**Short communication**

Translational stop codons in the precore sequence of hepatitis B virus pre-C RNA allow translation reinitiation at downstream AUGs

Nathalie Fouillot† and Jean-Michel Rossignol*

Laboratoire de Génétique des Virus, UPR 2431-CNRS, Avenue de la Terrasse, 91198 Gif sur Yvette cedex, France

Hepatitis B virus (HBV) wild-type pre-C RNA directs the synthesis of the HBeAg precursor but does not serve as mRNA for translation of the adjacent downstream C gene which encodes the core protein. Using bicistronic mRNA constructs that mimic pre-C RNA, we have demonstrated that this RNA likewise does not serve as messenger for translation of the P gene, which is located downstream of the C gene. However, when the pre-C RNA contains a translational stop codon at position 2 or 28 of the pre-C sequence (as in certain HBV mutants), it no longer directs synthesis of the HBeAg precursor but instead translation is initiated at downstream C and P gene AUGs. We propose that this occurs by a translation reinitiation mechanism.

Hepatitis B virus (HBV) pregenomic RNA and pre-C RNA are viral transcripts which comprise the entire length of the genome plus some 130 to 160 nucleotides respectively of terminal redundancy (Will et al., 1987). The pregenomic RNA has two functions: it serves as template for reverse transcription and directs the synthesis of the core protein from the C gene and the protein(s) involved in replication of the genome from the overlapping P gene. Unlike the pregenomic RNA, the pre-C RNA, a slightly longer transcript which includes the 29-codon pre-C sequence, is not encapsidated in the core particles (Nassal et al., 1990). This mRNA directs the synthesis of the HBe antigen (HBeAg) precursor but not synthesis of the core protein (Standing et al., 1988; Jean-Jean et al., 1989a). Circumstantial evidence suggests that the pre-C RNA does not serve as template for translation of the P protein(s) (Fouillot et al., 1993).

In some cases of chronic or fulminant hepatitis, mutant viruses have been detected on the basis of their inability to produce HBeAg (HBe− viruses). The most frequent mutation is the change of pre-C codon 28 from TGG to the translational stop codon TAG (Brunetto et al., 1990; Tong et al., 1990; Hasegawa et al., 1991; Pollicino et al., 1995). Another mutation leading to a translational stop at codon 2 has also been reported in a few cases (Bhat et al., 1990; Pollicino et al., 1995). As HBe− virus-infected patients are anti-HBe positive, it seems that the initial infection is due to the wild-type virus (Brunetto et al., 1991). However, how the HBe− viruses become predominant during persistent infection remains an open question. A possible answer is that the emergence of these mutants is an immunological escape mechanism of HBV, as it is generally assumed that HBeAg is an important target of cytotoxic T lymphocytes for HBV clearance (Schlicht et al., 1991). Another possibility could be a selective advantage of the HBe mutant viruses in the process of encapsidation of the pregenomic RNA which requires the P protein(s) acting primarily in cis (Bartenschlager et al., 1990). Very few, possibly one, P protein molecules are packaged with HBV RNA (Bartenschlager & Schaller, 1992). It has been suggested that the P gene is translated at a low level by a complex leaky scanning-reinitiation mechanism (Fouillot et al., 1993), by which the mutated pre-C RNA would be able to direct the synthesis of the P protein(s). Consequently, the precore mutants would produce more P proteins than the wild-type HBV, resulting in a slight advantage in the encapsidation process in which the P protein seems to be limiting (Bartenschlager & Schaller, 1992).

However, whether the mutated pre-C RNA would direct the synthesis of the P and the core proteins remained to be determined. To answer this question, we have designed plasmids producing bicistronic mRNAs that mimic the viral pregenome, the wild-type pre-C RNA and the pre-C RNA with a translational stop codon at either position 2 or 28 of the pre-C sequence. In
Fig. 1. Organization of HBV C region–lacZ gene recombinant plasmids. Positions of the restriction sites used for cloning the HBV sequences correspond to subtype ayw (GenBank). The truncated lacZ gene was cloned from the second MluI site of the lacZ gene of plasmid pCH110 (Pharmacia). Rectangles depict open reading frames (ORFs). Circles in these ORFs represent AUGs among which precore (pre-C), core (C), reverse transcriptase (P) and truncated β-galactosidase (lacZ') initiation codons are indicated. RSV, Rous sarcoma virus promoter.

addition, a plasmid carrying a translational stop codon at codon 18 (so far not detected in HBe− viruses) was constructed to evaluate the influence of the distance between stop codons and downstream AUGs on reinitiation efficiency under our experimental conditions.

Plasmid pRPCLAC contains the Rous sarcoma virus (RSV) promoter followed by the pre-C sequence, the C gene of HBV and by the first 150 nucleotides of the P gene fused to the last 975 nucleotides of the Escherichia coli lacZ reporter gene (Fig. 1). From plasmid pRPCLAC, using PCR and appropriate mutated primers, we obtained three singly mutated constructs carrying one translational stop codon in the pre-C sequence: plasmids pRPC2LAC and pRPC18LAC in which mutations at positions 1817 and 1865, respectively, change CAA codons into TAA codons; and plasmid pRPCsLAC in which a mutation at position 1896 changes the TGG codon into a TAG codon. Plasmid pRCLAC is identical to plasmid pRPCLAC except that it does not contain the pre-C sequence (Fig. 1). Plasmid pRLAC contains only the truncated lacZ gene downstream from the RSV promoter (Fig. 1).

Transient expression in human 293 cells of HBeAg precursor, core protein and P-β-gal fusion protein were studied by immunoprecipitation as described previously (Carlier et al., 1995), using anti-HBc (Dako) and anti-β-galactosidase (Sigma) antisera. Proteins were separated by electrophoresis on 12.5% SDS–polyacrylamide gels. After treatment with Amplify (Amersham) gels were dried prior to autoradiography with Fuji RX film.

Proteins immunoprecipitated from cells transfected with the different plasmids are shown in Fig. 2. Transfection with plasmid pRCLAC resulted in the synthesis of both the core protein (HBC, Fig. 2a) and the P-β-gal fusion protein (arrowhead, Fig. 2c). In contrast, with pRPCLAC transfected cells the 22 kDa precursor
of the HBeAg (P22, Fig. 2a) was detected. In agreement with previous data (Standring et al., 1988; Jean-Jean et al., 1989a), the 21 kDa core protein was not observed; the barely visible protein migrating below P22 is most likely a processed form of P22, as is another faint band (HBe, Fig. 2a) which has the expected molecular mass of mature HBeAg. Likewise, the 43 kDa P-β-gal fusion protein was not detected in this cell extract, consistent with the idea that the pre-C RNA cannot serve as template for P gene translation, as suggested by previous data (Fouillot et al., 1993). As expected, the secreted mature 15 kDa HBeAg was present in the cell culture medium (HBe, Fig. 2b).

Expression of the three mutated constructs, pRPC2LAC, pRPC18LAC and pRPC28LAC, led to the presence of both core (C, Fig. 2a) and P-β-gal fusion (arrowhead, Fig. 2c) proteins in the cell extracts while P22 was no longer detected (Fig. 2a). This suggests that a premature arrest of pre-C translation results in the synthesis of proteins initiated from the downstream AUGs. Interestingly, while the P-β-gal fusion protein was translated at approximately the same level in both cases (Fig. 2c), the core protein was synthesized in lesser amounts in pRPC28LAC transfected cells than in other transfected cells (Fig. 2a). In addition to the HBV-encoded proteins, a 32 kDa protein (LAC, Fig. 2c) was immunoprecipitated with the anti-β-galactosidase antiserum from extracts of transfected cells, except those transfected with pRPCCLAC. This protein most likely corresponds to the 279 amino acid β-gal protein translated from the lacZ AUG, as expected if leaky scanning occurred at the P gene AUG, as we observed previously (Jean-Jean et al., 1989b).

Our results show that both the C and P-LAC genes were translated from mRNAs carrying nonsense mutations in the pre-C sequence. Two arguments allow us to exclude the possibility that these genes are translated from minor mRNAs starting downstream from the pre-
C AUG: (i) the C and P-LAC genes were not translated from the non-mutated pre-C RNA; (ii) the RSV promoter has only one known transcription start site (Gorman et al., 1982). Thus, most likely, these two genes are translated from the mutated pre-C RNA by a translation reinitiation mechanism. This hypothesis is supported by our finding that the core protein was synthesized in higher amounts in cells transfected with pRPC<sub>C</sub>LAC and pRPC<sub>g</sub>LAC than with pRPC<sub>g</sub>LAC. As the stop codons are respectively 81, 33 and 3 nucleotides upstream of the C AUG in these constructs, our results are in good agreement with data from Kozak (1987) showing that the reinitiation efficiency of ribosomes is low when the downstream AUG is located close to the preceding ORF. It is worth noting that the amount of core protein synthesized was very similar in pRPC<sub>C</sub>LAC and pRPC<sub>g</sub>LAC transfected cells, leading to the conclusion that the reinitiation efficiency was similar whether the stop codon was located 33 or 81 nucleotides upstream from the next initiation codon. Likewise, we did not observe any differences in the amount of P-β-gal fusion protein whatever the position of the stop codon in the pre-C sequence. From these data, we conclude that the presence of a stop codon in the pre-C sequence results in reinitiation of translation at the downstream C and P genes.

Although our studies were limited to transfection of cells in culture, they suggest that in vivo a pre-C RNA carrying a nonsense mutation on codon 28 of the pre-C sequence could direct synthesis of P protein(s) from this mRNA; (ii) this pre-C RNA is probably encapsidated, as Nassal et al. (1990) have shown that stopping translation of the pre-C sequence before the encapsidation signal (i.e. at codon 13) allows the mutant pre-C RNA to be packaged. Nevertheless, such RNA yields linear DNA molecules which are defective for subsequent replication (Junker-Niepmann et al., 1990). It is also worth noting that nonsense mutations at codon 2 were found to be associated with nonsense mutations at codon 28 (Pollicino et al., 1995). Thus, it is likely that the effect of a mutation at codon 2 would again be a slight increase in the amount of P protein(s) molecules available for pregenome encapsidation.

In conclusion, we have demonstrated for the first time that the wild-type pre-C RNA cannot direct the synthesis of the P protein(s). This was not the case with pre-C RNA carrying nonsense mutations; we show that this RNA can serve as template for translation of the C and P genes. It is tempting to speculate that virus carrying such mutations could have a slight advantage in the encapsidation process over the wild-type virus, which could result, after multiple replication cycles during persistent infection, in the emergence of these mutants. So far, this hypothesis is not supported by any in vivo data.

We would like to thank Dr Hans Will for his helpful suggestions. This work was supported by grants from the Association pour la Recherche sur le Cancer and the Ligue Nationale contre le Cancer. Nathalie Fouillot was supported by a training fund from the Association pour la Recherche sur le Cancer.

References


GORMAN, C. M., MERLINO, G. T., WILLINGHAM, M. C., PASTAN, I. & HOWARD, B. H. (1982). The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eucaryotic
HBV pre-C stop codons and downstream reinitiation


(Received 1 December 1995; Accepted 8 February 1996)