondrial dysfunction. Mitochondrial dysfunction in neurons is, in turn, caused by A $\beta$  accumulated inside the cell to sufficient levels [72-75,92-108]. Whereas the scale of the postulated intraneuronal retention of beta-amyloid produced in the  $\beta$ APP mRNA amplification pathway is unparalleled in that it applies to the entire output of the amplification process, a notion of intracellular  $\beta$ APP-derived A $\beta$ , albeit as a fraction of its secreted counterpart, is not unprecedented. There are two potential sources of intracellular A $\beta$  produced in the  $\beta$ APP proteolytic/secretory pathway. The more studied and understood source is cellular uptake of secreted extracellular beta-amyloid. Data obtained conclusively show that soluble extracellular A $\beta$ 42 and A $\beta$ 40 use endocytosis [109] to enter the cell and that A $\beta$ 42 is taken up two times more efficiently than A $\beta$ 40 [110]. Beta-sheet-rich A $\beta$ 42 aggregates were observed to enter cells at low nanomolar concentrations [111]. In contrast, monomers were shown to bound to plasma membrane and to form aggregates there before cellular uptake and accumulation in endocytic vesicles [112], thus indicating that formation of A $\beta$  aggregates may be a prerequisite for cellular uptake [109,111-113]. Moreover, it was suggested that oligomer-specific A $\beta$  toxicity in cell models is mediated by its selective uptake [109]. Cellular uptake of A $\beta$  was also shown to be ApoE isoform-dependent and mediated by lipoprotein receptor LR11/SorLA [112]. ApoE4, a major genetic risk factor for AD, was shown to be much more efficient in mediating A $\beta$  uptake than ApoE3 and ApoE2 [112]. LRP, another member of the lipoprotein receptor family, binds to A $\beta$  directly or through ligands such as ApoE and undergoes endocytosis, thus facilitating cellular uptake of A $\beta$  [114]. The internalization of extracellular A $\beta$  can also be mediated by  $\alpha$ 7 nicotinic acetylcholine receptor [115-117], the scavenger receptor for advanced glycation, RAGE [118-120], the formyl peptide receptor-like 1, FPRL1 [121], and N-methyl-d-aspartate, NMDA, receptors [122]. A $\beta$  internalization was observed in multiple cell types, including neurons, astrocytes, glial cells and macrophages, where it appears to play different roles, and it occurs in cells of normal subjects as well as in cells of AD-affected individuals [123]. Because of its occurrence in healthy humans, this source of intracellular beta-amyloid appears to be insufficient alone to trigger mitochondrial dysfunction,  $\beta$ APP mRNA amplification, and, consequently, AD.

The other potential source of intracellular  $\beta$ APP-derived A $\beta$ , operating only in beta-amyloid producing cells, i.e. neurons in the subject under discussion, is its retention within the cell, apparently essential for reaching critical levels sufficient to cause mitochondrial distress and trigger the disease. Whether A $\beta$  is retained intracellularly or is secreted into the extracellular pool is defined by the location at which the immediate beta-amyloid precursor, the C99 fragment, is cleaved by the gamma-secretase complex. The vast majority of A $\beta$  produced in the  $\beta$ APP proteolytic pathway is generated by cleavage at the plasma membrane and is secreted. However, gamma cleavage can also occur in the Endoplasmic Reticulum (ER) [124], Golgi and Trans Golgi Network (TGN) [125], and at endosomal [124], lysosomal [124] and mitochondrial [126] membranes; such cleavages generate intracellularly retained A $\beta$ . It has been shown that different isoforms of intracellular A $\beta$  can be generated at different locations. For example, cleavage within the ER produces predominantly A $\beta$ 42 [127-131] whereas cleavage within the TGN mostly generates A $\beta$ 40 [132]. Interestingly, these locations of intracellular A $\beta$  generation are limited to neurons [128]. Recent evidence suggests that subcellular localization

of Presenilin 2 (PSEN2) directs the assembly of gamma-secretase complex to specific cellular compartments and thus contributes to the balance between intracellular accumulation and secretion of A $\beta$  [110,133]. Moreover, FAD-associated PSENs mutations were shown to pronouncedly augment the intracellular pool of A $\beta$  by determining localization and substrate specificity of gamma-secretase [133].

Whereas only a fraction of A $\beta$  produced in the  $\beta$ APP proteolytic pathway is retained intraneuronally, the entire output of the predominant  $\beta$ APP mRNA amplification pathway of beta-amyloid production presumably remains inside the AD-affected neurons (until they die as a result and release it into the extracellular pool, that is). Normally, proteolytic processing of the bulk of  $\beta$ APP occurs in the secretory pathway and culminates in gamma-secretase cleavage within the plasma membrane and subsequent secretion of newly generated A $\beta$  into the extracellular space. A nascent  $\beta$ APP molecule is chaperoned into the secretory pathway through its N-terminal signal peptide. There is, however, no N-terminal signal peptide in the translational end product of asymmetric amplification of human  $\beta$ APP mRNA. Indeed, the  $\beta$ APP mRNA amplification pathway results in the C99 fragment. C99 lacks N-terminal signal peptide and requires only gamma-secretase cleavage to produce A $\beta$ . This cleavage can occur on intracellular membranes at a variety of sites, such as ER and TGN [125-130], where gamma-secretase cleavage was shown to occur only in neurons [128], and even in mitochondria where all subunits of the gamma-secretase complex were shown to be present [132,134,135]. If this cleavage occurs in the ER, the A $\beta$ 42 isoform will be predominantly produced [128]. If both ER and TGN sites are utilized, a mixture of A $\beta$ 42/A $\beta$ 40 will result [128]. For A $\beta$  to be retained intraneuronally, its precursor, the C99 fragment, produced either in the proteolytic pathway or in the AD-specific  $\beta$ APP mRNA amplification pathway, has to be processed, i.e. cleaved by gamma-secretase, on an intracellular membrane. In the  $\beta$ APP proteolytic pathway this occurs infrequently, whereas in the chimeric  $\beta$ APP mRNA amplification pathway of A $\beta$  production, this is, apparently, the only way. In addition, if the heavily modified mRNA end product of RNA-dependent asymmetric human  $\beta$ APP mRNA amplification, encoding the C99 fragment, is translated in a compartmentalized manner, this may also contribute to selection of gamma-secretase cleavage sites on intracellular rather than plasma membranes. If the chimeric  $\beta$ APP mRNA amplification pathway were to progress into Tier Two of the amplification process, the impact of A $\beta$  produced in the iPCR pathway on the progression of the disease would be rather minor. This is because the iPCR pathway would produce the intact  $\beta$ APP; following its proteolytic processing, unless potentially compartmentalized translation of modified  $\beta$ APP mRNA leads to retention of the  $\beta$ APP-derived beta-amyloid, the bulk of the resulting A $\beta$  would be secreted and thus “disabled” in terms of its capacity to cause mitochondrial dysfunction.

As for how the intraneuronal A $\beta$ -initiated mitochondrial dysfunction activates the RNA-dependent mRNA amplification pathway, the missing links were provided in two recent studies [66,67]. In these investigations, it was shown that depolarization, a change in the electrical charge which occurs during mitochondrial dysfunction, activates the protease OMA1, which is located on the inner of the two mitochondrial membranes surrounding the organelle. In turn, the activated OMA1 cleaves or facilitates the cleavage of another mitochondrial protein, DELE1, which resides in the space between the two mitochondrial membranes and is associated with the inner membrane, the locality of OMA1. Following the cleavage, a fragment

of DELE1 is released to the cytosol where it binds to HRI and activates it. Activated HRI phosphorylates eIF2 $\alpha$ ; this leads to the integrated stress response. The ISR, as described above, suppresses global cellular protein synthesis but activates expression of selected transcription factors, notably ATF4, ATF5, and CHOP [66,67]. Some of these transcription factors either act as inducible components of the RdRp complex or enable their expression. Once the competent RdRp complex is assembled, the RNA-dependent mRNA amplification pathway is activated in neuronal cells (Figure 7). Importantly, as reasoned above, by the intrinsic logic of the process of ISR-activated mRNA amplification, the amplified mRNA has to be compatible with phosphorylated eIF2 $\alpha$  and freely translated under the ISR conditions, possibly due to its nucleotide modifications. This implies that the C99 fragment generated in the  $\beta$ APP mRNA amplification pathway is produced via preferential translation from its amplified and thus modified mRNA template and, consequently, that intraneuronally retained A $\beta$ , resulting from its processing, is further enriched (relative to the bulk of cellular proteins) in AD-affected neurons under mitochondrial distress and the ISR conditions, thus augmenting the efficiency of the mRNA amplification process.

### Activation of the $\beta$ APP mRNA amplification pathway in Alzheimer's disease: A Case Study

As an illustration of the concepts of activation of the  $\beta$ APP mRNA amplification pathway discussed above, it is instructive to briefly review the case of the Swedish  $\beta$ APP mutation. Swedish mutation is the first-discovered  $\beta$ APP mutation resulting in early onset Alzheimer's disease (familial AD). It affects amino acid residues 670 (normally lysine) and 671 (normally methionine) of  $\beta$ APP and replaces them with asparagine and leucine respectively [145]. These two residues are immediately adjacent to the site of beta-secretase cleavage, which occurs between positions 671 and 672 in both wild-type and the Swedish  $\beta$ APP mutant. Therefore, it was initially assumed that the Swedish mutation causes the disease by increasing the efficiency of  $\beta$ APP cleavage by beta-secretase and thus elevating the production of A $\beta$  [145,146]. This is, however, not the case; the efficiency of beta-cleavage of the Swedish  $\beta$ APP mutant is similar to that of wild-type  $\beta$ APP. And yet, the Swedish  $\beta$ APP mutation reliably and uniformly triggers early onset Alzheimer's disease in its carriers. The proffered interpretation, that in this case the beta-cleavage outcompetes the alpha-cleavage of  $\beta$ APP and thus increases the production of A $\beta$  [146], appears insufficient and necessitates an alternative explanation.

In the framework of the  $\beta$ APP mRNA amplification theory of Alzheimer's disease, where the emergence of symptomatic AD is equated with the activation of the  $\beta$ APP mRNA amplification pathway, it can be assumed that the occurrence of the Swedish mutation substantially accelerates the commencement of A $\beta$  generation in the mRNA amplification pathway. In this respect, it could be argued that the operation of the  $\beta$ APP mRNA amplification pathway of A $\beta$  production is inconsistent with the Swedish mutation because the mutant  $\beta$ APP mRNA no longer encodes Met671, postulated to initiate translation of the amplified  $\beta$ APP mRNA. This argument, however, would not be valid for the following reason. In wild-type  $\beta$ APP mRNA, the nucleotide sequence encoding residues 670-672 is: AAG AUG GAU. In the Swedish mutant this sequence is changed to: AAU CUG GAU. The CUG triplet, which replaces the AUG in the mutant  $\beta$ APP mRNA, was shown to initiate translation with 82% efficiency of the AUG when it is situated in the optimal translation initiation context [138]. It is in the Swedish  $\beta$ APP mRNA mutant (the A in

position -3 and the G in position +4 relative to the C of the CUG), and thus would be operational in the initiation of translation.

Therefore, if and when activated, the generation of A $\beta$  via the  $\beta$ APP mRNA amplification pathway would be (almost) perfectly efficient in cells carrying the Swedish mutation. The question is how the occurrence of this mutation accelerates activation of the  $\beta$ APP mRNA amplification pathway? The generic answer is: Location, location, location. In this particular case, it is the location of  $\beta$ APP cleavage by beta-secretase. Whereas in wild-type it occurs in the endosomal compartments [147] and leads, eventually, to secretion of A $\beta$ , in Swedish mutation carriers it takes place predominantly in the Golgi complex [145,146], where gamma-secretase was also shown to be present and gamma-cleavage to occur [125]. It follows that in Swedish mutation carriers, the bulk of A $\beta$  produced by  $\beta$ APP proteolysis would be retained intracellularly. Initially, however, it was speculated [146] that the processing of  $\beta$ APP containing the Swedish mutation occurs not in the Golgi complex itself but in Golgi-associated secretory vehicles, and, therefore, it culminates in secretion of the resulting A $\beta$ . Subsequent investigations indicated that this is not the case. Indeed, the notion of intracellular retention of beta-amyloid generated by the proteolytic processing of  $\beta$ APP containing the Swedish mutation was validated experimentally. It was shown [148] that  $\beta$ APP containing the Swedish mutation is processed in a pathway which is distinctly different from the conventional  $\beta$ APP secretory proteolytic process (that occurs in a post-Golgi compartment). Not only is it processed in the Golgi complex itself, but the processing also results in the intracellularly retained A $\beta$  and, consequently, in its rapid accumulation within the cell [148]. Thus, what is accelerated (or, rather, substantially shortened) in carriers of the Swedish  $\beta$ APP mutant is the duration of A $\beta$  accumulation to critical intraneuronal levels sufficient to trigger mitochondrial dysfunction, to initiate the ISR and to activate the  $\beta$ APP mRNA amplification pathway and elicit symptomatic manifestation of the disease. Consequently, in Swedish mutation carriers, symptomatic AD uniformly ensues much sooner (early onset) than in sporadic late onset cases seen in wild-type  $\beta$ APP carriers where it takes most of the lifetime to build up sufficiently high levels of intracellular A $\beta$ .

### The engine that drives Alzheimer's Disease

To summarize the preceding section, the intraneuronal accumulation of A $\beta$ , conventionally produced in the  $\beta$ APP proteolytic pathway to levels sufficient to trigger mitochondrial dysfunction and, consequently, activation of the  $\beta$ APP mRNA amplification pathway, plays the role of a "starter motor" in getting car engine moving in a self-sustainable manner. In this context, importantly, the rate of intraneuronal retention of  $\beta$ APP-derived beta-amyloid, combined with that of its uptake, apparently plays the decisive role in defining the susceptibility to AD. The "starter motor", once switched on, activates the engine that drives Alzheimer's disease. Indeed, mitochondrial distress leads to activation of the mitochondrial protease OMA1 followed by the cleavage, still within the mitochondria, of DELE1, release of one of DELE1 fragments into the cytosol, its binding to and activation of the HRI kinase, consequent phosphorylation of eIF2 $\alpha$ , and commencement of the integrated stress response. Within the framework of the ISR, it also leads to the activation of the RNA-dependent mRNA amplification machinery and, provided  $\beta$ APP mRNA is eligible, of the predominant, AD-specific, pathway of A $\beta$  generation, the  $\beta$ APP mRNA amplification process [72-75]. The entire output of this pathway is retained intraneuronally. Drastically

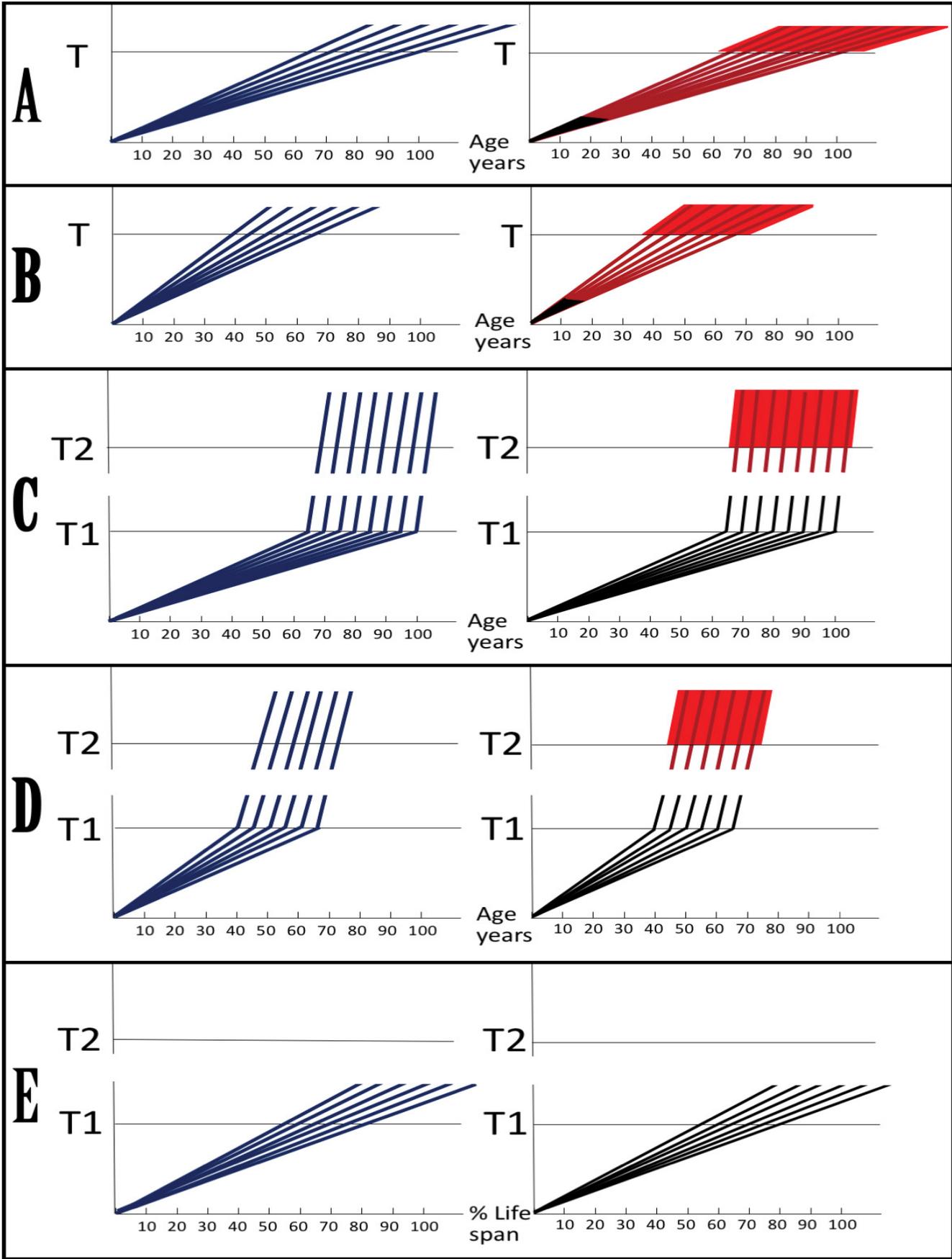
increased levels of intracellular A $\beta$  promote, in turn, further mitochondrial dysfunction, the OMA1 to DELE1 to HRI signaling pathway is maintained, and the integrated stress response and, consequently, the  $\beta$ APP mRNA amplification process, are sustained, thus generating self-perpetuating beta-amyloid overproduction/mitochondrial dysfunction mutual feedback cycles. These relationships are presented diagrammatically in Figure 16. It depicts the mutual feedback cycles as a two-stroke engine, the engine that drives beta-amyloid overproduction and, consequently, Alzheimer's disease; ultimately, it triggers neuronal death. To develop sporadic AD, it takes, apparently, a lifetime of conventionally produced intraneuronal A $\beta$  accumulation to critical levels sufficient to trigger mitochondrial dysfunction, the integrated stress response and, consequently, the RNA-dependent mRNA amplification pathway, combined with certain susceptibility factors (e.g. appropriate TSS usage permitting the operation of  $\beta$ APP mRNA amplification). In familial AD, because of abnormal  $\beta$ APP proteolysis and/or other factors, critical intraneuronal levels of conventionally produced A $\beta$ , or of its more "toxic" isoforms with increased capacity to aggregate and to trigger mitochondrial dysfunction, are reached sooner and the disease occurs earlier in life.

### The new paradigm: Alzheimer's Disorder is a "Fast" disease preventable by therapeutic intervention initiated even late in life, and treatable at the early symptomatic stages

The new theory of AD posits that Alzheimer's disorder is a "fast" disease [74,75]. This is in sharp contrast to a view, prevailing until recently, that AD is a quintessential "slow" disease, which develops throughout the life as one prolonged process. Dynamics of the disease in these two paradigms is presented in Figure 17. Its interpretation in the old paradigm is shown in panels A and B. This dynamics is single-phased and can be divided into asymptomatic (red lines) and symptomatic (red blocks) portions. A $\beta$ , it is assumed, is overproduced (and secreted) solely in the  $\beta$ APP proteolytic/secretory pathway. As its extracellular levels increase, it triggers neurodegeneration (red lines) starting early in life. Damages accumulate and manifest symptomatically (red blocks) late in life in sporadic cases (panel A). In familial AD cases, where mutations in the  $\beta$ APP gene or in presenilins increase production of either common A $\beta$  isoform or of its more toxic isoforms, neurodegeneration reaches critical threshold sooner and AD symptoms occur earlier in life, mostly in the late 40s and 50s (panel B). In this paradigm, the disease is considered untreatable in the symptomatic phase and there are currently no preventive AD therapies, but if they were available, according to this viewpoint, it would be largely futile to intervene late in life in SAD case or at mid-age in case of FAD because, although AD symptoms have not yet manifested, the irreversible damage has already occurred during the preceding decades. In this paradigm, to be effective, preventive therapeutic intervention should be initiated early and continued for life.

Dynamics of the disease in the new paradigm, illustrated in panels C (SAD) and D (FAD) of Figure 17, is radically different [73-75]. This dynamic is biphasic. In the first phase, only the  $\beta$ APP proteolytic pathway of A $\beta$  production is in operation. This phase is a slow process of intraneuronal beta-amyloid accumulation. It occurs via cellular uptake of secreted A $\beta$  and the intracellular retention of a fraction of  $\beta$ APP-derived A $\beta$ . These processes are common to *Homo sapiens*, including healthy humans, and to non-human mammals, and result neither in noticeable damage (black lines), nor in any manifestation of





**Figure 17: Dynamics of Alzheimer's disease in two paradigms.** *Left panels:* Dynamics of A $\beta$  production; *Right panels:* Dynamics of neurodegeneration. *Blue lines:* Levels of beta-amyloid; *Red lines:* Extent of neurodegeneration; *Black lines:* Indicator lines, no noticeable neurodegeneration; *Red blocks:* Symptomatic manifestation of AD. **T:** Threshold of symptomatic manifestation of AD (reflects levels of extracellular A $\beta$  and assumed consequent extent of neurodegeneration); **T1:** Threshold of activation of RNA-dependent  $\beta$ APP mRNA amplification (reflects intraneuronal A $\beta$  levels and the consequent extent of A $\beta$ -coupled mitochondrial dysfunction); Numerous genetic factors such as the occurrence of various alleles of ApoE gene, as well as certain epigenetic factors, influence the age when thresholds **T** and **T1** are reached, hence, the fanning lines. **T2:** Threshold of symptomatic occurrence of AD (reflects levels of intraneuronal A $\beta$ , degree of mitochondrial dysfunction and consequent extent of neurodegeneration). *Panels A, B: View of the dynamics of AD in the old paradigm (A: Dynamics of SAD; B: Dynamics of FAD).* Levels of extracellular A $\beta$  increase, neurodegeneration starts early and accumulates throughout the life. When threshold **T** is reached, AD symptoms manifest. *Panels C, D: The outlook on the dynamics of AD in the new paradigm (C: Dynamics of SAD; D: Dynamics of FAD).* Levels of intraneuronal A $\beta$  increase, the extent of mitochondrial dysfunction reaches threshold **T1**, the integrated stress response is triggered, and RNA-dependent  $\beta$ APP mRNA amplification is activated. There is no noticeable neurodegeneration until after a *lag period* (when amplified RNA encoding the C99 fragment of  $\beta$ APP accrues and intraneuronal A $\beta$  further accumulates) following the crossing of **T1** threshold and activation of  $\beta$ APP mRNA amplification; when the extent of neurodegeneration reaches threshold **T2**, AD symptoms manifest. *Panel E: Dynamics of A $\beta$  production and neurodegeneration in non-human mammals and inhumans not susceptible to AD or successfully treated:* **T1** threshold is crossed but  $\beta$ APP mRNA is not amplified in animals because it is not eligible for RNA-dependent amplification process. There is no noticeable neurodegeneration; **T2** threshold is not reached, no AD symptoms manifest, no disease occurs. *Note:* Scenario depicted in panel (E) would occur in humans not susceptible to Alzheimer's disease due to variations in  $\beta$ APP TSSs utilization or for other reasons, or when the  $\beta$ APP mRNA amplification pathway of A $\beta$  production is effectively suppressed by therapeutic intervention.

mitochondrial defects and normalized lifespan of GRU102 that otherwise was substantially reduced, thus establishing metabolic mitochondrial dysfunction as a promising and viable therapeutic intervention target in Alzheimer's disease. Moreover, there are good reasons to believe that with a drug blocking the  $\beta$ APP-independent A $\beta$  production pathway, asymmetric  $\beta$ APP mRNA amplification, or with related therapeutic approaches detailed in [75], it would be possible not only to preempt the disease but also to stop and reverse it even when early AD symptoms have already manifested. This assumption is based on observations made in studies where AD symptoms, as well as neurodegeneration, were reversed by suppressing, via BACE inhibition, overproduction of beta-amyloid in mouse AD models generating A $\beta$  solely in the  $\beta$ APP proteolytic pathway [136,137]. This paradigm of Alzheimer disease establishes the feasibility of prevention of and interference with the disease and suggests an array of novel therapeutic targets [74,75]. Most importantly, after years of despair in the field, it makes a case for optimism [75].

### Induced pluripotent stem cells-based optimal experimental AD models

The feasibility of a reporter for the chimeric pathway of mammalian RNA-dependent mRNA amplification, described above, suggests a blueprint for the construction of, arguably, the optimal cell-based model of Alzheimer's disease. In this design, human iPSCs are stably transfected with a DNA construct of a type detailed in the top panel of Figure 12. It encodes a tag peptide that can be produced only if construct-transcribed RNA is amplified in the RNA-dependent mRNA amplification pathway and the AUG codon, capable of initiating translation of a tag peptide, is contributed to the chimeric RNA end product of amplification by its antisense RNA component; otherwise RNA transcripts from this construct are translationally "silent". The occurrence of the tag peptide, therefore, reports the activity of the RNA-dependent mRNA amplification pathway. To imitate processes occurring in AD, the resulting modified iPSCs are differentiated into neurons and the neuronal cultures are transiently transfected with amplification-ineligible RNA (efficient DNA transfection requires cell division and would be problematic in neurons) encoding human A $\beta$ 42 that would be retained within neuronal cells. If necessary, cells may be transfected multiple times at few days intervals; A $\beta$ 42 may also be delivered into the neurons by other feasible means. Intracellular accumulation of A $\beta$ 42, translated from transiently transfected RNA or delivered by other means, to sufficient levels would trigger mitochondrial distress. This will initiate a chain of events, namely activation of OMA1 leading to cleavage of DELE1 followed by activation of the HRI kinase, phosphorylation of eIF2 $\alpha$ , and elicitation of the integrated stress response, which

ultimately would result in the activation of the RNA-dependent mRNA amplification pathway and consequent production of an easily detectable reporter peptide.

With the mRNA amplification machinery activated and simultaneously with the expression of a reporter peptide, the eligibility of human  $\beta$ APP mRNA for mRNA amplification would enable the Type II asymmetric chimeric pathway described above and verifiable by the occurrence of rigorously predictable and uniquely specific antisense/sense  $\beta$ APP RNA junction sequences.  $\beta$ APP mRNA amplification would result in the generation of endogenous C99 produced independently of  $\beta$ APP and processed, in a neuron-specific manner, into intracellularly retained beta-amyloid peptides possibly enriched in the A $\beta$ 42 isoform. This, in turn, would lead to drastically increased levels of intracellular A $\beta$ , which would promote, without additional transient transfections or other exogenous A $\beta$  deliveries, further mitochondrial dysfunction. Consequently, the OMA1 to DELE1 to HRI signaling pathway would be maintained, the integrated stress response sustained, and the  $\beta$ APP mRNA amplification continued; i.e. the "engine" that drives AD would be operational. This would further propagate the production of endogenous A $\beta$  via  $\beta$ APP mRNA amplification pathway and potentially lead to manifestation of AD features, such as neurofibrillary tangles, previously observed in neuronal cultures overproducing A $\beta$  [140], that can be evaluated in assessing the efficiency of candidate drugs. The employment of iPSCs derived from Alzheimer's patients would strongly increase the probability that the TSS (-149) of the  $\beta$ APP gene is utilized and, therefore, that endogenous  $\beta$ APP mRNA is amplification-eligible. In such system, the  $\beta$ APP mRNA amplification-mediated production of A $\beta$  is expected to correlate with that of a reporter peptide. If this correlation is substantiated, such a system would meet the acute need for an adequate and easily monitored experimental AD model and advance the development of therapeutic approaches, currently non-existent. Any treatment affecting the  $\beta$ APP mRNA amplification-mediated production of AD-causing intracellularly retained A $\beta$  would equally affect the production of a tag peptide, which would, therefore serve as a "reporter" in more ways than one.

In an alternative design variant for experimental AD model a tag peptide-encoding sequence in a construct otherwise conceptually identical to that described above is replaced with the sequence encoding human A $\beta$ 42 or another desired isoform of beta-amyloid tagged by one of the FAD-associated mutations. This construct is then used to stably transfect human iPSCs; A $\beta$  can be produced from it only if the mRNA amplification pathway is activated. The "tagging" is to distinguish, upon modified iPSCs' differentiation into neurons

and activation of the RNA-dependent mRNA amplification pathway by A $\beta$ 42-encoding RNA transfections, the resulting construct-derived beta-amyloid from that produced endogenously and from A $\beta$  molecules translated from transiently transfected RNA. In this variant, construct-derived, mutation-tagged A $\beta$  itself serves as a reporter of the activity of the RNA-dependent mRNA amplification pathway, in addition to its functional role. Once the mRNA amplification process is activated, expression from the construct would result in intraneuronally retained beta-amyloid that will sustain the activity of self-perpetuating A $\beta$  overproduction/mitochondrial dysfunction mutual feedback cycles as well as that of the RdRp complex. It will also trigger manifestation of the AD features even in the absence of endogenous  $\beta$ APP mRNA amplification due, for example, to the TSS utilization pattern. It should be emphasized that whereas all animal AD models are intrinsically compromised to a certain degree because they express, non-physiologically, beta-amyloid in multiple cell types and at various developmental stages, human iPSCs-originated, cell-based experimental models such as described above are, arguably, the optimal ones because, unlike other experimental systems, they allow differentiation of stem cells into neurons, unimpeded and unaffected by non-physiological production of beta-amyloid or any other protein in undifferentiated precursor cells.

**Alzheimer's "Holy Grail": Beta-amyloid generated in the  $\beta$ APP mRNA amplification pathway is naturally "tagged" and can be distinguished directly in AD-affected brain from A $\beta$  produced conventionally by  $\beta$ APP proteolysis**

In fact, the AD models discussed in the preceding section could be significantly simplified, and experiments considered "unthinkable" previously can be performed since, apparently, nature itself "tags" A $\beta$  peptide generated in the  $\beta$ APP mRNA amplification pathway.

In the  $\beta$ APP proteolytic pathway the beta-cleavage, which generates the N-terminus of A $\beta$ , occurs between methionine (Met) in position 671 and aspartate (Asp) in position 672 of  $\beta$ APP. Accordingly, the Asp672 constitutes the N-end of the C99 fragment of  $\beta$ APP and, after gamma-cleavage of the latter, of A $\beta$ . In the human  $\beta$ APP mRNA amplification pathway of A $\beta$  generation, translation initiates with the Met671. Above, it was suggested that this methionine is removed by N-terminal methionine aminopeptidase. In such a case, the N-terminal amino acid residue of the resulting C99 fragment, in this situation a primary translation product, and of A $\beta$  derived from it by the gamma-cleavage, would also be the Asp672. Consequently, A $\beta$  generated in the  $\beta$ APP mRNA amplification pathway would contain no origin-specific identifying feature, and thus be identical in every aspect to conventional  $\beta$ APP-derived beta-amyloid peptide and indistinguishable from it.

However, apparently, this is not the case. As was pointed out by A. Varshavsky (Caltech) in a personal communication, the N-terminal methionine in a translation product IS NOT REMOVED by N-terminal methionine amino peptidase (and, moreover, the retained N-terminal Met could possibly be acetylated) if it is followed by aspartate or other residues that are larger than valine [141-143]. This is precisely the scenario whereby translation of chimeric mRNA resulting from human  $\beta$ APP mRNA amplification initiates with the Met671 (in this case, actually, Met1) followed immediately by the Asp672 (or rather Asp2, in fact). In this situation, following translation of C99-encoding mRNA generated in the  $\beta$ APP mRNA amplification pathway, the N-terminal methionine will be retained and possibly acetylated ((Ac)Met) and, following the gamma-cleavage, the sequence of the

resulting product of interest, starting from the N-terminus, would be Met-A $\beta$  or (Ac)Met-A $\beta$ . The occurrence of N-terminal Met or acetylated N-terminal Met at the N-terminus of A $\beta$  would, therefore, amount to an "identifier" of its origin and a "reporter" as well as major evidence of its generation in the  $\beta$ APP mRNA amplification pathway. Because of its potential extraordinary significance in defining the etiology of AD and affecting development of the field, the detection of presumably AD-causing, N-terminal Met- or (Ac)Met-carrying A $\beta$  variant directly in the AD-affected brain and the ability to distinguish between A $\beta$  variants generated conventionally by  $\beta$ APP proteolysis and in the  $\beta$ APP mRNA amplification pathway would arguably constitute the Alzheimer's proverbial "Holy Grail". N-terminal methionine-containing A $\beta$  peptide was not and probably could not be seen previously in the human brain because only the massively aggregated, regular detergents-insoluble, extracellular,  $\beta$ APP-derived and secreted beta-amyloid was analyzed ([144], for example); whereas A $\beta$  generated in the  $\beta$ APP mRNA amplification pathway and apparently containing Met or (Ac)Met at its N-end is presumably retained intracellularly, is potentially detergent-soluble, and could have been lost in discarded regular detergent-extracted fractions of brain tissues [144] during preparation of samples (a standard beta-amyloid isolation procedure designed, apparently, to substantially enrich the A $\beta$  content in a sample but potentially throwing the proverbial baby out with the bathwater). Accordingly, it appears that a search for the predicted AD-associated, N-terminal methionine-containing A $\beta$  variant should center on the previously ignored, regular detergent-soluble fraction of brain tissues. Moreover, even if the intracellular proteome were analyzed, within the framework of the proposed mechanism of  $\beta$ APP mRNA amplification described above the detection of such a molecule in the AD-affected human brain could be challenging. Indeed, following the activation of  $\beta$ APP mRNA amplification in neuronal cells, A $\beta$  (or, rather Met-A $\beta$  or (Ac)Met-A $\beta$ ) would accumulate rapidly and trigger processes resulting in cell death and removal. Therefore, the lifespan of AD-affected neurons following the activation of  $\beta$ APP mRNA amplification would be short, the number of intact neuronal cells where the mRNA amplification pathway has been activated would be limited at any given time, and a temporal window of Met-A $\beta$  or (Ac)Met-A $\beta$  detection would be small. On the other hand, the importance of the demonstration of its occurrence, possibly the only feasible way to obtain direct evidence of AD-correlated  $\beta$ APP mRNA amplification in the human brain, would be considerable. As a practical application, the ability to distinguish between A $\beta$  variants generated in the  $\beta$ APP mRNA amplification pathway and in the  $\beta$ APP proteolytic process would greatly simplify the design of model systems for Alzheimer's disease by enabling the use of unmodified human neuronal cell cultures as an AD model.

Importantly, whereas in Alzheimer's disease the presumed activation of the  $\beta$ APP mRNA amplification pathway and consequent neuronal death occur on a massive scale, it is possible, even plausible, that in the normal human brain, as a part of physiologically occurring aging process, a fraction of individual neurons accumulate intracellular, conventionally generated A $\beta$  to levels sufficient to trigger mitochondrial distress, initiate the ISR, and activate  $\beta$ APP mRNA amplification. This would lead, in a time-dependent manner, to persistent death of initially small but increasing (with age) number of neuronal cells and may contribute significantly to cognitive changes occurring in normal aging. If this is the case, the timing and the severity of the "normal" aging-related cognitive decline would depend mainly on the rate of intracellular accumulation of conventionally

produced A $\beta$  in individual neurons to levels triggering mitochondrial dysfunction, the ISR and  $\beta$ APP mRNA amplification, and on the frequency of such occurrences, i.e. the number of cells involved. It follows that Alzheimer's disease differs from the "normal" aging-related cognitive decline only in the rate of intracellular accumulation of conventionally generated beta-amyloid and the amount of neurons where its levels reach the threshold leading to the activation of  $\beta$ APP mRNA amplification. It also follows that mechanistic difference between AD and the "normal" aging-associated cognitive decline is only quantitative, and that designations "normal" and "disease" with regard to aging-related cognitive decline versus AD are purely operational; the processes involved in both states are, apparently, identical, continuous, or at least overlapping. In experimental terms, this implies that N-terminal Met- or (Ac)Met-capped A $\beta$  could be detected not only in AD-affected brain but also in the normal, control, human brain albeit at lower levels. The difference, thus, would be quantitative rather than qualitative. Moreover, in such a case, since the extent of the intraneuronal accumulation of conventionally produced beta-amyloid and, consequently, the rate of activation of the  $\beta$ APP mRNA amplification pathway are functions of time, it could be expected that levels of N-terminal Met- or (Ac)Met-containing beta-amyloid in the normal human brain would increase as aging progresses.

### Mammalian mRNA Amplification, Exogenous RNA-Based Vaccines and mRNA Therapeutics

The significance of the research described above and of its inferences and consequences is not limited to the increased conceptual understanding of one of the fundamental processes of molecular biology, the flow of the protein-encoding genomic information. It includes the increased understanding of the nature of some diseases previously considered untreatable and elucidation of the novel approaches for their prevention and cure. It also includes the potential powerful applications, both medical and biotechnological, of principles of mammalian RNA-dependent mRNA amplification. The recent demonstration, by investigators from NIH/Moderna and Pfizer/BioNTech, of the feasibility of exogenous RNA-based vaccines suggests one of such usages. In this potential application, aimed to significantly increase the efficiency of exogenous RNA-mediated antigen production, an antigen-encoding mRNA is designed in such a way as to be eligible for amplification in the Type I chimeric mRNA amplification pathway. The RNA of interest is then delivered into the cellular environment, i.e. the induced or ongoing mRNA amplification process, where it can be amplified. This can be accomplished by two different approaches. In the first approach, the RNA-dependent mRNA amplification process can be induced, preferably by a substance delivered jointly with an exogenous RNA of interest and capable of eliciting the integrated stress response by activating any of the four eIF2 $\alpha$  kinases: PKR, PERK, GCN2, and HRI, or leading to the ISR-independent assembly and activation of the competent RdRp complex. The second, possibly more immediately practical, approach is to plug in the naturally occurring mammalian RNA-dependent mRNA amplification process. In this procedure, an exogenous RNA is delivered to cells where the mRNA amplification process is operational, for example the erythroblasts. Since in this application the exogenous RNA of interest is, by design, amplification-eligible, it will be amplified and translated alongside the native amplification-eligible mRNA species; as was shown in [8,9], the output of this process, both in terms of the amplified mRNA and, consequently, of

the polypeptide translated from it, can be highly substantial.

Any cells with ongoing RNA-dependent mRNA amplification, for example cells overproducing the ECM proteins, would be capable of amplifying a suitably designed exogenous mRNA. Targeting specific cell types can provide certain control benefits. In the chimeric pathway of mRNA amplification, the duration of the amplification process in the absence of the input of new amplification templates, such as gene-transcribed mRNA, is limited to the lifetime of the initial exogenous mRNA input and could be relatively short. In the iPCR pathway, the duration of the amplification process apparently depends on the persistent cytoplasmic elongation of the 3'-terminal poly(A) segment and could be long. However, in certain cell types, for example erythroblasts, RNA-dependent mRNA amplification occurs for only a short duration and ceases when cells progress to the reticulocyte stage; in such a setting the production of an antigen would be definitively transient. Additional measures of control can be achieved by designing an exogenous RNA in a certain way, for example so as to be translated into a desired protein only in the Type IV chimeric mRNA amplification pathway. In such a case, a desired protein would be produced only in cells with the actively ongoing mRNA amplification process and only via the chimeric amplification pathway, whereas designing an mRNA for the Type I chimeric pathway would ensure that it would be amplified and expressed in both the chimeric and the iPCR pathways, as well as translated conventionally. On the other hand, following the strategy for construction of a reporter for Tier Two, or utilizing a variant of Type II chimeric pathway and applying the principle of the Two-Tier Paradox, an exogenous mRNA could be designed in such a way as to ascertain its expression solely in the iPCR pathway. The choice of mRNA design between these three options would define the duration of expression of an exogenous mRNA. Moreover, in accordance with the same Two-Tier Paradox principle and with the appropriate mRNA design, two different polypeptides could be produced in parallel from the same exogenous mRNA, one in the Type II chimeric pathway and another in the iPCR process.

The utilization, via targeted intracellular amplification of the antigen-encoding mRNA, of mammalian RNA-dependent mRNA amplification in the development of RNA-based vaccines could confer a decisive advantage of achieving a rapid and substantial production of antigen of interest, sufficient to trigger a fast and potent immune response, by delivering only a small amount of suitably designed exogenous RNA into an appropriate cellular environment. Importantly, similar advantage in maximizing the efficient production of a desired protein can be attained by using the above described intracellular RNA-dependent mRNA amplification strategies in mRNA therapeutics. Moreover, these strategies could be instrumental in addressing a persistent challenge of mRNA-based therapies, namely extending and, ultimately, achieving a sustained expression of exogenously introduced mRNA. Interestingly, eIF2 $\alpha$  kinases are often found activated in cancer cells [20,44]; for example activated GCN2 was shown to support, through initiation of the ISR and stimulation of ATF4 expression, tumor cells survival in response to nutrient deprivation [27]. It follows that the ISR and, consequently, the RNA-dependent mRNA amplification process are operational in many types of cancer cells, and that properly designed therapeutic anti-cancer mRNA delivered into these cells could be efficiently and sustainably expressed and be highly effective.

Another conceivable approach to sustainably amplify and express an exogenous mRNA of interest is to exploit mechanisms involved in the replication and expression of HDV RNA [17]. HDV RNA

genome is amplified via a double rolling circle mechanism, with HDAG serving, as was discussed above, as a processivity co-factor of cellular RdRp activity, presumably the RNA polymerase II complex or its components, and intrinsic ribozyme domains cleaving, *in cis*, long RNA molecules into individual units that are subsequently circularized. In parallel, apparently also with the help of HDAG, a linear subgenomic transcription of the HDAG mRNA, 5'-capped and 3'poly(A)-containing, takes place during viral infection. Both mechanism, a rolling circle and a linear transcription, can be adopted for sustainable intracellular production, via RNA-dependent RNA synthesis, of exogenous mRNA. This would require a proper design of an exogenous mRNA of interest and its delivery simultaneously with that of exogenous HDAG mRNA. Moreover, the availability of an exogenous mRNA-encoded processivity co-factor (HDAG) may enable the assembly of the competent RdRp complex and the activation of both Tiers of amplification of the eligible mRNA molecules in any cell type.

## Conclusion

The present article attempts to assess and systematize the current state of the fast-developing field of eukaryotic RNA-dependent mRNA amplification, a novel mode of protein-encoding RNA-to-RNA-to-Protein genomic information transfer, constituting, in addition to the "DNA-to-RNA-to-Protein" modes of genomic information movement, the extension of the Central Dogma of Molecular Biology in the context of mammalian cells. Considering the relative novelty of the field, the volume of the accumulated knowledge, both conceptual and factual, is remarkable. Not only have the fundamental principles of the mRNA amplification process been established, many, if not most, of its mechanistic and regulatory aspects are also well understood, at least in broad outline, by now. Among those, it was determined that the amplification process consists of two intimately linked components. In the first, every conventionally produced, genome-transcribed, amplification-eligible mRNA molecule acts, potentially and repeatedly, as a template for synthesis of additional mRNA molecules. Considering that there could be thousands of conventionally gene-transcribed molecules of a particular mRNA species in the cell, the mRNA amplification process is potentially equivalent in its extent to a thousand-fold gene amplification; at the peak of the erythroid differentiation, for example, the amount of globin mRNA produced in the chimeric amplification pathway is about 1500-fold higher than the amount of conventionally generated globin mRNA in the same cells. Whereas the first component of mRNA amplification, the chimeric pathway, is linear, the second component could be exponential. In it, remarkably, one of the end products of the first component of mRNA amplification can serve as an initial progenitor template in the process that constitutes physiologically occurring intracellular polymerase chain reaction, iPCR. In the course of investigations, a number of regulatory principles, such as the probable involvement, through expression of select transcription factors, of the integrated stress response in the activation of the mRNA amplification process, have been established, and powerful research tools, such as reporter constructs for both components of the amplification process, have been developed. Importantly, a plausible involvement of the mammalian mRNA amplification process in the etiology of various diseases has been convincingly determined. The diseases involved include pathologies associated both with the deficiency of a protein normally produced by this mechanism and with the overproduction of a protein encoded by mRNA species normally not involved in such a process. Certain classes of beta-thalassemia exemplify the former category

whereas the latter is represented by Alzheimer's disease, the prevailing cause of dementia, notoriously impervious to all previous attempts to develop therapeutic treatments and even to produce adequate experimental models. The determination that RNA-dependent amplification of beta-amyloid precursor protein mRNA is driving the disease opens up new experimental approaches and indicates novel therapeutic opportunities. The delineated principles of mammalian RNA-dependent mRNA amplification also suggest potentially powerful applications, both medical and biotechnological. Thus, in light of the recently demonstrated feasibility of RNA-based vaccines (NIH/Moderna and Pfizer/BioNTech Covid-19 vaccines), it can be proposed that the targeted intracellular amplification of exogenously introduced amplification-eligible antigen-encoding mRNAs via the induced or naturally occurring mammalian RNA-dependent mRNA amplification pathway could be of substantial benefit in triggering a fast and potent immune response and instrumental in the development of future vaccines; similar approaches can be effective in the development of efficient and sustainable exogenous mRNA-based therapies. In conclusion, the current understanding of the mammalian RNA-dependent mRNA amplification process constitutes a formidable framework for further developments in the field. The outcomes can be significant.

## Authors Contributions

All concepts described in the present article were conceived and developed by the Authors; V.V. wrote the manuscript.

## Acknowledgements

The concept of the physiologically occurring mammalian intracellular polymerase chain reaction was first formulated in a conversation of one of the Authors (V.V.) with Dr. Philipp Kapranov. The notion of the retention of N-terminal methionine (or acetylated methionine) in human beta-amyloid peptide variant possibly produced in the  $\beta$ APP mRNA amplification pathway was conceptualized in the Authors' conversation with Alexander Varshavsky. Authors are grateful to Dr. Rachel Rits-Volloch for suggesting the title and to Dr. Bjorn R. Olsen for his involvement in the previous joint publications forming basis of the present one and for reading the manuscript and helping with preparation of the figures.

## Conflicts of interest

Authors declare no conflicts of interest.

## Funding

NIH R21 GM056179; NIH RO1 AR036819.

## References

1. Volloch V. Cytoplasmic Synthesis of Globin RNA in Differentiated Murine Erythroleukemia Cells: Possible Involvement of RNA-dependent RNA Polymerase. *Proc Natl Acad Sci USA*. 1986;83:1208-12.
2. Downey K, Byrnes J, Jurmark B, So A. Reticulocyte RNA-Dependent RNA Polymerase. *Proc Natl Acad Sci USA*. 1973;70:3400-4.
3. Volloch V, Schweitzer B, Rits S. Synthesis of Globin RNA in Enuclated Differentiating MEL Cells. *J Cell Biol*. 1987;105:137-43.
4. Richards O, Ehrenfeld E. Poliovirus RNA replication. *Curr Topics Microbiol Immun*. 1990;161:89-119.
5. Volloch V, Schweitzer B, Rits S. Evolutionarily Conserved Elements in the 5'-untranslated Region of  $\beta$  Globin mRNA Mediate Site-specific Priming of a Unique Hairpin Structure during cDNA Synthesis. *Nucl Acids Res*. 1994;22:5302-9.
6. Volloch V, Schweitzer B, Rits S. Antisense Globin RNA in Murine Erythroid Tissues: Structure, Origin and Possible Function. *Proc Natl Acad Sci USA*. 1996;93:2476-81.

7. Volloch V. Determination and characterization of RNA termini by ligation-mediated RNA amplification. In: PCR-Essential Techniques, 1997, BIOS, Oxford, J. Burke, editor.
8. Volloch V. Protein-encoding RNA to RNA information transfer in mammalian cells: Principles of RNA-dependent mRNA amplification. *Ann Integr Mol Med.* 2019;1(1):1002. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6750253/>
9. Rits S, Olsen B, Volloch V. Protein-encoding RNA to RNA information transfer in mammalian cells: RNA-dependent mRNA amplification. Identification of chimeric RNA intermediates and putative RNA end products. *Ann Integr Mol Med.* 2019;1(1):1003. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6814175/>
10. Volloch V, Rits S, Olsen B. RNA-dependent amplification of mammalian mRNA encoding extracellular matrix proteins: Identification of chimeric RNA intermediates for  $\alpha 1$ ,  $\beta 1$ , and  $\gamma 1$  chains of laminin. *Ann Integr Mol Med.* 2019;1(1):1004. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6818727/>
11. Lai M. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis delta virus. *J Virol.* 2005;79:7951-8.
12. Taylor J. Replication of human hepatitis delta virus: recent developments. *Trends Microbiol.* 2003;11:185-90.
13. Lehmann E, Brueckner F, Cramer P. Molecular basis of RNA-dependent RNA polymerase II activity. *Nature.* 2007;450:445-9.
14. Wagner S, Yakovchuk P, Ponicsan S, Drullinger L, Kugel J, Goodrich J. RNA polymerase II acts as an RNA-dependent RNA polymerase to extend and destabilize non-coding RNA. *EMBO J.* 2013;32:781-90.
15. Maida Y, Yasukawa Y, Masutomi K. De Novo RNA Synthesis by RNA-Dependent RNA Polymerase Activity of Telomerase Reverse Transcriptase. *Mol Cell Biol.* 2016;38:1248-59.
16. Kapranov P, Ozsolak F, Kim S, Foissac S, Lipson D, Hart C, et al. New class of gene-termini-associated human RNAs suggests a novel RNA copying mechanism. *Nature.* 2010;466:642-6.
17. Tseng C, Lai M. Hepatitis delta virus RNA replication. *Viruses.* 2009;1:818-31.
18. Brazas R, Ganem D. A cellular homolog of hepatitis delta antigen: implications for viral replication and evolution. *Science.* 1996;272:90-4.
19. Huang C, Lo S. Evolution and diversity of human hepatitis D virus genome. *Adv Bioinformatics.* 2010.
20. Pakos-Zebrucka K, Koryga I, Mnich K. The integrated stress response. *EMBO Rep.* 2016;17:1374-95.
21. Ron D. Translational control in the endoplasmic reticulum stress response. *J Clin Invest.* 2002;110:1383-8.
22. Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calton M, et al. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell.* 2003;11:619-33.
23. Brostrom CO, Prostko CR, Kaufman RJ, Brostrom MA. Inhibition of translational initiation by activators of the glucose-regulated stress protein and heat shock protein stress response systems. Role of the interferon-inducible double-stranded RNA-activated eukaryotic initiation factor 2alpha kinase. *J Biol Chem.* 1996;271:24995-5002.
24. Dever TE, Feng L, Wek RC, Cigan AM, Donahue TF, Hinnebusch AG. Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell.* 1992;68:585-96.
25. Wek RC, Jiang HY, Anthony TG. Coping with stress: eIF2 kinases and translational control. *Biochem Soc Trans.* 2006;34:7-11.
26. Rzymiski T, Milani M, Pike L, Buffa F, Mellor HR, Winchester L, et al. Regulation of autophagy by ATF4 in response to severe hypoxia. *Oncogene.* 2010;29:4424-35.
27. Ye J, Kumanova M, Hart LS, Sloane K, Zhang H, De Panis DN, et al. The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J.* 2010;29:2082-96.
28. Garcia MA, Meurs EF, Esteban M. The dsRNA protein kinase PKR: virus and cell control. *Biochimie.* 2007;89:799-811.
29. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature.* 1999;397:271-4.
30. Berlanga JJ, Herrero S, de Haro C. Characterization of the hemin-sensitive eukaryotic initiation factor 2alpha kinase from mouse nonerythroid cells. *J Biol Chem.* 1998;273:32340-6.
31. Chen JJ, Throop MS, Gehrke L, Kuo I, Pal JK, Brodsky M, London IM. Cloning of the cDNA of the heme-regulated eukaryotic initiation factor 2 alpha (eIF-2 alpha) kinase of rabbit reticulocytes: homology to yeast GCN2 protein kinase and human double-stranded-RNA-dependent eIF-2 alpha kinase. *Proc Natl Acad Sci USA.* 1991;88:7729-33.
32. Shi Y, Vattem KM, Sood R, An J, Liang J, Stramm L, Wek RC. Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, PEK, involved in translational control. *Mol Cell Biol.* 1998;18:7499-509.
33. Meurs E, Chong K, Galabru J, Thomas NS, Kerr IM, Williams BR, et al. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell.* 1990;62:379-90.
34. Ramirez M, Wek RC, Hinnebusch AG. Ribosome association of GCN2 protein kinase, a translational activator of the GCN4 gene of *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1991;11:3027-36.
35. Lavoie H, Li JJ, Thevakumaran N, Therrien M, Sichi F. Dimerization-induced allostery in protein kinase regulation. *Trends Biochem Sci.* 2014;39:475-86.
36. Donnelly N, Gorman AM, Gupta S, Samali A. The eIF2alpha kinases: their structures and functions. *Cell Mol Life Sci.* 2013;70:3493-511.
37. Korennykh A, Walter P. Structural basis of the unfolded protein response. *Annu Rev Cell Dev Biol.* 2012;28:251-77.
38. Wang M, Kaufman RJ. Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature.* 2016;529:326-35.
39. Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat Cell Biol.* 2000;2:326-32.
40. Carrara M, Prisci F, Nowak PR, Kopp MC, Ali MM. Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to initiate ER stress signaling. *eLife.* 2015;4:e03522.
41. Gardner BM, Walter P. Unfolded proteins are Ire1-activating ligands that directly induce the unfolded protein response. *Science.* 2011;333:1891-4.
42. de la Cadena SG, Hernandez-Fonseca K, Camacho-Arroyo I, Massieu L. Glucose deprivation induces reticulum stress by the PERK pathway and caspase-7- and calpain-mediated caspase-12 activation. *Apoptosis.* 2014;19:414-27.
43. Moore CE, Omikorede O, Gomez E, Willars GB, Herbert TP. PERK activation at low glucose concentration is mediated by SERCA pump inhibition and confers preemptive cytoprotection to pancreatic beta-cells. *Mol Endocrinol.* 2011;25:315-26.
44. Hart LS, Cunningham JT, Datta T, Dey S, Tameire F, Lehman SL, et al. ER stress-mediated autophagy promotes Myc-dependent transformation and tumor growth. *J Clin Invest.* 2012;122:4621-34.
45. Donnelly N, Gorman AM, Gupta S, Samali A. The eIF2alpha kinases: their structures and functions. *Cell Mol Life Sci.* 2013;70:3493-511.
46. Vazquez de Aldana CR, Wek RC, Segundo PS, Truesdell AG, Hinnebusch AG. Multicopy tRNA genes functionally suppress mutations in yeast eIF-2 alpha kinase GCN2: evidence for separate pathways coupling GCN4 expression to unchanged tRNA. *Mol Cell Biol.* 1994;14:7920-32.
47. Castilho BA, Shanmugam R, Silva RC, Ramesh R, Himme BM, Sattlegger E. Keeping the eIF2 alpha kinase Gcn2 in check. *Biochim Biophys Acta.* 2014;1843:1948-68.
48. Deng J, Harding HP, Raught B, Gingras AC, Berlanga JJ, Scheuner D, et al. Activation of GCN2 in UV-irradiated cells inhibits translation. *Curr Biol.* 2002;12:1279-86.
49. Lu W, Laszlo CF, Miao Z, Chen H, Wu S. The role of nitric-oxide synthase in

- the regulation of UVB light-induced phosphorylation of the alpha subunit of eukaryotic initiation factor 2. *J Biol Chem.* 2009;284:24281-8.
50. Clemens MJ, Elia A. The double-stranded RNA-dependent protein kinase PKR: structure and function. *J Interferon Cytokine Res.* 1997;17:503-24.
  51. Lemaire PA, Anderson E, Lary J, Cole JL. Mechanism of PKR activation by dsRNA. *J Mol Biol.* 2008;381:351-60.
  52. Balachandran S, Roberts PC, Brown LE, Truong H, Pattnaik AK, Archer DR, et al. Essential role for the dsRNA-dependent protein kinase PKR in innate immunity to viral infection. *Immunity.* 2000;13:129-41.
  53. Lee SB, Esteban M. The interferon-induced double-stranded RNA-activated human p68 protein kinase inhibits the replication of vaccinia virus. *Virology.* 1993;193:1037-41.
  54. Shimazawa M, Hara H. Inhibitor of double stranded RNA-dependent protein kinase protects against cell damage induced by ER stress. *Neurosci Lett.* 2006;409:192-5.
  55. Lee ES, Yoon CH, Kim YS, Bae YS. The double-strand RNA-dependent protein kinase PKR plays a significant role in a sustained ER stress-induced apoptosis. *FEBS Lett.* 2007;581:4325-32.
  56. Garcia MA, Gil J, Ventoso I, Guerra S, Domingo E, Rivas C, et al. Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiol Mol Biol Rev.* 2006;70:1032-60.
  57. Williams BR. PKR; a sentinel kinase for cellular stress. *Oncogene.* 1999;18:6112-20.
  58. Zhou HR, He K, Landgraf J, Pan X, Pestka JJ. Direct activation of ribosome-associated double-stranded RNA-dependent protein kinase (PKR) by deoxynivalenol, anisomycin and ricin: a new model for ribotoxic stress response induction. *Toxins.* 2014;6:3406-25.
  59. Reineke LC, Lloyd RE. The stress granule protein G3BP1 recruits protein kinase R to promote multiple innate immune antiviral responses. *J Virol.* 2015;89:2575-89.
  60. Han AP, Yu C, Lu L, Fujiwara Y, Browne C, Chin G, et al. Heme-regulated eIF2alpha kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. *EMBO J.* 2001;20:6909-18.
  61. Suragani RN, Zachariah RS, Velazquez JG, Liu S, Sun CW, Townes TM, et al. Heme-regulated eIF2alpha kinase activated Atf4 signaling pathway in oxidative stress and erythropoiesis. *Blood.* 2012;119:5276-84.
  62. Ill-Raga G, Tajés M, Busquets-García A, Ramos-Fernandez E, Vargas LM, Bosch-Morato M, et al. Physiological control of nitric oxide in neuronal BACE1 translation by heme-regulated eIF2alpha kinase HRI induces synaptogenesis. *Antioxid Redox Signal.* 2015;22:1295-307.
  63. McEwen E, Kedersha N, Song B, Scheuner D, Gilks N, Han A, et al. Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. *J Biol Chem.* 2005;280:16925-33.
  64. Yerlikaya A, Kimball SR, Stanley BA. Phosphorylation of eIF2alpha in response to 26S proteasome inhibition is mediated by the haem-regulated inhibitor (HRI) kinase. *Biochem J.* 2008;412:579-88.
  65. Lu L, Han AP, Chen JJ. Translation initiation control by heme-regulated eukaryotic initiation factor 2alpha kinase in erythroid cells under cytoplasmic stresses. *Mol Cell Biol.* 2001;21:7971-80.
  66. Guo X, Aviles G, Tian R. Mitochondrial stress is relayed to the cytosol by an OMA1-DELE1-HRI pathway. *Nature.* 2020;579:427-32.
  67. Fessler E, Eckl E, Shmitt S. A pathway coordinated by DELE1 relays mitochondrial stress to the cytosol. *Nature.* 2020;579:433-7.
  68. Volloch V. A mechanism for  $\beta$ -amyloid overproduction in Alzheimer's disease: Precursor-independent generation of  $\beta$ -amyloid via antisense RNA-primed mRNA synthesis. *FEBS Lett.* 1996;390:124-8.
  69. Volloch V. Mechanism for  $\beta$ -amyloid overproduction in sporadic Alzheimer's Disease: Possible antisense RNA-mediated generation of a 5'-truncated  $\beta$ APP mRNA encoding 12 kDa C-terminal fragment of  $\beta$ APP, the immediate precursor of A $\beta$ . In: *Molecular Mechanisms of Dementia.* 1997, W. Wasco and R. Tanzi, Eds.
  70. Volloch V. Possible mechanism for resistance to Alzheimer's disease (AD) in mice suggests new approach to generate a mouse model for sporadic AD and may explain familial resistance to AD in man. *Exp Neurobiol.* 1997;144:214-8.
  71. Volloch V, Rits S. Results of beta-secretase-inhibitor clinical trials support amyloid precursor protein-independent generation of beta amyloid in sporadic Alzheimer's disease. *Med Sci.* 2018.
  72. Volloch V, Olsen B, Rits S. Precursor-independent overproduction of beta-amyloid in AD: Mitochondrial dysfunction as possible initiator of asymmetric RNA-dependent  $\beta$ APP mRNA amplification. An engine that drives Alzheimer's disease. *Ann Integr Mol Med.* 2019;1(1):1005. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6922309/>
  73. Volloch V, Olsen B, Rits S. AD "Statin": Alzheimer's disorder is a "fast" disease preventable by therapeutic intervention initiated even late in life and reversible at the early stages. *Ann Integr Mol Med.* 2020;2(1):1006. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7083596/>
  74. Volloch V, Olsen B, Rits S. Alzheimer's disease is driven by intraneuronally retained beta-amyloid produced in the AD-specific, bAPP-independent pathway: Current perspective and experimental models for tomorrow. *Ann Integr Mol Med.* 2020;2(1):1007. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7331974/>
  75. Volloch V, Olsen B, Rits S. Alzheimer's disease prevention and treatment: Case for optimism. *Ann Integr Mol Med.* 2020;2(1):1008. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7546530/>
  76. Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell.* 2001;107(7):881-91.
  77. Jiang S, Zhang E, Zhang R, Li X. Altered activity patterns of transcription factors induced by endoplasmic reticulum stress. *BMC Biochem.* 2016;17:8.
  78. Liang Y, Song Y, Zhang F, Sun F, Wag N. Effect of a Single Nucleotide Polymorphism in the LAMA1 Promoter Region on Transcriptional Activity: Implication for Pathological Myopia. *Curr Eye Res.* 2016;10:1379-86.
  79. Piccinni S, Bolcato-Bellemin A, Klein A. Kruppel-like factors regulate the Lama1 gene encoding the laminin chain. *J. Biol. Chem.* 2004;279:9103-14.
  80. Okano R, Mita T, Matsui T. Characterization of a novel promoter structure and its transcriptional regulation of the murine laminin B1 gene. *Biochem. Biophys. Acta.* 1992;1132:49-57.
  81. Lodish H, Berk A, Zipurski A. In: *Molecular Cell Biology.* New York, 2000, Freeman and Co.
  82. Cao A, Galanello R. Beta-thalassemia. *Genet Med.* 2010;12:61-76.
  83. Egan M, Kost J, Tariot P. Randomized trials of verubecestat for mild-to-moderate Alzheimer's disease. *N Engl J Med.* 2018;378:1691-703.
  84. Breimer L, Denny P. Alzheimer amyloid aspects. *Nature.* 1987;326:749-50.
  85. Dyrks T, Dyrks E, Monning U, Urmoneit B, Turner J, Beyreuther, K. Generation of beta A4 from the amyloid protein precursor and fragments thereof. *FEBS Lett.* 1993;335:89-93.
  86. Haass C, Lemere C, Capell A, Citron M, Seubert P, Schenk D. The Swedish mutation causes early-onset Alzheimer's disease by beta-secretase cleavage within the secretory pathway. *Nat Med.* 1995;1:1291-6.
  87. Iizuka T, Shoji M, Kawarabayashi T, Sato M, Kobayashi T. Intracellular generation of amyloid beta-protein from amyloid beta-protein precursor fragment by direct cleavage with beta- and gamma-secretase. *Biochem. Biophys. Res Commun.* 1996;218:238-42.
  88. Citron M, Haass C, Selkoe D. Production of amyloid beta peptide by cultured cells: no evidence for internal initiation of translation at Met596. *Neurobiol Aging.* 1993;14:571-3.
  89. Mita S, Sadlock J, Herbert J, Schon E. A cDNA specifying the human amyloid beta precursor protein encodes a 95-kDa polypeptide. *Nucl Acids Res.* 1988;16:9351.

90. Mita S, Sadlock J, Herbert J, Schon E. A cDNA specifying the human amyloid beta precursor protein encodes a 95-kDa polypeptide: Correction. *Nucl Acids Res.* 1988;16:11402.
91. Salbaum J, Weidemann A, Lemaire H, Masters C, Beyreuther K. The promoter of Alzheimer's disease amyloid A4 precursor gene. *EMBO J.* 1988;7:2807-13.
92. Zhu X, Perry G, Moreira PI, Aliev G, Cash AD, Hirai K, et al. Mitochondrial abnormalities and oxidative imbalance in Alzheimer disease. *J Alzheimers Dis.* 2006;9:147-53.
93. Blass JP. The mitochondrial spiral. An adequate cause of dementia in the Alzheimer's syndrome. *Ann N Y Acad Sci.* 2000;924:170-83.
94. Manczak M, Park BS, Jung Y, Reddy PH. Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease: implications for early mitochondrial dysfunction and oxidative damage. *Neuromolecular Med.* 2004;5:147-62.
95. Qin W, Haroutunian V, Katsel P, Cardozo CP, Ho L, Buxbaum JD, et al. PGC-1alpha expression decreases in the Alzheimer disease brain as a function of dementia. *Arch Neurol.* 2009;66:35261.
96. Du H, Guo L, Yan S, Sosunov AA, McKhann GM, Yan SS. Early deficits in synaptic mitochondria in an Alzheimer's disease mouse model. *Proc Natl Acad Sci U S A.* 2010;107:18670-5.
97. Lin MT, Simon DK, Ahn CH, Kim LM, Beal MF. High aggregate burden of somatic mtDNA point mutations in aging and Alzheimer's disease brain. *Hum Mol Genet.* 2002;11:133-45.
98. Calkins M, Manczak M, Mao P, Shirendeb U, Reddy PH. Impaired mitochondrial biogenesis, defective axonal transport of mitochondria, abnormal mitochondrial dynamics and synaptic degeneration in a mouse model of Alzheimer's disease. *Hum Mol Genet.* 2011.
99. Anandatheerthavarada HK, Biswas G, Robin MA, Avadhani NG. Mitochondrial targeting and a novel transmembrane arrest of Alzheimer's amyloid precursor protein impairs mitochondrial function in neuronal cells. *J Cell Biol.* 2003;161:41-54.
100. Caspersen C, Wang N, Yao J. Mitochondrial Abeta: a potential focal point for neuronal metabolic dysfunction in Alzheimer's disease. *FASEB J.* 2005;19:2040-1.
101. Chen JX, Yan SS. Role of mitochondrial amyloid-beta in Alzheimer's disease. *J Alzheimers Dis.* 2010;22:S569-78.
102. Hansson C, Petersen CA, Alikhani N, Behbahani H. The amyloid beta-peptide is imported into mitochondria via the TOM import machinery and localized to mitochondrial cristae. *Proc Natl Acad Sci U S A.* 2008;105:13145-50.
103. de la Monte SM, Luong T, Neely TR, Robinson D, Wands JR. Mitochondrial DNA damage as a mechanism of cell loss in Alzheimer's disease. *Lab Invest.* 2000;80:1323-35.
104. Brooks WM, Lynch PJ, Ingle CC, Hatton A, Emson PC, Faull RL, et al. Gene expression profiles of metabolic enzyme transcripts in Alzheimer's disease. *Brain Res.* 2007;1127:127-35.
105. Wang X, Perry G, Smith MA, Zhu X. Amyloid-beta-derived diffusible ligands cause impaired axonal transport of mitochondria in neurons. *Neurodegener Dis.* 2007;7:56-9.
106. Wang X, Su B, Fujioka H, Zhu X. Dynamin-like protein 1 reduction underlies mitochondrial morphology and distribution abnormalities in fibroblasts from sporadic Alzheimer's disease patients. *Am J Pathol.* 2008;173:470-82.
107. Wang X, Su B, Lee HG, Li X, Perry G, Smith MA, Zhu X. Impaired balance of mitochondrial fission and fusion in Alzheimer's disease. *J Neurosci.* 2009;29:9090-103.
108. Wang X, Su B, Siedlak SL, Moreira PI, Fujioka H, Wang Y, et al. Amyloid-beta overproduction causes abnormal mitochondrial dynamics via differential modulation of mitochondrial fission/fusion proteins. *Proc Natl Acad Sci U S A.* 2008;105:19318-23.
109. Chafekar S, Baas F, Scheper W. Oligomer-specific amyloid-beta toxicity in cell models is mediated by selective uptake. *Biochem Biophys Acta.* 2008;9:523-31.
110. Wesen E, Jeffries G, Dzebo M, Esbjornner M. Endocytic uptake of monomeric amyloid- $\beta$  peptides is clathrin- and dynamin-independent and results in selective accumulation of A $\beta$ (1-42) compared to A $\beta$ (1-40). *Sci Rep.* 2017;7:2021.
111. Hu X, Crick S. Amyloid seeds formed by cellular uptake, concentration, and aggregation of the amyloid-beta peptide. *Proc Natl Acad Sci USA.* 2009;106:20324-9.
112. Yajima R, Tokutake T, Koyama A. ApoE-isoform-dependent cellular uptake of amyloid- $\beta$  is mediated by lipoprotein receptor LR11/SorLA. *Biochem Biophys Res Comm.* 2015;456:482-8.
113. Omtri R, Davidson M, Arumugam B. Differences in the cellular uptake and intracellular itineraries of amyloid beta proteins 40 and 42: ramifications for the Alzheimer's drug discovery. *Mol Pharmaceutics.* 2012;9:1887.
114. Bu G, Cam J, Zerbini C. LRP in amyloid-b production and metabolism. *Ann NY Acad Sci.* 2006;1086:35-53.
115. Wang H. b-Amyloid1-42 binds to  $\alpha 7$  nicotinic acetylcholine receptor with high affinity. Implications for Alzheimer's disease pathology. *J Biol Chem.* 2000;275:5626-32.
116. Nagele R, D'Andrea, M, Anderson, W, Wang H. Intracellular accumulation of Ab42 in neurons is facilitated by the  $\alpha 7$  nicotinic acetylcholine receptor in Alzheimer's disease. *Neuroscience.* 2002;110:199-211.
117. Oddo S. Chronic nicotine administration exacerbates tau pathology in a transgenic model of Alzheimer's disease. *Proc Natl Acad Sci USA.* 2005;102:3046-51.
118. Yan S. RAGE and amyloid-b peptide neurotoxicity in Alzheimer's disease. *Nature.* 1996;382:685-91.
119. Sasaki N. Immunohistochemical distribution of the receptor for advanced glycation end products in neurons and astrocytes in Alzheimer's disease. *Brain Res.* 2001;888:256-62.
120. Deane R. RAGE mediates amyloid-b peptide transport across the blood-brain barrier and accumulation in brain. *Nature Med.* 2003;9:907-13.
121. Iribarren P, Zhou Y, Hu J, Le Y, Wang, J. Role of formyl peptide receptor-like 1 (FPR1/FPR2) in mononuclear phagocyte responses in Alzheimer disease. *Immunol Res.* 2005;31:165-76.
122. Snyder E. Regulation of NMDA receptor trafficking by amyloid-b. *Nature Neurosci.* 2005;8:1051-8.
123. LaFerla F, Green K, Oddo S. Intracellular amyloid-beta in Alzheimer's disease. *Nat Rev Neurosci.* 2007;8(7):499-509.
124. Kinoshita A. Demonstration by FRET of BACE interaction with the amyloid precursor protein at the cell surface and in early endosomes. *J Cell Sci.* 2003;116:3339-46.
125. Xu H, Greengard P, Gandy S. Regulated formation of Golgi secretory vesicles containing Alzheimer b- amyloid precursor protein. *J Biol Chem.* 1995;270:23243-5.
126. Mizuguchi M, Ikeda K, Kim S. Differential distribution of cellular forms of b-amyloid precursor protein in murine glial cell cultures. *Brain Res.* 1992;584:219-25.
127. Cook D. Alzheimer's Ab42 is generated in the endoplasmic reticulum/intermediate compartment of NT2N cells. *Nature Med.* 1997;3:1021-3.
128. Hartmann T. Distinct sites of intracellular production for Alzheimer's disease Ab40/42 amyloid peptides. *Nature Med.* 1997;3:1016-20.
129. Wild-Bode C. Intracellular generation and accumulation of amyloid beta-peptide terminating at amino acid 42. *J Biol Chem.* 1997;272:16085-8.
130. Lee S. A detergent-insoluble membrane compartment contains Ab *in vivo*. *Nature Med.* 1998;4:730-4.
131. Skovronsky D, Doms R, Lee V. Detection of a novel intraneuronal pool of insoluble amyloid b protein that accumulates with time in culture. *J Cell Biol.* 1998;141:1031-9.
132. Manczak M. Mitochondria are a direct site of Ab accumulation in Alzheimer's disease neurons: implications for free radical generation and oxidative damage in

- disease progression. *Hum Mol Genet.* 2006;15:1437-49.
133. Sannerud R. Restricted Location of PSEN2/gamma-secretase determines substrate specificity and generates an intracellular Abeta pool. *Cell.* 2016;166:193-208.
134. Hansson C. Nicastrin, presenilin, APH-1, and PEN-2 form active b-secretase complexes in mitochondria. *J Biol Chem.* 2004;279:51654-60.
135. Caspersen C. Mitochondrial Ab: a potential focal point for neuronal metabolic dysfunction in Alzheimer's disease. *FASEB J.* 2005;19:2040-1.
136. Keskin A, Kekuš M, Adelsberger H, Neumann U, Shimshek D. BACE inhibition-dependent repair of Alzheimer's pathophysiology. *Proc Natl Acad Sci USA.* 2017;114:8631-6.
137. Hu X, Das B, Hou H, He W, Yan R. BACE1 deletion in the adult mouse reverses preformed amyloid deposition and improves cognitive functions. *J Exp Med.* 2018;10:1084.
138. Kearse M, Wilusz J. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* 2017;31:1717-31.
139. Teo E, Ravi S, Barardo D, Kim H, Fong S. Metabolic stress is a primary pathogenic event in transgenic *Caenorhabditis elegans* expressing pan-neuronal human amyloid beta. *eLife.* 2019.
140. Choi SH, Kim YH, Hebisch M, Sliwinski C, Lee S, D'Avanzo C, et al. A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature.* 2014;515:274-8.
141. Frottin F, Martinez A, Peinot P, Mitra S, Holz R, Giglione C, Meinnel T. The proteomics of N-terminal methionine cleavage. *Mol Cell Proteomics, MCP* 2006;5(12):2336-49.
142. Xiao Q, Zhang F, Nacev B, Liu J, Pei D. Protein N-Terminal Processing: Substrate Specificity of *Escherichia coli* and Human Methionine Aminopeptidases. *Biochemistry* 2010;49:5588-99.
143. Varshavsky A. The N-end rule pathway and regulation by proteolysis. *Protein Sci* 2011;20:1298-1345.
144. Naslund J, Schierhorn A, Hellman U, et al. Relative abundance of Alzheimer Aβ amyloid peptide variants in Alzheimer disease and normal aging. *Proc Natl Acad Sci* 1994, 91: 8378-8382.
145. Mullan M, Crawford F, Axelman K, Houlden H, Lilius L, Winblad B, et al. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nat Genet* 1992;1(5):345-7.
146. Haass C, Lemere C, Capell A, Citron M, Seubert P, Schenk D, Lannfelt L, Selkoe D. The Swedish mutation causes early-onset Alzheimer's disease by beta-secretase cleavage within the secretory pathway. *Nat Med* 1995;1(12):1291-6.
147. O'Brien R, Wong P. Amyloid precursor protein processing and Alzheimer's Disease. *Annu Rev Neurosci* 2011;34:185-204.
148. Martin B, Schrader-Fischer G, Busciglio J, Duke M, Paganetti P, Yankner B. Intracellular accumulation of beta-amyloid in cells expressing the Swedish mutant amyloid precursor protein. *J Biol Chem* 1995;270:26727-30.