The EMBO Journal vol.4 no.1 pp.249–255, 1985

Co-transcribed 3′ host sequences augment expression of integrated hepatitis B virus DNA

Arndt Freytag von Loringhoven, Stefan Koch, Peter Hans Hofschneider and Rajen Koshy[1]

Max-Planck-Institut für Biochemie, Am Klopferspitz 13, D-8033 Martinsried, FRG

[1]Present address: Department of Medicine, Health Sciences Center, T-17 080, State University of New York at Stony Brook, NY 11794-8174, USA

Communicated by P.H.Hofschneider

We have previously reported the cloning and structural analysis of integrated hepatitis B virus DNA copies from the human hepatoma cell line PLC/PRF/5. Here we show that the cloned DNA fragments of 10.7 kb and 10.5 kb contain intact coding sequences for HBsAg since Ltk$^-$ cells transfected with these DNAs secrete considerable amounts of HBsAg. We show for the 10.7-kb fragment that multiple readthrough messages composed of viral as well as cellular sequences are transcribed. These RNAs differ only in their 3′ sequences. Furthermore, the 10.7-kb insert leads to a substantial increase in HBsAg produced compared with HBV DNA and with the 10.5-kb insert. We provide evidence that the different 3′ sequences on the HBsAg transcripts account for the augmentation of expression.

Key words: hepatoma cell line/integrated hepatitis B virus DNA/HBV surface antigen/hybrid host-virus mRNA/increased expression

Introduction

Hepatitis B virus (HBV) infects human and primate hepatocytes and causes hepatitis and probably primary liver carcinoma (PLC) in humans (Szmuness, 1978). We have previously reported (Koshy *et al.*, 1983; Koch *et al.*, 1984b) the cloning of different human DNA sequences with integrated HBV DNA from PLC/PRF/5 cells (Alexander *et al.*, 1976), which were derived from human primary liver carcinoma tissue. The cloned inserts were characterized by restriction mapping and partially by nucleotide sequencing (Koshy *et al.*, 1983; Koch *et al.*, 1984a, 1984b). Two of the fragments – designated A-10.7 and A-10.5 – contain cellular flanking sequences only on one side of the integrated viral DNA (Figure 1). On the basis of restriction mapping and nucleotide sequencing it appears that in A-10.7 and A-10.5 the HBV DNA sequences as well as 187 nucleotides – with the exception of three bases – of adjacent human flanking DNA are identical, but the remainder of the cellular sequences are unrelated (see Figure 1; Koch *et al.*, 1984a). The viral sequences are co-linear with the published HBV genome (Valenzuela *et al.*, 1980).

Both A-10.7 and A-10.5 contain the HBV surface antigen (HBsAg) gene (S gene) and two sequences implicated as the S gene promoter, a 'TATA box' preceding the pre-S gene (Pourcel *et al.*, 1982; Gough, 1983; Rall *et al.*, 1983) and a region homologous to the SV40 late promoter sequence in the pre-S gene (Cattaneo *et al.*, 1983). However, integration was such that both inserts lack the viral S gene termination region downstream of the putative S gene polyadenylation signal TATAAA as well as the viral poly(A) signal itself, which normally lies 1083 bp 3′ of the S gene (Figure 1; Pourcel *et al.*, 1982; Gough, 1983; Cattaneo *et al.*, 1983).

To study the functional properties of the integrated viral sequences and their possible interaction with flanking host DNA, the cloned DNAs were transfected into mouse Ltk$^-$ cells. The S gene transcription and translation products are characterized here in detail. Compared with the unintegrated viral S gene, both cloned fragments lead to significantly augmented HBsAg expression. This can be ascribed to distal human sequences co-transcribed on hybrid host-virus mRNAs.

Results

HBsAg secretion by transfected Ltk cells

To investigate whether HBsAg can be expressed despite the loss of viral regulatory signals at the 3′ end of the HBsAg transcription unit, lambda A1 10.7 and lambda A1 10.5 (lambda recombinants containing A-10.7 and A-10.5, respectively) were co-transfected with pFG5 (Colbere-Garapin *et al.*, 1979), containing the thymidine kinase (*tk*) gene of herpes simplex virus I, into mouse Ltk$^-$ cells by the calcium phosphate precipitation procedure (Graham and van der Eb, 1973; Pellicer *et al.*, 1978). Cells were grown in selective medium containing hypoxanthine, aminopterin and thymidine (HAT). Two weeks after transfection colonies were pick-

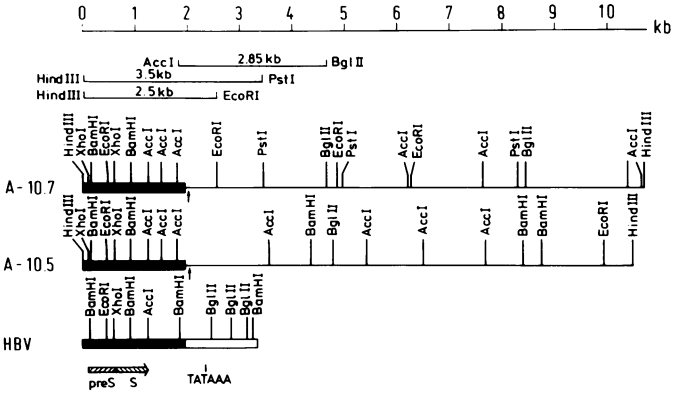

Fig. 1. Restriction maps of the cloned DNA fragments from PLC/PRF/5 cells. Cleavage sites are shown for the inserts A-10.7, A-10.5 and linearized HBV DNA (subtype ayw) for comparison. The thin lines represent human DNA, solid boxes HBV DNA and the open box HBV DNA not present in A-10.7 or A-10.5. The viral sequences in A-10.7 and A-10.5 were truncated by the cloning procedure due to an unexpected *Hind*III site within viral DNA (Koshy *et al.*, 1983). The thick arrow at the bottom marks the location of the pre-S and S gene. The position of the viral S gene poly(A) signal (TATAAA) is indicated. In A-10.7 and A-10.5, 187 nucleotides of human flanking sequences – from the host-virus junction to the positions marked by the arrows – are identical. The following subfragments of A-10.7 are shown at the top: *Hind*III-*Eco*RI (2.5 kb) and *Hind*III-*Pst*I (3.5 kb); both were used to map the region responsible for the augmented HBsAg expression; the *Acc*I-*Bgl*II fragment (2.85 kb) was ligated to the herpes simplex virus I *tk* gene to check for enhancer activity.

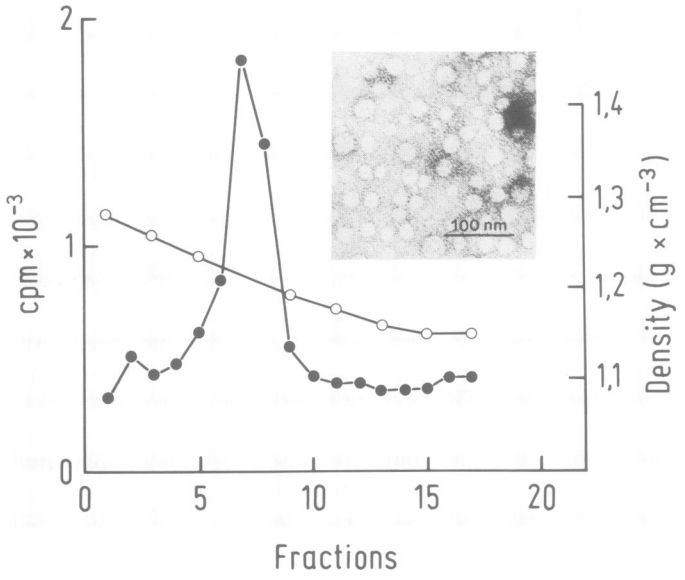

Fig. 2. CsCl gradient sedimentation profile and electron micrograph of HBsAg particles from supernatant medium of A-10.7 transfected cells. Medium from ~1.5 x 10^8 Ltk 2425 cells – transfected with pA1 10.7 (A-10.7 in pBR327) – was centrifuged at 15 000 *g* for 24 h. The pellet was resuspended in 1 ml of medium and submitted to CsCl density gradient centrifugation (1.21 g/ml; 45 000 r.p.m. at 17°C for 60 h in a Beckman SW 41 rotor). Fractions were collected from the bottom, diluted 1:100 and HBsAg measured using the Abbott AUSRIA II radioimmunoassay (solid circles). Open circles represent fraction densities. The peak fraction was submitted to a second, identical CsCl gradient, dialysed and concentrated by precipitation with 10% polyethyleneglycol 6000, adsorbed on carbon-coated grids and particles visualized by staining with 2% uranyl formeate.

ed and transferred to culture wells.

Supernatant media from cells in culture wells were tested for HBsAg by radioimmunoassay (AUSRIA II, Abbott). Sixteen of 18 and four of 12 colonies transfected with lambda A1 10.7 DNA and lambda A1 10.5 DNA, respectively, secreted HBsAg (see below). Thus A-10.7 and A-10.5 are able to express HBsAg despite the absence of the viral S gene termination and polyadenylation signals in both cases.

To characterize the HBsAg product, supernatant medium from one of the Ltk colonies (Ltk 2425) was submitted to CsCl density gradient centrifugation. When fractions were analyzed for HBsAg by radioimmunoassay a peak was seen at 1.21 g/cm³ (Figure 2). This is typical for 22 nm HBsAg particles, which are frequently found in infected human serum (Gerin *et al.*, 1969). Spherical particles of 20–25 nm were detected by electron microscopy (Figure 2) after further purification. ^{35}S-labeled polypeptides were enriched and examined by SDS-polyacrylamide gel electrophoresis. The presence of bands of 24 000 and 28 000 daltons (Figure 3a) is in agreement with other reports (Stibbe and Gerlich, 1982). When the same sample was immunoprecipitated with a human anti-HBsAg positive serum (obtained from the Max-von-Pettenkofer-Institut, Munich) again polypeptides of the same size are present (Figure 3b). Thus by all criteria applied the HBsAg secreted by Ltk 2425 cells is indistinguishable from that secreted by HBV-infected human hepatocytes.

Cell line Ltk 2425 was also used to quantify HBsAg expression. Up to 1.2 μg/10^7 cells/24 h are produced (for details see Freytag von Loringhoven, 1984) which is in the range of HBsAg expression under the control of retroviral LTR sequences (Stratowa *et al.*, 1982).

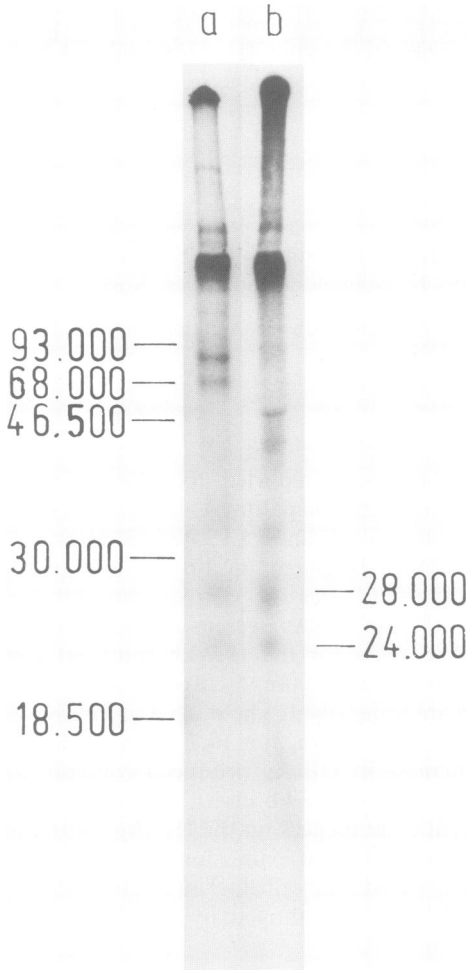

Fig. 3. SDS-PAGE analysis of HBsAg polypeptides. [^{35}S]Cysteine-labeled proteins (10^6 c.p.m.) from supernatant medium of Ltk 2425 cells were precipitated with 10% polyethyleneglycol 6000 and analyzed by electrophoresis through a 13% polyacrylamide gel **(lane a)**. The same sample was immunoprecipitated with 10 μl of anti-HBs containing human serum **(lane b)**. Proteins of 24 000 and 28 000 daltons are present in both **lanes a** and **b**. The positions of ^{14}C-labeled protein standards are also indicated.

Comparison of HBsAg expression levels directed by A-10.7, A-10.5 and HBV DNA

The level of HBsAg expression in individual colonies varied as expected (Table I), presumably reflecting a combination of factors including different DNA copy number, influence of flanking Ltk DNA and chromatin structure. Surprisingly, however, the average level of HBsAg expression in lambda A1 10.7-derived colonies was eight times higher than that in lambda A1 10.5-derived colonies. This was corroborated by further analysis of five representative colonies from each series. Cells were grown to mass culture, their HBsAg production and cell number quantitated and the copy number determined by dot blotting (see Freytag von Loringhoven, 1984). While the number of integrated copies was between one and three per genome and on average equal in lambda A1 10.7 and lambda A1 10.5 colonies, the average level of HBsAg produced per 10^7 cells per day was again eight times higher in lambda A1 10.7 than in lambda A1 10.5 colonies (see Freytag von Loringhoven, 1984). These data exclude the possibility that the difference in expression of HBsAg is due

Table I. Comparison of HBsAg levels secreted by Ltk colonies after transfection with lambda A1 10.7 and lambda A1 10.5

Cell clone	λA1 10.7 P/N	λA1 10.5 P/N
1	2.1	10.6
2	12.3	6.2
3	114.1	11.6
4	65.2	(2.0)
5	3.1	2.4
6	5.5	–
7	140.2	–
8	15.8	–
9	8.3	–
10	9.5	–
11	28.6	–
12	15.1	–
13	24.2	
14	14.9	
15	30.9	
16	23.7	
17	–	
18	–	
Mean	28.7	3.6

5 μg of lambda A1 10.7 or lambda A1 10.5 DNA were mixed with 50 ng of pFG5 (containing the HSV *tk* gene) and transfected into Ltk$^-$ cells by the calcium phosphate method. Colonies were picked 14 days later and transferred to culture wells (Costar). When cells were confluent the supernatant media (1 ml) of 48 h were analyzed for HBsAg by radioimmunoassay (Abbott AUSRIA II). Measurements are expressed as P/N. P/N ratios lower than 2.0 are indicated as negative values.

Table II. Comparison of HBsAg levels secreted by Ltk colonies after transfection with pKK HBs 34, pA1 10.7 and two pA1 10.7 subfragments

Cell clone	pKK HBs 34 P/N	pA1 10.7 P/N	pA1 10.7 (H-E) P/N	pA1 10.7 (H-P) P/N
1	2.2	445	213	104
2	2.1	407	27.6	7.7
3	11.5	33.1	18.4	9
4	6.5	8.3	16.2	5
5	21.8	–	9.3	91
6	–	–	2.6	32
7	–	–	97	53
8	–		2.5	440
9	–		27.2	–
10	–		544	
11	–		97	
12	–		454	
13	–		10.3	
14	–		–	
15	–			
16	–			
Mean	3.8	128	108.6	82.6

Equimolar amounts of pKK HBs 34 (2.8 μg), pA1 10.7 (5 μg), pA1 10.7 (H-E, the 2.5-kb *Hind*III-*Eco*RI fragment cloned in pBR327) (1.9 μg) and the 3.5-kb *Hind*III-*Pst*I fragment (H-P) of pA1 10.7 (1.1 μg) were precipitated with 50 ng of pFG5 (contains the HSV *tk* gene). All plasmids were previously digested to completion with *Hind*III. pA1 10.7 subfragments are also indicated in Figure 1. Colonies were picked 14 days later and transferred to culture wells (Costar). When cells were confluent the supernatant media (1 ml) of 24 h were analyzed for HBsAg by radioimmunoassay (Abbott AUSRIA II). Measurements are expressed as P/N.

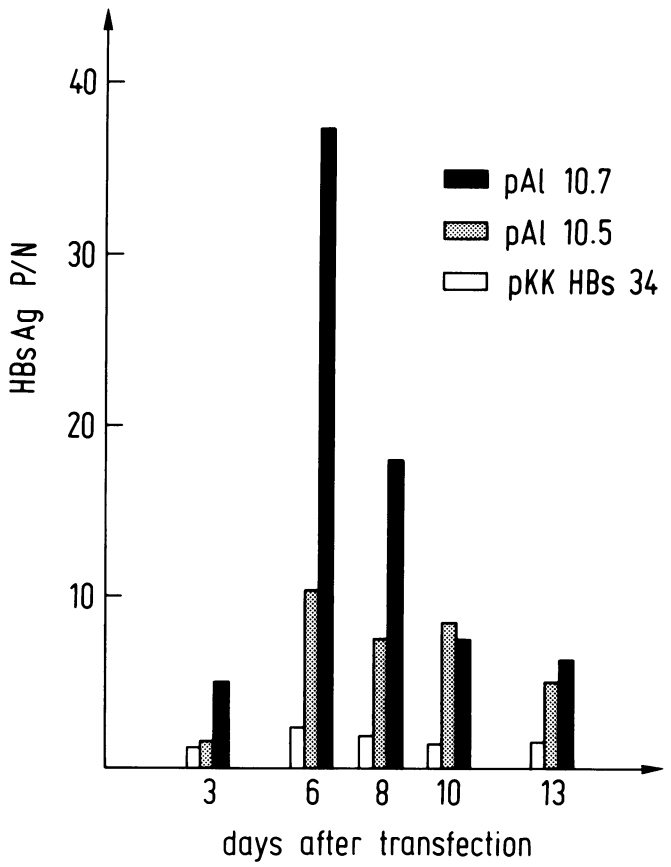

Fig. 4. Comparison of HBsAg levels secreted by unselected Ltk cell cultures. Equimolar amounts of *Hind*III-digested pA1 10.7 (3 μg), pA1 10.5 (2.55 μg) and pKK HBs 34 (1.65 μg) were transfected into Ltk$^-$ cells. Cells were maintained as monolayers and supernatants were assayed for HBsAg by radioimmunoassay on the indicated days. Results are the averages of duplicate experiments.

to different copy numbers of cloned inserts in transfected Ltk cells.

In a second transfection experiment the HBsAg expression by pA1 10.7 – pBR327 containing A-10.7 (Koshy *et al.*, 1983) – was compared with that of cloned HBV DNA of the same subtype (adw$_2$; Valenzuela *et al.*, 1980; Koch *et al.*, 1984b). Plasmid pKK HBs 34, which is composed of the 2.8-kb *Bgl*II-*Bgl*II HBV DNA fragment containing the functional S gene and the pBR322-derived vector pKK92-c (Wang *et al.*, 1982), was used. Equimolar amounts of each DNA were transfected along with *tk* DNA into Ltk$^-$ cells as described above. When supernatant media from cell colonies were assayed for HBsAg the average level of HBsAg expression of pA1 10.7-derived colonies was 33 times higher than that of pKK HBs 34-derived colonies (Table II).

The levels of HBsAg expression of A-10.7, A-10.5 and HBV DNA were also compared in unselected cell culture (Figure 4). Equimolar amounts of pA1 10.7, pA1 10.5 and pKK HBs 34 were individually transfected into Ltk$^-$ cells and supernatants assayed for HBsAg over a period of 13 days. Under these conditions the possibility is minimized that the expression levels are determined by the integration position in the Ltk genome, since they are compared early after transfection. Furthermore, a far higher amount of transfected inserts can be screened than in *tk*-selected colonies. Here again, the

same pattern is evident as in the experiment comparing tk^+ colonies. A-10.7 was clearly more efficient in directing HBsAg expression than A-10.5, which in turn was more efficient than HBV DNA (day 6: $P/N_{A\text{-}10.7} = 37.3$, $P/N_{A\text{-}10.5} = 10.4$, $P/N_{HBV\ DNA} = 2.4$). The differences, however, were somewhat less pronounced than in *tk*-selected colonies.

Mapping of the DNA region responsible for the high HBsAg expression

The integrated HBV sequences as well as 187 nucleotides of adjacent human DNA – with the exception of three bases – are identical in A-10.7 and A-10.5 (Koch *et al.*, 1984a). Therefore the DNA region responsible for the different expression of HBsAg cannot lie within the HBV-specific sequences, but must reside in cellular sequences at least 187 nucleotides away from the host-virus junction. To map this region more precisely two subfragments of A-10.7 were isolated and their expression compared with that of the complete A-10.7 as follows. (i) A 2.5-kb *Hind*III-*Eco*RI fragment (see Figure 1) containing all of the HBV sequences and 0.5 kb of flanking human DNA was subcloned into pBR327 (Soberon *et al.*, 1980) and linearized with *Hind*III. (ii) A 3.5-kb *Hind*III-*Pst*I fragment (see Figure 1) containing all of the HBV sequences and 1.5 kb of flanking DNA was purified. Equimolar amounts of these DNAs were included in the experiment mentioned above in which the expression of pA1 10.7 and pKK HBs 34 DNA were compared (Table II). The average level of surface antigen expression directed by the 2.5-kb *Hind*III-*Eco*RI subfragment of pA1 10.7 was P/N = 108.6, that of the 3.5-kb *Hind*III-*Pst*I fragment P/N = 82.6. Both differed only marginally from the mean expression of pA1 10.7 transfected colonies (P/N = 128; Table II). Therefore, the DNA region responsible for the differential HBsAg expression of A-10.7 and A-10.5 lies within a region of cellular DNA encompassing ~300 bp. This region extends from a position 187 nucleotides away from the host-virus junction (extent of identical sequences in A-10.7 and A-10.5) to the nearest *Eco*RI site in A-10.7.

To test whether the augmentation of HBsAg expression is due to an enhancer sequence situated in this region, a 2.85-kb *Acc*I-*Bgl*II fragment (containing the region mapped as being responsible for the high HBsAg expression; see Figure 1) from A-10.7 was ligated in the same (pF8) and in the opposite (pF3) orientation next to a *Bam*HI-*Bam*HI fragment containing the functional *tk* gene (for details see Freytag von Loringhoven, 1984). As these constructs were assayed in the same cell line (mouse Ltk) as all other experiments reported here, an enhancer potential should result in a large increase in *tk*-positive colonies after transfection (Jolly *et al.*, 1983; Hearing and Shenk, 1983). However, this was not the case. pF3 (1 colony/ng) and pF8 (0.89 colonies/ng) resulted in very similar colony numbers to pFG5 (1.17 colonies/ng) which does not contain A-10.7 sequences. Thus the presence of a cellular enhancer sequence responsible for the elevated HBsAg expression is unlikely.

Characterization of the HBsAg mRNA

Since the viral S gene polyadenylation and termination signals are not present in A-10.7 and A-10.5 the HBsAg expression described could be explained by assuming transcriptional readthrough into human flanking DNA and termination therein. Such hybrid host-virus transcripts may be exceptionally stable, thus accounting for the observed increase in HBsAg expression. The possibility of readthrough transcrip-

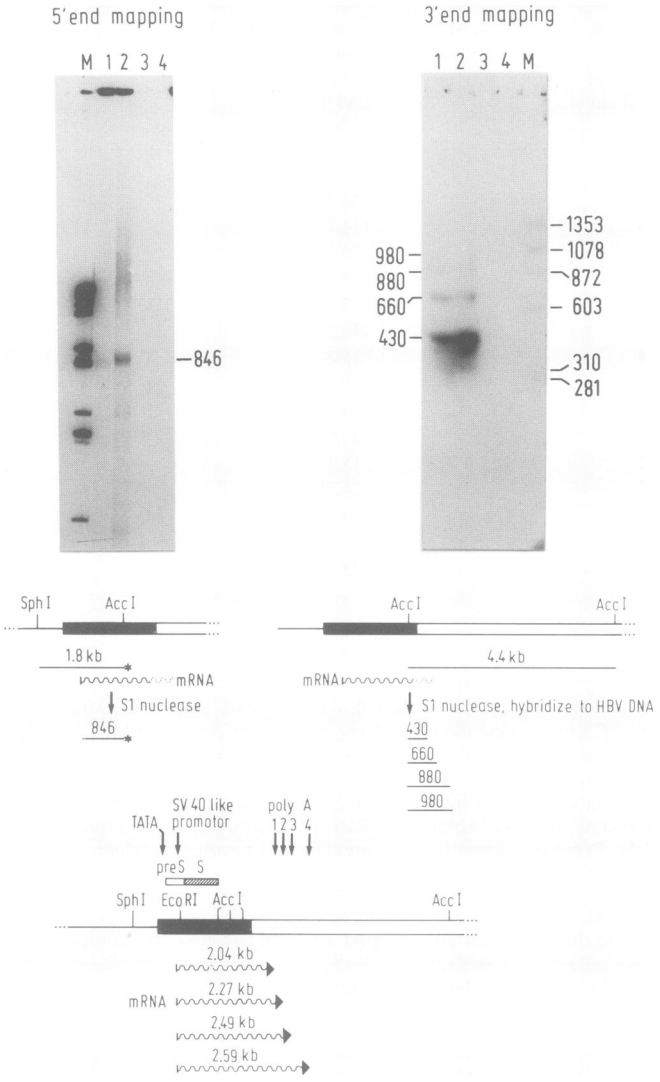

Fig. 5. Nuclease S1 mapping of the 5′ and 3′ ends of the HBsAg mRNA in A-10.7 transfected cells. 5′ end **(top left)**: a 1.9-kb *Acc*I-*Acc*I subfragment of pA1 10.7 was 5′ end-labeled and digested further with *Sph*I to give an asymmetrically labeled fragment of 1.8 kb. 50 000 c.p.m. of this DNA were hybridized to 30 μg of total RNA from Ltk 2425 cells or to 30 μg of yeast tRNA at 53°C for 16 h. Hybrids were treated with different amounts of S1 nuclease at 37°C for 30 min. After alkali hydrolysis of the RNA, protected fragments were electrophoresed through a 6% polyacrylamide sequencing gel. **Lanes 1, 2**, labeled DNA after hybridization to Ltk 2425 RNA and digestion with 30 and 10 units of S1 nuclease, respectively. **Lanes 3, 4**, labeled DNA after hybridization to yeast tRNA and digestion with 30 and 10 units of S1 nuclease, respectively. The ^{32}P-end-labeled marker DNA was a mixture of *Hind*III and *Eco*RI digested lambda DNA and *Hpa*II digested pBR322. Underneath pA1 10.7 (straight line, plasmid DNA; filled box, HBV DNA; open box, human DNA), the hybrid of mRNA and the end-labeled *Acc*I-*Sph*I fragment and the S1 protected fragment are schematically depicted. 3′ end **(top right)**: an unlabeled 4.4-kb *Acc*I-*Acc*I subfragment of pA1 10.7 was hybridized to 30 μg of total Ltk 2425 RNA or yeast tRNA at 59°C for 16 h. Hybrids were treated with 10 or 30 units of S1 nuclease at 37°C for 30 min and the RNA hydrolysed with alkali. Protected fragments were separated on a 2% agarose gel, blotted onto nitrocellulose filter and hybridized to a ^{32}P-labeled HBV-DNA probe. **Lanes 1, 2**, *Acc*I-*Acc*I fragment hybridized to Ltk 2425 RNA and digested with 10 and 30 units of S1 nuclease, respectively. **Lanes 3, 4**, *Acc*I-*Acc*I fragment hybridized to 30 μg yeast tRNA and digested with 10 and 30 units of S1 nuclease, respectively. *Hae*III-digested φX174 was used as mol. wt. marker. Below a scheme of the experiment is given. A summary of the results is provided at the bottom of the figure. An excerpt of pA1 10.7 and the positions of the pre-S and S gene as well as the two promoters (TATA and SV40 late-like) and the four poly(A) signals in the human DNA are shown. Below the locations of the S1 mapped RNAs are given.

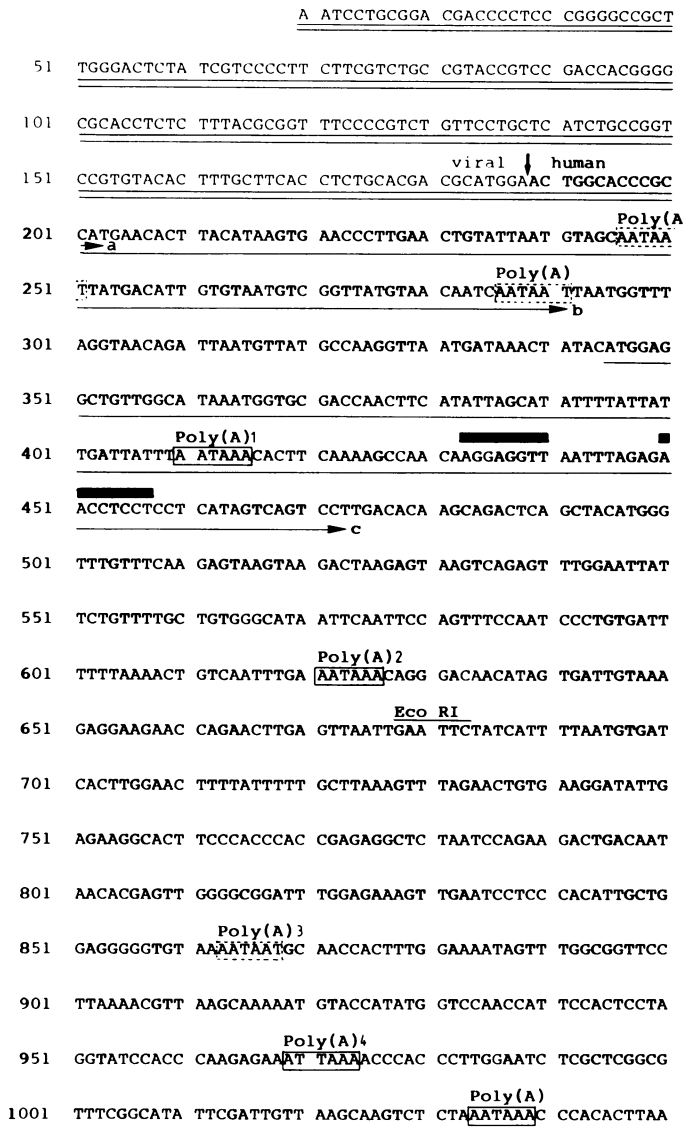

Fig. 6. Partial nucleotide sequence of A-10.7. The sequence shown covers the region from the rightmost *Acc*I site within HBV DNA to a position 418 bp 3′ of the *Eco*RI site within human DNA (compare Figure 1). Reading frames (a,b,c) are indicated by arrows. a and b are extensions of the viral reading frames P and X (Valenzuela *et al.*, 1980), respectively, both of which are truncated in the integrated HBV DNA. Conventional polyadenylation sequences (AATAAA, ATTAAA) are marked by solid boxes, tentative signals (AATAAT) by dotted boxes. Poly(A) sites accounting for the S1 mapped fragments (Figure 5) are numbered. The 8-bp perfect inverted repeat suggested to form a stem-loop structure is indicated by thick bars.

tion was investigated by S1 mapping. At the same time it was tested which of the so far described HBs gene promoters (Pourcel *et al.*, 1982; Gough, 1983; Cattaneo *et al.*, 1983; Rall *et al.*, 1983; Standring *et al.*, 1984) operates in this system.

To map the 5′ end of the Ltk 2425 HBsAg mRNA a 5′-labeled *Acc*I-*Sph*I fragment (1.8 kb; Figure 5) from pA1 10.7 was hybridized to total Ltk 2425 RNA and the hybrids subsequently digested with the single-strand specific nuclease S1. After alkali hydrolysis of the RNA, analysis of the DNA on a 6% polyacrylamide sequencing gel showed that a fragment of 846 nucleotides length was protected from S1 digestion (Figure 5, lane 2). This corresponds well with the results of Cattaneo *et al.* (1983) and demonstrates the transcription of the S gene from the SV40 late-like promoter (Figure 5). The possibility of the 5′ end of the HBsAg mRNA being a splice acceptor site rather than the proper start of the message has been excluded by Cattaneo *et al.* (1983). Transcription from the SV40 late-like promoter was also demonstrated for HBV DNA of subtype adw$_2$ in mouse Ltk cells (Standring *et al.*, 1984). This system was used here for comparisons of expression levels (Figure 4, Table II).

In the case of the 'TATA box' being the S gene promoter, a 1230-bp fragment should have been protected from S1 digestion. A fragment of this size is not visible in this experiment (Figure 5). Minor amounts, however, of transcripts promoted by the TATA sequence cannot be excluded by this analysis.

The 3′ end of the HBsAg mRNA was mapped as follows: a 4.4-kb *Acc*I-*Acc*I fragment (see Figure 5) was isolated from pA1 10.7 and hybridized to total Ltk 2425 RNA. Hybrids were digested with S1 nuclease. Remaining RNA was hydrolysed by alkali treatment and single-stranded DNA fragments separated on a 2% agarose gel. The DNA was blotted onto a nitrocellulose filter and hybridized to a ^{32}P-labeled HBV DNA probe. Four fragments of 430, 660, 880 and 990 nucleotides were visible upon autoradiography, all of which span the region (188 bp) between the *Acc*I site within HBV DNA and the host-virus junction and additionally cover different amounts of flanking human DNA. This conclusively proves transcriptional readthrough into human DNA. It is noteworthy that the different fragments are present in significantly different amounts. Intensities decrease as a function of distance from the HBV DNA. These results were corroborated by Northern blot analysis of Ltk 2425 RNA, again using HBV DNA as hybridization probe. A major band at position 2.1 – 2.2 kb and several bands of slightly higher mol. wt. were visible (data not shown).

Nucleotide sequence of A-10.7 flanking human DNA

The nucleotide sequence of the human DNA flanking the integrated HBV DNA copy in A-10.7 was determined (Figure 6). In good agreement with the S1 nuclease mapping of the 3′ end of the HBsAg mRNA, the following poly(A) signals are located in the positions indicated (measured from the rightmost *Acc*I site in HBV DNA): AATAAA (Proudfoot and Brownlee, 1976): sites 1 and 2 (positions 410 and 621) and ATTAAA (Jung *et al.*, 1980): site 4 (position 969). In addition, there is a sequence AATAAT (site 3; position 863) which could give rise to the S1 nuclease fragment of 880 nucleotides. This sequence has not previously been reported to function as a poly(A) signal. A number of other candidate poly(A) signals (Figure 6) apparently do not lead to polyadenylated transcripts.

Discussion

Transcription of hybrid host-virus mRNAs

Two copies of integrated HBV DNA (designated as A-10.7 and A-10.5) cloned from the human hepatoma cell line PLC/PRF/5 were assayed for their ability to direct the synthesis of HBsAg after transfection into mouse Ltk$^-$ cells. Both DNAs led to secretion of HBsAg proving that their HBsAg coding sequences are intact. The gene product was indistinguishable from naturally occurring HBsAg. The HBsAg mRNA transcribed in Ltk 2425 cells, which derive from transfection with pA1 10.7, was analyzed. Of the two sequences reported to function as the S gene promoter, a TATA box preceding the pre-S gene (Pourcel *et al.*, 1982;

Gough, 1983; Rall *et al.*, 1983) and an SV40 late-like promoter region (Cattaneo *et al.*, 1983), the latter was clearly proven here to be functional in this system. The same must apply to A-10.5, since its HBV sequences appear to be identical to those in A-10.7.

Transcriptional readthrough takes place into human flanking DNA. The transcription of hybrid RNA molecules composed of host and virus sequences has been long suspected (Wall and Darnell, 1971) and was first proven to occur in SV40-transformed cells (Sambrook *et al.*, 1980). With respect to integrated HBV sequences hybrid host-virus transcripts have so far not been shown. However, based on preliminary data Edman *et al.* (1980) had already suggested readthrough transcription in PLC/PRF/5 cells.

Detailed analysis revealed four HBV-specific hybrid RNA species of different lengths. The multitude of hybridizing fragments is explained by the presence of four polyadenylation signals in the flanking sequences of A-10.7 used at differential efficiencies. The use of more than one poly(A) site in a single transcription unit has been shown in other systems, e.g., the major late adenovirus-2 transcription unit [five poly(A) sites; Nevins and Wilson, 1981] and the dihydrofolate reductase gene in methotrexate-resistant mouse cells [four poly(A) sites; Setzer *et al.*, 1980]. As the viral region necessary for transcriptional termination and polyadenylation is also missing in A-10.5, we assume that in this case too transcriptional readthrough occurs into human DNA.

Augmented HBsAg expression is conferred by 3′ co-transcribed host sequences

Cell clones with stably integrated 10.7-kb DNA expressed eight times more HBsAg than cells containing similar amounts of 10.5-kb DNA and 33 times more than cells containing cloned HBV DNA (pKK HBs 34) of the same subtype. Similar differences were found in unselected cell culture, which proves that the different expression levels are not due to different integration events.

The region responsible for the augmented HBsAg expression resides in cellular sequences not directly adjacent to the 3′ end of the structural gene as can be inferred from the comparison between A-10.7 and A-10.5 which differ only in a distal part of their host sequences (Koch *et al.*, 1984a, 1984b). The augmented expression was studied for A-10.7 in greater detail. No indication at all was found for an enhancer effect.

As HBsAg mRNAs of A-10.7, A-10.5 and HBV DNA all contain different sequences 3′ of the S gene it appears probable that these account for the observed different levels in HBsAg expression. There are only a few other eucaryotic examples of a similarly drastic influence of sequences 3′ of a gene on the level of polyadenylated mRNA (e.g., Kaufman and Sharp, 1982; Zaret and Sherman, 1982; McDevitt *et al.*, 1984; Woychik *et al.*, 1984). With respect to HBV surface antigen expression this is the first time an augmentation has been shown by foreign sequences 3′ of the gene.

We consider the following two hypotheses to account for the different HBsAg expression levels. First, expression levels correlate with the actual sequences of the polyadenylation signals. In cells transfected with A-10.7, poly(A) signals 1 and 2 (both AATAAA) lead to higher RNA levels (Figure 5) than signals 3 and 4 (AATAAT and ATTAAA, respectively). Conceivably, in this system AATAAA is a more potent polyadenylation signal than ATTAAA and AATAAT. As the poly(A) signal of the viral S gene (TATAAA) also differs from AATAAA it may also be used less efficiently than AATAAA. Second, an 8-bp perfect inverted repeat is located 12 nucleotides 3′ of poly(A) signal 1 (Figure 6) which leads to the highest amount of RNA. No such repeat is located immediately 3′ of the viral poly(A) site. Hairpin structures have been detected predominantly in non-coding regions of procaryotic and eucaryotic genes and have been suggested to have regulatory significance (Mueller and Fitch, 1982; McDevitt *et al.*, 1984; Woychik *et al.*, 1984). The inverted repeat in A-10.7 might form a stable stem-loop structure protecting poly(A) site 1 from a nuclear single-stranded 3′ – 5′ exonuclease activity prior to polyadenylation. Since it is known that the poly(A) addition site is 10 – 30 nucleotides 3′ of the poly(A) signal (Nevins and Wilson, 1981) it is also conceivable that the hairpin structure may expose the poly(A) addition site in its loop. We feel that a combination of these hypotheses would best explain the high expression levels observed. Both hypotheses presented are testable by *in vitro* mutagenesis experiments.

Implications for hepatoma development

In contrast to the epidemiology, molecular biological studies have so far not provided much evidence for a causal relationship between HBV and primary liver carcinoma. Based on these experimental results the following hypothesis may be considered. As cellular sequences of A-10.7 and A-10.5 are co-transcribed at very high levels in Ltk cells, the same is likely true in PLC/PRF/5 cells. These sequences might be expressed at low levels or not at all in normal hepatocytes but become expressed at high levels as a consequence of HBV DNA integration thus contributing to cell transformation. Such a mechanism would be similar to proviral insertion oncogenesis by the slow acting transforming retroviruses which do not contain '*onc*' genes (Hayward *et al.*, 1981). A short reading frame of 129 nucleotides (reading frame c, Figure 6) is present in the transcribed cellular DNA in A-10.7. Furthermore, two hybrid host-virus reading frames (a,b, see Figure 6) are contained in A-10.7. The gene products of these reading frames may be related to the transformed state of PLC/PRF/5 cells. Although A-10.7 is not able to transform NIH 3T3 cells, sequences in the inserts described in this paper may well have transforming potential in hepatocytes, as transformation may be cell type specific. Experiments introducing the cloned HBV specific fragments into primary hepatocyte cultures are being attempted.

Materials and methods

Transfection of cells

Mouse Ltk$^-$ cells were grown in Dulbecco's modified Eagle's (DME) medium supplemented with 10% fetal calf serum and 50 μg/ml kanamycin. Transfection by the calcium phosphate method was carried out as described by Stratowa *et al.* (1982). *Hind*III-digested plasmid DNA (for amounts see legend to Figure 5) was made up to 10 μg with calf thymus DNA, precipitated and transfected. Cells were kept in monolayers for 13 days and supernatants assayed for HBsAg by radioimmunoassay. For *tk* selection 5 μg of phage DNA or *Hind*III-digested plasmid DNA (for amounts see Table legends), and 50 ng of pFG5 (containing the thymidine kinase gene of herpes simplex virus I; Colbere-Garapin *et al.*, 1979) were made up to 20 μg with salmon sperm DNA and precipitated. Two days after transfection, selection in hypoxanthine, aminopterin, thymidine (HAT) medium was started. About 2 weeks later, individual colonies were picked with cloning cylinders, transferred to culture wells (Costar) and supernatants assayed for HBsAg by radioimmunoassay. Some cell colonies were later grown to mass cultures.

HBsAg detection

HBsAg was measured with AUSRIA II radioimmunoassay kits (Abbott). Samples were counted in a gamma counter. Positive controls were between

7000 and 9000 c.p.m. and negative controls were around 140 c.p.m. As measurement was only linear between 20 and 40% (1000–2000 c.p.m.) of positive controls, test samples were diluted, if necessary, to lie within this range. Results are expressed as P/N (experimental result over negative control). Supernatants with values of P/N ≥ 2.1 are considered to contain HBsAg.

Purification of HBsAg particles

Cell media were centrifuged at 15 000 *g* for 24 h. The pellet was dissolved in 1 ml of medium and HBsAg particles were further purified by two subsequent CsCl density centrifugations (density 1.21 g/ml) in a Beckman SW 41 rotor at 45 000 r.p.m. at 17°C for 60 h and precipitation with 10% polyethyleneglycol 6000. Pellets were resuspended in phosphate buffered saline (PBS).

Electron microscopy

Samples were adsorbed on carbon-coated grids and stained with 2% uranyl formeate. Grids were examined under a Siemens Elmiscop 102.

Polypeptide analysis

Ltk 2425 cell cultures, which were 80% confluent, were incubated for 20 h in HAT medium containing 0.1 mg/l cysteine, 400 μCi of [^{35}S]cysteine and no cystine. Supernatants were collected and precipitated with 10% polyethyleneglycol 6000. Pellets dissolved in PBS were immunoprecipitated with human anti-HBs containing serum (obtained from the Max-von-Pettenkofer Institut, Munich) essentially as described (Stratowa *et al.*, 1982). Samples in 5% SDS and 0.5 M dithiothreitol were boiled for 4 min and electrophoresed through slab gels of 13% polyacrylamide. Gels were fluorographed, dried and exposed to X-ray film for 48 h.

S1 nuclease analysis

RNA was isolated by the guanidium isothiocyanate method (Chirgwin *et al.*, 1979).

The 5′ end mapping experiment was carried out essentially as described by Weaver and Weissmann (1979). Briefly, 30 μg of total Ltk 2425 RNA or as blank tRNA from yeast were hybridized to a 5′ end-labeled 1.8-kb *Acc*I-*Sph*I fragment from pA1 10.7 (see Figure 5) in 20 μl 80% formamide, 40 mM Pipes (pH 6.4), 1 mM EDTA and 0.4 M NaCl at 53°C for 16 h. Hybrids were incubated for 30 min at 37°C with 1 (tRNA only), 10 or 30 units of S1 nuclease (PL Biochemicals) in 200 μl of 0.25 M NaCl, 30 mM Na acetate, 1 mM $ZnSO_4$ and 100 μg salmon sperm DNA and subsequently treated with 0.2 M NaOH for 2 h at 45°C. Protected DNA fragments were electrophoresed through a 6% DNA sequencing gel (Maxam and Gilbert, 1980).

The 3′ end was mapped by hybridizing Ltk 2425 RNA or yeast tRNA at 59°C to a 4.4-kb *Acc*I-*Acc*I fragment from pA1 10.7 (see Figure 5), followed by S1 nuclease digestion and alkali treatment as described above. Protected fragments were separated on a 2% agarose gel, blotted on a nitrocellulose filter and hybridized to ^{32}P-labeled HBV-DNA.

DNA sequencing

Nucleotide sequencing was done following the method of Maxam and Gilbert (1980).

Acknowledgements

We are grateful to Irene Dick, Sabine Hankel, Ester Piravandi-Weber, Christine Schmidt and Ingrid Scholz for excellent technical assistance and to Dr Ralf Schnabel for advice on RNA experiments. We further wish to thank Dr V.Gauss-Mueller for providing human anti-HBs serum and Akis Sorbas for ϕX174 marker DNA. This work was supported by the Deutsche Stiftung für Krebsforschung.

References

Alexander,J.J., Bey,E.M., Geddes,E.W. and Lecatsas,G. (1976) *S. Afr. Med. J.*, **50**, 2124-2128.

Cattaneo,R., Will,H., Hernandez,N. and Schaller,H. (1983) *Nature*, **305**, 336-338.

Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry (Wash.)*, **18**, 5294-5299.

Colbere-Garapin,F., Chousterman,S., Horodniceanu,F., Kourilsky,P. and Garapin,A.-C. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3755-3759.

Edman,J.C., Gray,P., Valenzuela,P., Rall,L.B. and Rutter,W.J. (1980) *Nature*, **286**, 535-538.

Freytag von Loringhoven,A. (1984) Ph.D. Thesis, Ludwig-Maximilians-Universität, Munich.

Gerin,L., Purcell,R.H., Hoggan,M.D., Holland,P.V. and Chanock,R.M. (1969) *J. Virol.*, **4**, 763-768.

Gough,N. (1983) *J. Mol. Biol.*, **165**, 683-699.

Graham,F.L. and van der Eb,A.J. (1973) *Virology*, **52**, 456-467.

Hayward,W.S., Neel,B.G. and Astrin,S.M. (1981) *Nature*, **290**, 475-480.

Hearing,P. and Shenk,T. (1983) *Cell*, **33**, 695-703.

Jolly,J.D., Esty,A.C., Subramani,S., Friedman,A.T. and Verma,I. (1983) *Nucleic Acids Res.*, **11**, 1855-1872.

Jung,A., Sippel,A.E., Grez,M. and Schuetz,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5759-5763.

Kaufman,R.C. and Sharp,P.A. (1982) *Mol. Cell. Biol.*, **2**, 1304-1319.

Koch,S., Freytag von Loringhoven,A., Hofschneider,P.H. and Koshy,R. (1984a) *EMBO J.*, **3**, 2185-2189.

Koch,S., Freytag von Loringhoven,A., Kahmann,R., Hofschneider,P.H. and Koshy,R. (1984b) *Nucleic Acids Res.*, **12**, 6871-6886.

Koshy,R., Koch,S., Freytag von Loringhoven,A., Kahmann,R., Murray,K. and Hofschneider,P.H. (1983) *Cell*, **34**, 215-223.

Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.

McDevitt,M.A., Imperiale,M.J., Ali,H. and Nevins,J.R. (1984) *Cell*, **37**, 993-999.

Mueller,U.R. and Fitch,W.M. (1982) *Nature*, **298**, 582-585.

Nevins,J.R. and Wilson,M.C. (1981) *Nature*, **290**, 113-118.

Pellicer,A., Wigler,M., Axel,R. and Silverstein,S. (1978) *Cell*, **14**, 133-141.

Pourcel,C., Louise,A., Gervais,M., Chenciner,N., Dubois,M.-F. and Tiollais,P. (1982) *J. Virol.*, **42**, 100-105.

Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211-214.

Rall,L.B., Standring,D.N., Lamb,O. and Rutter,W.J. (1983) *Mol. Cell. Biol.*, **3**, 1766-1773.

Sambrook,J., Greene,R., Stringer,J., Mitchison,T., Hu,S.-L. and Botchan, M. (1980) *Cold Spring Harbor Symp. Quant. Biol.*, **44**, 569-584.

Setzer,D.R., McGrogan,M., Nunberg,J.H. and Schimke,R.T. (1980) *Cell*, **22**, 361-370.

Soberon,X., Covarrubias,L. and Bolivar,F. (1980) *Gene*, **9**, 287-305.

Standring,D.N., Rutter,W.J., Varmus,H.E. and Ganem,D. (1984) *J. Virol.*, **50**, 563-571.

Stibbe,W. and Gerlich,W.H. (1982) *Virology*, **123**, 436-442.

Stratowa,C., Doehmer,J., Wang,Y. and Hofschneider,P.H. (1982) *EMBO J.*, **1**, 1573-1578.

Szmuness,W. (1978) *Prog. Med. Virol.*, **24**, 40-51.

Valenzuela,P., Quiroga,M., Zaldivar,J., Gray,P. and Rutter,W. (1980) in Fields,B.N., Jaenisch,R. and Fox,C.F. (eds.), *Animal Virus Genetics.*, Academic Press, NY, pp. 57-70.

Wall,R. and Darnell,J. (1979) *Nature New Biol.*, **232**, 73-76.

Wang,Y., Schaefer-Ridder,M., Doehmer,J. and Hofschneider,P.H. (1982) *EMBO J.*, **1**, 1213-1216.

Weaver,R.F. and Weissmann,C. (1979) *Nucleic Acids Res.*, **7**, 1175-1193.

Woychik,R.P., Lyons,R.H., Post,L. and Rottman,F.M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3944-3948.

Zaret,K.S. asnd Sherman,F. (1982) *Cell*, **28**, 563-573.

Received on 11 October 1984; revised on 19 November 1984