Hepatitis B virus polymerase gene: expression of the long open reading frame using the baculovirus expression system

E. McGlynn, S. Reutener, A. Matter, G. Wildner, H. Will and N. B. Lydon

1Research Department, Pharmaceuticals Division, Ciba-Geigy Limited, K-125.4.12, CH-4002 Basel, Switzerland and
2Max-Planck Institut für Biochemie, 8033 Martinsried, Germany

A recombinant baculovirus was constructed containing a copy of the hepatitis B virus (HBV) genome which was inserted to produce an in-frame fusion of the precore (pre-C) coding region with the first 11 amino acids of the polyhedrin gene. The recombinant baculovirus expressed the 25K pre-C protein and two novel proteins, of approximately 93K and 72K. Both the 93K and 72K proteins are recognized by an anti-polymerase monoclonal antibody. Northern blot analysis of the mRNA produced during infection of Spodoptera frugiperda cells by the HBV recombinant baculovirus detected only one HBV mRNA species, suggesting that the three HBV-specific proteins expressed are translated from the same mRNA. No larger fusion proteins cross-reacting with either anti-core or polymerase antibodies were detected. These findings suggest that the two proteins encoded within the HBV polymerase gene are not produced via a core–polymerase fusion intermediate but by internal binding of ribosomes. These results are the first clear demonstration of efficient expression of two bona fide unprocessed polymerase proteins in a 1:1 ratio from an unspliced pre-C mRNA-like transcript. With the successful expression of the polymerase gene in insect cells it is now possible to produce large amounts of these proteins, allowing a more detailed structural and functional analysis of these proteins.

The human hepatitis B virus (HBV) is the prototype of the Hepadnaviridae, a family of viruses which exhibit a narrow host range and closely resemble each other in genomic organization. The genome of HBV is a small, partially dsDNA molecule. DNA-dependent DNA polymerase activity, which can repair the single-stranded gap in the HBV DNA molecule, has been found associated with HBV particles (Kaplan et al., 1973; Robinson, 1975). This enzyme has not yet been isolated in a functional form, making comparative biochemical studies of the HBV and mammalian polymerases difficult (Aden et al., 1979; Hirschmann & Garfinkel, 1977; Hess et al., 1980; Goto et al., 1984).

Following the cloning of HBV DNA, sequence analysis revealed that the viral genome is composed of four overlapping open reading frames (ORFs) (Galibert et al., 1979; Pasek et al., 1979). It has been assumed that the long ORF encodes the viral DNA polymerase. It has the capacity to encode a protein of 93K which is in the size range of known viral DNA polymerases (Kornberg, 1980) and contains regions which resemble those of retroviral reverse transcriptases including the consensus sequence of RNase H (Toh et al., 1983; Wintersberger, 1990). The Hepadnaviridae replicate the negative DNA strand of their genomes via an RNA intermediate using reverse transcriptase activity (Summers & Mason, 1982). The similarity to reverse transcriptases and the predicted size of the gene product have reinforced the assignment of the long ORF as the polymerase gene.

No mRNA capable of exclusively encoding the HBV polymerase gene product has been detected. However, such a message may be present in very low amounts, being transiently expressed for a limited period during HBV infection. Alternatively, the polymerase could be encoded by one of the genomic size messages which encode the core and precore (pre-C) gene products. The organization of the ORFs in HBV is similar to that in several retroviruses and retroid elements (Fuetterer & Hohn, 1987). In particular the 3' end of the core gene overlaps the 5' end of the proposed polymerase gene, an arrangement analogous to the gag polymerase of Rous sarcoma virus and some other retroviruses. These ORFs differ in that the postulated polymerase gene of HBV is in the + 1 frame with respect to the core gene, whereas the polymerase genes of retroviruses are in the - 1 frame with respect to the gag gene. A number of mechanisms could account for expression of the HBV polymerase gene. Firstly, the polymerase gene product could be expressed by frameshifting either as a pre-C–polymerase or core–polymerase fusion protein, which is then
proteolytically processed to liberate the polymerase (in a manner similar to that of retroviruses). Alternatively, expression of the polymerase gene could result from internal initiation of translation. Although parts of the polymerase ORF have been expressed in heterologous systems (Stemler et al., 1988; Weimer et al., 1989; Jean-Jean et al., 1989; Chang et al., 1989a), the expression of the full-length polymerase gene products have so far not been demonstrated biochemically or immunologically. As a result, analysis of the polymerase gene product has been severely hampered. We have attempted to overcome these problems by expressing the HBV polymerase gene in insect cells using the baculovirus expression system. The present report describes the production of two proteins encoded within the HBV polymerase ORF which are expressed from a pre-C-like mRNA using a recombinant baculovirus.

The plasmid pHBV130.4 (Gough & Murray, 1982) which carries four genomic copies of HBV was digested with FspI. This digestion released a fragment containing almost the complete HBV genome (nucleotides 3096 to 3083; nucleotide positions are derived from the DNA sequences reported in the GenBank sequence databank) which is referred to as the pre-C gene (Fig. 1). The baculovirus transfer vector pAc360, designed to facilitate the production of fusion proteins between a foreign gene and the first 11 N-terminal codons of the polyhedrin gene, was cleaved with BamHI. The 5' overhangs were repaired with the Klenow enzyme and then ligated to the FspI fragment carrying the pre-C gene. The resulting plasmid, pPrecore, was sequenced to confirm that the pre-C gene was in-frame with the polyhedrin ATG. The 5' overhangs were repaired with the Klenow enzyme and then ligated to the FspI fragment carrying the pre-C gene. The resulting plasmid, pPrecore, was sequenced to confirm that the pre-C gene was in-frame with the polyhedrin ATG. Spodoptera frugiperda (Sf9) cells were cotransfected with the plasmid pPrecore and Autographa californica nuclear polyhedrosis virus (AcNPV) DNA (Luckow & Summers, 1988). Homologous recombination between AcNPV sequences present in the transfer vectors and viral DNA gave rise to recombinant viruses that were identified by dot blot analyses (Fung et al., 1988). A recombinant virus (vPrecore1) was selected for further study.

The recombinant baculovirus vPrecore1 and AcNPV were propagated in Sf9 cells. After 72 h viral DNA was extracted as described (Luckow & Summers, 1988) and digested with EcoRV and KpnI. Southern blotting using an HBV-specific probe revealed a single fragment of approximately 3.6 kb (3.1 kb of HBV DNA and 0.5 kb of flanking sequence), indicating that the recombinant virus contained a single intact copy of the pre-C gene (Fig. 2a).

To determine the size of the HBV-specific RNAs, total cellular RNA from vPrecore1 virus-infected Sf9 cells was hybridized with a labelled HBV-specific probe (Luckow & Summers, 1988). An RNA molecule of approximately 4 kb (Fig. 2b), the expected size of the combined pre-C gene and polyhedrin mRNA was found; transcription started at nucleotide – 48 in the polyhedrin gene (Howard et al., 1986).

SF9 monolayers infected with the recombinant vPrecore1 virus were harvested at various times after infection and analysed using SDS–PAGE. A novel protein with an Mr of approximately 25K, which represented approximately 5% of total cell protein (Fig. 3a), was expressed in cells infected with vPrecore1 virus. Peptides from infected cell extracts were fractionated by SDS–PAGE and transferred to nitrocellulose. The membranes were probed with an anti-HBV core antigen monoclonal antibody (MAb) which reacted with the novel 25K protein, confirming its status as the pre-C protein (Fig. 3b). No larger proteins were recognized by the anti-core MAb. A duplicate Western blot developed using an anti-HBV polymerase peptide MAb revealed two proteins with approximate Mr's of 93K and 72K (Fig. 4). These two proteins were not detected with MAb specific for HBV core, surface or X proteins. The anti-polymerase antibody was raised against a bacterially expressed recombinant polymerase fusion protein (Will et al., 1986) and recognized the amino terminus (amino acids 28 to 199) of the polymerase protein. This was determined by immunoprecipitation of the protein products of in vitro transcription/translation experiments and by immunoblotting using recombinant MS2–poly-

\[ \text{HBV genome} \]

\[ \text{pPrecore} \]

\[ \text{25K} \]

\[ \text{93K} \]

\[ \text{72K} \]

Fig. 1. (a) Schematic representation of the HBV genome showing the positions of the direct repeats (Dr), precore (PC), core (C), polymerase (P), surface antigen (S) and X genes. (b) Insert cloned in plasmid pPrecore. (c) Predicted size of encoded proteins from HBV pre-C and polymerase ORFs.
merase fusion proteins (G. Wildner & H. Will, unpublished results).

Thus, a recombinant baculovirus has been constructed containing a copy of the HBV genome. It was inserted to produce an in-frame fusion of the pre-C coding region with the first 11 amino acids of the polyhedrin gene. The recombinant baculovirus expressed the 25K pre-C protein and two novel proteins, of approximately 93K and 72K. The pre-C antigen when expressed in insect cells was not properly processed, a result which has also been demonstrated by Lanford & Notval (1990). The 93K and 72K proteins were recognized by an HBV polymerase-specific antibody. However, antibodies directed against other HBV proteins (surface, core and X) did not detect either 93K or 72K protein. The polymerase reading frame has the coding capacity for a protein of 93K. The larger protein probably represents the full-length product of the HBV polymerase gene, whereas the smaller protein may represent a cleavage product of the larger polypeptide or result from translation initiation from the internal ATG codon located 333 nucleotides downstream from the polymerase gene ATG (Fig. 1c). Polymerase-specific proteins of similar size have also been detected in HBV virion particles using the Western blotting technique (Mack et al., 1988; Bavand & Laub, 1988). It has been proposed that the 93K protein is the terminal genome-linked protein used to prime reverse transcription (Bartenschlager & Schaller, 1988), whereas
Fig. 4. Western blot analysis of vPrecore1 virus-infected insect cells. Cells were harvested 48 h after infection and lysed in SDS sample buffer. Proteins were subsequently fractionated by SDS-PAGE and transferred to nitrocellulose. Membranes were incubated with an anti-HBV polymerase peptide MAb. Following washing in PBS/Tween, membranes were incubated with peroxidase-conjugated secondary antibody and finally developed using the Bio-Rad peroxidase substrate kit. Lane 1, vPrecore1 virus-infected Sf9 cells; lane 2, AcNPV-infected cells.

the 72K protein is possibly the reverse transcriptase (Jean-Jean et al., 1989).

The recombinant baculovirus construct described in this paper was designed with the polymerase initiation codon being the fourth ATG from the 5' end of the mRNA. Exhaustive studies were performed in an attempt to detect a fusion protein between the pre-C and polymerase gene products. No larger proteins reactive with anti-core antibodies were detected in extracts from cells infected with the recombinant vPrecore1 baculovirus. Northern blot analysis of the mRNA produced during infection of Sf9 insect cells by the HBV recombinant baculovirus detected only one HBV mRNA species suggesting that the three HBV-specific proteins expressed are translated from the same mRNA. These findings, in agreement with gene mutation and complementation studies involving mutation of the sequences located in the overlap between the core–polymerase genes of duck hepatitis B virus, indicate that the polymerase gene is expressed by internal initiation and not via a core–polymerase fusion intermediate (Schlicht et al., 1989; Chang et al., 1989b, 1990). Other studies involving mutational analysis of the HBV polymerase gene, transfection of the mutated pregenome and analysis of the effect on DNA/RNA intermediates produced during replication (Okamoto et al., 1990; Radziwill et al., 1990; Wu et al., 1991; Roychoudhury & Shih, 1990) also support the idea that the polymerase gene is expressed by internal initiation. These studies suggest that internal initiation of the polymerase gene is required for expression of the polymerase gene and not a pre-C–polymerase fusion intermediate. This was further supported by transient assay experiments and in vitro transcription/translation experiments (Jean-Jean et al., 1989; Ou et al., 1990).

As both the 93K and 72K polymerase-specific proteins are expressed in insect cells, regulation of expression of these proteins must be independent of a hepatic element or other virus-encoded polypeptides produced during the HBV infectious cycle.

It has been shown for picornaviruses that initiation of translation occurs at internal ATGs via internal ribosome binding as opposed to that involved in Kozak’s scanning model (Kozak, 1986; Pelletier & Sonenberg, 1988; Jang et al., 1989). Such a mechanism of translation initiation may account for the production of both the 93K and 72K proteins. It is also possible that production of the latter results from scanning of ribosomes that fail to initiate at the AUG of the 93K protein, but which initiate at the AUG codon lying 333 nucleotides downstream. One other possibility is that the 72K protein represents a degradation or specific cleavage product of the 93K protein.

In our studies we show that efficient polymerase gene expression is observed from a pre-C-like mRNA which is in contrast to previously published evidence, in which polymerase protein translation from pre-C mRNA was observed at low efficiency (Ou et al., 1990). Interestingly we have shown that the 93K and 72K polymerase gene products are relatively stable in vivo unlike that shown in previous experiments using other expression systems (Ou et al., 1990). This may indicate that the polymerase gene products are protected in some way, possibly by encapsidation into pre-C-like particles (Miyanohara et al., 1986). With the successful expression of the polymerase gene products in insect cells it is now possible to produce sufficient amounts of these proteins for biochemical analysis.

We would like to thank Professor K. Murray for kindly providing plasmid pHBV130.4. We would also like to thank Professor D. Bishop and Drs M. Bailey and K. Takehara for help in establishing the
baculovirus expression system. Part of this work was supported by a grant from the Deutsche Forschungsgemeinschaft (Forschergruppe Fa138/3-3 and SFB 2170).

References


Received 18 October 1991; Accepted 17 February 1992.