Effect of frameshift mutation in the pre-C region of hepatitis B virus on the X and C genes

Seong Kee Kim, Sung Key Jang and Hyune Mo Rho

1Department of Molecular Biology and Research Center for Cell Differentiation, Seoul National University, Seoul 151-742 and 2Department of Life Science, Pohang Institute of Science and Technology, Pohang 790-784, Korea

We have previously cloned a mutant hepatitis B virus (HBV) genome which had one thymidine addition in the pre-C region resulting in a frameshift mutation in the pre-C region and fusion of the X and C genes. We constructed plasmids containing serially deleted and/or back-mutated (authentic) pre-C regions to study the effect of the frameshift mutation. COS cells transfected with plasmids containing the frameshifted pre-C region produced a 21K C protein (P21c) but not a 22K partially processed pre-C protein (P22). On the other hand, COS cells transfected with plasmids containing the back-mutated pre-C region produced P22. This result was also observed in HepG2-K8 cells producing the mutant HBV particles. Therefore, the pre-C region of HBV is likely to be non-essential for virus replication. COS cells transfected with the plasmid containing a fused X–C open reading frame (ORF) produced a 40K X–C fusion protein. This X–C fusion protein exerted transcriptional trans-activation. These results suggest that the mutant HBV has a C gene with a defective pre-C region and a fused X–C ORF, and hence cannot synthesize 16K HBeAg (P16e).

Genomic DNA of hepatitis B virus (HBV) has four open reading frames (ORFs) which are designated pre-C/C, P, pre-S/S and X (Ganem & Varmus, 1987; Tiollais et al., 1985). In HBV the coding region for the C protein (P21c) is preceded by an in-phase ORF termed the pre-C region (Ou et al., 1986). Translation initiation from the pre-C initiation codon produces a pre-C protein (P25) that contains the entire C protein sequence and a 29-residue amino-terminal extension (pre-C region). It has been demonstrated that the first 19 amino acids of the pre-C region form a signal sequence to direct the pre-C protein to the endoplasmic reticulum, where it is cleaved, forming the partially processed pre-C protein (P22) (Bruss & Gerlich, 1988; Garcia et al., 1988; Junker et al., 1987; Ou et al., 1986, 1988; Roossinck et al., 1986; Weimer et al., 1987). The pre-C region produced P22. This result was also observed in HepG2-K8 cells producing the mutant HBV particles. Therefore, the pre-C region of HBV is likely to be non-essential for virus replication. COS cells transfected with the plasmid containing a fused X–C open reading frame (ORF) produced a 40K X–C fusion protein. This X–C fusion protein exerted transcriptional trans-activation. These results suggest that the mutant HBV has a C gene with a defective pre-C region and a fused X–C ORF, and hence cannot synthesize 16K HBeAg (P16e).

Meyers et al., 1986; Moriarty et al., 1985). It has been demonstrated that the X protein is a transcriptional activator of various enhancer–promoter combinations, including the HBV enhancer and C gene promoter, the simian virus 40 (SV40) enhancer and early promoter, and the β-interferon gene (Spandau & Lee, 1988; Twu & Schloemer, 1987; Zahn et al., 1988).

A novel HBV protein of 35K to 40K was discovered in HBV-transfected HepG2 cells (Bchini et al., 1990) and in core particles isolated from HBV-infected livers (Farza et al., 1988; Feitelson, 1986). It was characterized by Western blotting and proposed to be an X–C fusion protein. However, HBV does not have a fused X–C ORF. The biological function and the mechanism of production of such an X–C fusion protein are as yet unknown. The possibility of ribosomal frameshifting causing X–C fusion protein production has been excluded (Lo et al., 1990).

Previously, we have cloned a mutant HBV (adr) genome from a Korean hepatitis B patient (Choi et al., 1984) and determined the complete sequence of 3213 nucleotides (nt) (Rho et al., 1989). The DNA sequence analysis revealed the addition of one thymidine at nt 1821 of the pre-C region, which resulted in a frameshift mutation in the pre-C region and fusion of the X and C genes. In this paper, we characterized the proteins expressed from the pre-C/C and fused X–C ORFs of the mutant HBV. As the fused X–C ORF contained a region

0001-1967 © 1994 SGM
encoding part of the X protein, we also tested whether the fused X-C ORF products possess trans-activating activity.

In the mutant HBV, the putative initiation codon of the elongated pre-C region at nt 1758 was located 53 nt upstream of the initiation codon of the authentic pre-C region at nt 1811 (Fig. 1a). Therefore we tested whether the mutant HBV contained an elongated pre-C region instead of the authentic pre-C region which may be involved in generating HBeAg. To investigate the utilization of the AUG codons of the pre-C region for the HBV core antigen (HBcAg) and HBeAg synthesis, we constructed recombinant plasmids containing serially deleted and/or back-mutated (authentic) pre-C regions (Fig. 1). To construct 5' deletion mutant plasmids, pUSCP containing the C gene and pre-C region was opened with BamHI and serially deleted with Bal 31 nuclease. After filling of the staggered ends with the Klenow fragment, BamHI linkers were attached. The resulting series of 5' deletion mutants was selected by DNA sequencing (data not shown). To recover the promoter region, the SV40 late promoter region of pUSCP was cloned between BamHI and HindIII sites of the deletion mutant plasmids. Mutants with endpoints at -180 (pUSCP1), -129 (pUSCP3) and -32 (pUSCP) were selected (numbering system starts with the A of the C gene initiation codon). For the construction of plasmids pUSCP10 and pUSCP30 containing the back-mutated pre-C region, site-directed mutagenesis was performed using PCR (Kammann et al., 1989; Kim et al., 1992b; Saiki et al., 1988) with a mutagenic primer. Plasmids pUSCP10 and pUSCP30 were selected by DNA sequencing (data not shown).

Each of the vectors was introduced into COS cells
Table 1. Relative levels of expression of HBcAg and HBeAg in recombinant plasmid-transfected COS cells

<table>
<thead>
<tr>
<th>Plasmid*</th>
<th>Radioactivity (c.p.m.)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell extract</td>
</tr>
<tr>
<td>pUC19</td>
<td>607</td>
</tr>
<tr>
<td>pUSCP1</td>
<td>17262</td>
</tr>
<tr>
<td>pUSCP10</td>
<td>19184</td>
</tr>
<tr>
<td>pUSCP3</td>
<td>24624</td>
</tr>
<tr>
<td>pUSCP30</td>
<td>25423</td>
</tr>
<tr>
<td>pUSC</td>
<td>46012</td>
</tr>
<tr>
<td>pSVXC</td>
<td>11786</td>
</tr>
</tbody>
</table>

* Plasmids are described in Fig. 1(b) and 4(a). † Cell extracts and culture media were one-fifth of the cell lysate, and one-fifth of the supernatant, respectively, from a 100 mm dish. The values are the mean of three determinations.

using the calcium phosphate method (Chen & Okayama, 1987). At 72 h post-transfection, cells and culture media were collected separately. Antigens from COS cells transfected with the recombinant plasmids were assayed with the Abbott HBc radioimmunoassay kit (Abbott Laboratories) (Table 1). Note that this assay cannot differentiate between HBcAg and HBeAg. In the case of plasmids pUSCP10 and pUSCP30 (Fig. 1) which contained the C gene with the back-mutated (wild-type) pre-C region, HBcAg/HBeAg was detected in both cell extracts and culture media. These results agreed with the reports of previous studies (Bruss & Gerlich, 1988; Ou et al., 1986; Roossinck et al., 1986). In the case of plasmids pUSCP1, pUSCP3 and pUSC, HBcAg/HBeAg was detected only in cell extracts but not in the culture media (Table 1). These results suggest that our mutant HBV cannot synthesize pre-C region-derived HBeAg. This observation was further confirmed by Western blotting experiments.

Characterization of core antigens by Western blotting (Sambrook et al., 1989) was performed to estimate the relative size of the processed antigens. We expressed the C gene in Escherichia coli (Choi et al., 1991). The recombinant core particles produced by E. coli cells were purified (Kim & Rho, 1992) and used as an M protein marker. A rabbit polyclonal anti-HBc antiserum against recombinant HBcAg was prepared using techniques reported previously (Sambrook et al., 1989). As shown in Fig. 2(a), COS cells transfected with pUSCP1, pUSCP3 and pUSC produced only a 21K protein (P21c), whereas COS cells transfected with pUSCP10 produced a minor 22K protein (P22) as well as a 21K protein (P21c). After deletion of the ATG codon region of the elongated pre-C (pUSCP30), the 22K protein became the major species. The 22K protein (P22) then becomes cleaved by a protease(s) and secreted into the medium (Ou et al., 1986, 1988; Salfeld et al., 1989). The production of P22 from pUSCP10 was much lower compared with that from pUSCP30 (Table 1 and Fig. 2). This result suggests that the AUG codon of the elongated pre-C may interfere with translational initiation at the AUG codon of the C protein. We have also established a mutant-HBV-producing cell line, HepG2-K8, by transfection of mutant HBV DNA into HepG2 cells (Kim et al., 1992a). In HepG2-K8 cells, only the 21K protein was detected (Fig. 2a). These results confirm that our mutant HBV has a defective pre-C region, and hence cannot synthesize HBcAg/HBeAg.

![Fig. 2. Western blotting analysis of HBV antigens. Extracts of COS cells transfected with recombinant plasmids and HepG2 cells producing the mutant HBV particles were separated by SDS-PAGE and transferred to a nitrocellulose filter. Core proteins (P21c and P22) (a) and X-C fusion protein (b) were detected by an anti-HBcAg antiserum and 125I-labelled Protein A. The COS control refers to pUC19-transfected COS cells. M values are shown on the right. M and D indicate the positions of the monomer (21K) and dimer (42K) forms of P21c purified from E. coli.](image-url)
the following references: HBV viruses. Numbers indicate map positions. Sequences were taken from immunoassay results (Table 1). This implies that the mutant HBV did not utilize the AUG codon of the elongated pre-C region for HBcAg and HBeAg synthesis. However, if initiation from the AUG codon of the elongated pre-C region is not essential for virus replication. Introduction of a frameshift and nonsense mutations in the duck hepatitis B virus (DHBV) pre-C region has only a pre-C/C ORF. Furthermore, the 5' end of the DHBV C gene shows no homology with the X gene of mammalian hepadnaviruses (Feitelson & Miller, 1988). Therefore, the DHBV C gene cannot be considered as a fusion between the C and X genes.

To determine whether an X-C fusion protein is synthesized, we constructed the plasmid pSVXC which contained a fused X-C ORF under the control of the SV40 promoter. Extracts of COS cells transfected with pSVXC contained HBcAg/HBeAg activity (Table 1). In RNA analysis experiments, a 1.9 kb transcript which had the fused X-C ORF coding sequences was detected (data not shown). However, a transcript of approx. 1 kb which encodes the C protein, transcribed from the initiation site of C mRNA, was not detected. Approximately 70% of the transcripts from a region near the initiation site of the X mRNA did not stop at the HBV transcription termination signal (data not shown). This value was higher than that in a previous report (Guo et al., 1991) which showed that 50% of X gene transcripts could read through the transcription termination signal. As shown in Fig. 2(b), COS cells transfected by pSVXC produced a 40K X-C fusion protein. Bands above the 40K band appear to be due to proteins interacting non-specifically with the anti-HBcAg antiserum. The fused X-C ORF could encode an X-C fusion protein of 359 amino acids, consisting of 151 amino acids encoded by the wild-type X ORF and 208 amino acids encoded by the pre-C/C ORF region.

Previous results have shown that the carboxyl-terminal amino acids of X protein are dispensable for its function as a transcriptional trans-activator (Levrero et al., 1990; Ritter et al., 1991); we therefore examined the trans-activating activity of the X-C fusion protein. The reporter plasmid pSV2CAT containing the SV40 enhancer and early promoter driving the chloramphenicol acetyltransferase (CAT) gene was cotransfected with test plasmids into HepG2 cells, and a CAT assay was then performed (Gorman et al., 1982). Plasmid pHX contained the wild-type X gene (Won et al., 1989) under the control of the HBV enhancer/X promoter and the poly(A) signal of the herpes simplex virus thymidine kinase gene (HSV TK). Plasmid pHXC, on the other hand, was similar to pHX but contained the fused X-C ORF instead of the X gene (Fig. 4). Plasmid pHXCM is similar to pHXC except that the poly(A) signal sequence UAAUAAA was mutated to UAACAG. In the case of plasmid pHXC, the poly(A) signal is located inside the C gene. To exclude the effect of the TATAAA sequence when it encounters the transcription termination signal,
Fig. 4 Transcriptional trans-activation of SV40 enhancer/promoter by the fused X-C ORF product. (a) Schematic maps of test plasmids pHX, pHXC and pHXCM, the correct constructions of which were confirmed by DNA sequence analysis. Numbers in parentheses indicate map positions on the HBV genome. E/P, HBV enhancer/X promoter; open boxes, ORFs; open circles, HBV poly(A) signal; solid circles, herpes simplex virus TK poly(A) signal; star, mutated HBV poly(A) signal. (b) CAT assay for the trans-activation function of the X-C fusion protein. HepG2 cells were cotransfected with 5 µg of pSV2CAT and 15 µg of each of the plasmids pUC19, pHX, pHXCM or pHXC. Activation (fold) represents the ratio of the percentage of chloramphenicol acetylation in cotransfection with the X and X-C gene expression vectors to that with control plasmid pUC19. All values are given as the mean ± s.d. obtained from experiments carried out in triplicate.


mammalian cells by plasmid DNA. *Molecular and Cellular Biology* 7, 2745–2752.


Tong, S. P., Doun, C., Grison, P., Li, J. S., Vitytiski, L., Trepo, C. &


(Received 19 July 1993; Accepted 26 October 1993)