Biochem. J. (1987) 248, 1–11 (Printed in Great Britain)

REVIEW ARTICLE

The bovine papillomavirus genome and its uses as a eukaryotic vector

Paul E. STEPHENS and Christopher C. G. HENTSCHEL
Department of Molecular Genetics, Celltech Ltd., R16 Bath Road, Slough, Berkshire SL1 4DY, U.K.

INTRODUCTION

The development and use of eukaryotic expression systems has become widespread in recent years. The limited capacity of microbial systems to express authentic and functional eukaryotic proteins is one factor which fuels these developments [1,2]. Another is the desire of numerous investigators to be able to study diverse aspects of gene expression with highly manipulatable and yet physiological model systems. The characteristics of the bovine papillomavirus make it an attractive choice for such a system. These include the episomal nature of the genome in transformed cells and the fact that these cells remain viable over many generations. The question of which type of eukaryotic vector system should be used for a particular study or project can provoke arguments. In this Review we do not intend to pursue such comparisons as ends in their own right, but rather accept the view that all the current systems are appropriate for some uses and inappropriate for others. We will confine ourselves to the development of the bovine papillomavirus genome (specifically BPV1) as a eukaryotic vector and to the increasingly detailed understanding of its molecular genetics which has paralleled this development.

The biology and biochemistry of papillomaviruses have been reviewed recently [3–5] and will not be discussed here. Nor will the evidence that viruses can play a promoting role in a number of specific neoplasia [5].

FUNCTIONAL ELEMENTS OF THE BPV GENOME

Papillomaviruses were once classified along with polyomaviruses in the family Papovaviridae. However, sequence and functional analysis now indicates that these two classes of virus are not related. The polyomaviruses are much smaller and RNAs are transcribed from both DNA strands in a bidirectional fashion. Papillomaviruses will transform fibroblasts in culture and *in vivo*. However, viable viruses are only produced in terminally differentiated papillomas and cannot be cultured *in vitro*.

A diagrammatic representation of the BPV1 genome is shown in Fig. 1 and indicates the positions of the viral open reading frames (ORFs), the sizes and locations of the viral transcripts and some of the relevant regulatory functions. The BPV1 genome is a double-stranded DNA molecule of 7945 nucleotides whose sequence was first determined by Chen *et al.* [6] and by Stenlund *et al.* [7]. It is conventionally shown as a linear map opened at the unique *Hpa*I site with numbering starting at the G of this recognition site. The sequence analysis confirmed previous observations that all of the open reading frames are on the same strand [8,9]. The sequences of several animal papillomaviruses have been elucidated. They show significant homologies, allowing the classification of the virus into subtypes and the delineation of gene functions [3].

Lowy *et al.* [10] showed that a 69% *Bam*HI–*Hin*dIII fragment is capable of inducing cellular transformation. Coding information required for transformation must therefore be contained in this region and ORFs from this region are referred to as early (E) ORFs [11]. Five RNA species all mapping to this region were found in transformed mouse cells in culture [12]. RNA transcripts which hybridized with the non-transforming 31% region were only found in tumours and fibropapillomas [8,13]. These transcripts coincide with the late region ORFs L1 and L2. A 900 bp region with no significant ORFs exists between the late and early ORFs (co-ordinates 7093 to 48). This region is now known to contain several regulatory sequences required for transcription and replication (see below).

The viral open reading frames

Ten viral ORFs have been identified, eight early ORFs which are required for early viral functions, replication and transformation, and two late ORFs L1 and L2. Functions have been ascribed to seven of the ten ORFs; these are summarized in Table 1. The relative positions of the ORFs are shown in Fig. 1.

L1 and L2 open reading frames. The L1 and L2 proteins are encoded by two ORFs in the same phase which span the 31% non-transforming region (co-ordinates 4171–7095) with a termination codon at co-ordinate 5593 [6]. The ORFs code for the major and minor viral capsid proteins [14] and mRNAs corresponding to these ORFs are consequently only detected in terminally differentiated papillomas [13]. Clear evidence that L1 is a major viral capsid protein comes from the fact that an L1 fusion protein produced in *Escherichia coli* is capable of producing a neutralizing antiserum which prevents BPV transformation in C127 cells [14]. In addition, mRNA isolated from a BPV-infected wart can direct the synthesis *in vitro* of a corresponding 55 kDa major capsid protein [13].

E1 open reading frame. The E1 ORF (co-ordinates 813–2663) has the largest coding capacity of all the BVP ORFs with the potential of producing a protein approx.

Abbreviations used: BPV(1), bovine papillomavirus (1); ORF, open reading frame; RI, replication intermediate; LTR, long terminal repeat; tPA, tissue plasminogen activator; MT, metallothionein.

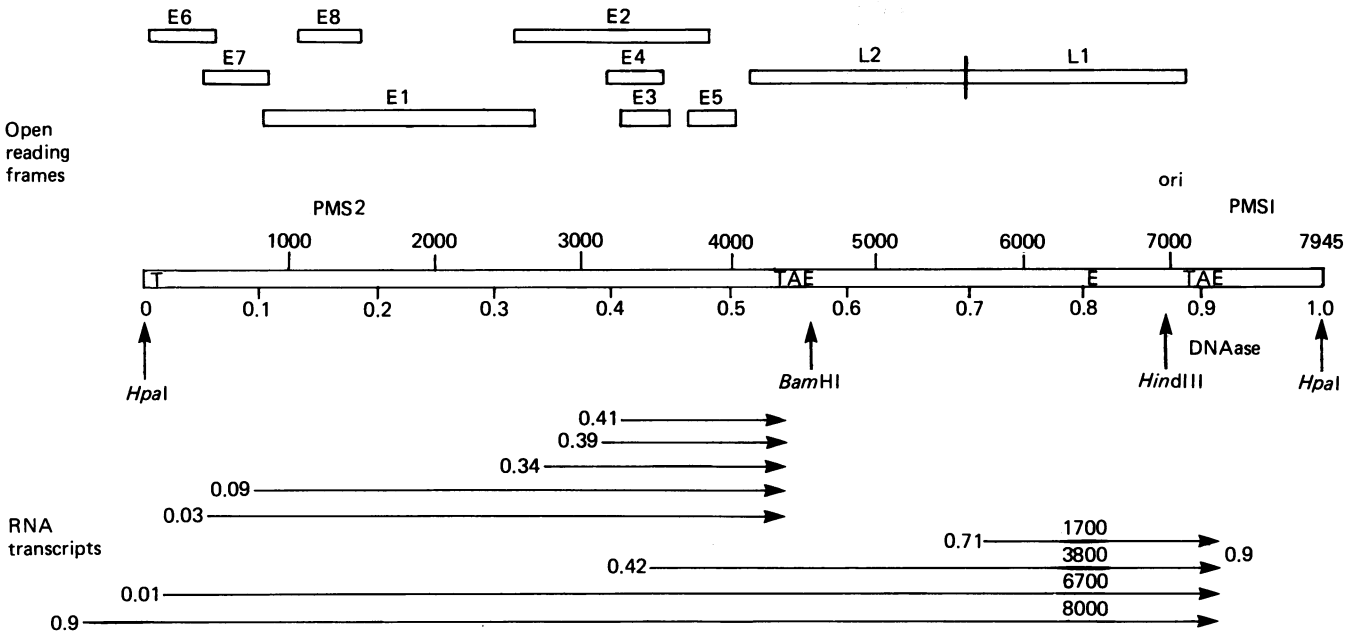

Fig. 1. Diagrammatic representation of the BPV1 genome

The BPV genome is shown as a linear map opened at the unique *Hpa*I site. The map is divided into map units 0–10 below the line and nucleotides 1–7945 above the line [6]. The position of several potential regulatory elements are shown: TATAAA and TATATA promoter elements, T; polyadenylation recognition sequences (AATAAA), A; enhancer elements, E. The location of the origin of replication (ori), the DNAase I hypersensitive site (DNAse) and two plasmid maintenance sequences (PMS1 and 2) are also indicated. The positions of the ten ORFs (E1–E8, L1 and L2) are indicated; the vertical bar between L1 and L2 indicates the position of the termination codon. The BPV-specific polyadenylated RNA transcripts and their sizes as found in transformed cells in culture [35] or in papillomas [13] are shown as arrowed lines below the map.

Table 1. Viral open reading frames

ORF	Coding capacity (amino acid residues)	Function and characteristics
E1	600	Control of plasmid replication. An E1$^-$ mutant has a REP$^-$ phenotype, i.e. the virus integrates into the host genome. The 3′ end has homology with SV40 large T.
E2	400	Transactivator of viral enhancer.
E5	44	Transforming hydrophobic peptide associated with cellular membranes.
E6	140	Transforming protein in C127 but not NIH 3T3 cells. Possible nucleic acid binding function. Found in association with cell membranes and the nucleus.
E7	100	Control of copy number. E7$^-$ mutant has copy number of 1–5 (COP$^-$ phenotype).
L1	500	Major capsid protein.
L2	500	Minor capsid protein.

600 amino acids long. The E1 ORF overlaps with E2, E7 and E8 ORFs, a feature typical of all the early region ORFs [6]. There is considerable evidence implicating E1 in the control of viral replication [15–20, and see below]. Thus, mutations in E1 lead to a phenotype in which the BPV fails to replicate and becomes integrated into the genome, the REP$^-$ phenotype. Interestingly, there exists sequence homology between the *C*-terminal halves of the E1 gene product and the large T antigen from simian virus 40 (SV40) [21,22]. This region corresponds to the ATPase and nucleotide-binding functions of the T antigen, possibly indicating that these activities may be associated with E1.

E2 open reading frame. The E2 ORF maps to the 3′ end of the transforming region of the genome (co-ordinates 2581–3837) and overlaps with the E3 and E4 ORFs [6]. Sequence information from a separate BPV isolate

indicates the presence of an extra guanosine residue at position 3444, thus extending the E2 ORF to co-ordinate 3864 [7]. The E2 gene product was initially implicated as a transforming protein, based on mutational analysis [20]. This view appeared to be supported by studies based on the expression of subgenomic fragments or cDNAs from the 3′ end of the transforming region [19,23]. However, more detailed analysis, either by the introduction of nonsense mutants specifically into the E2 ORF [24] or insertional mutagenesis of the corresponding cDNA clone [25], leads to the conclusion that the E2 product is not a transforming protein but a transactivator of the BPV regulatory unit [26]. Thus some of the effects reportedly caused by mutations in the E2 gene, such as decreased efficiency of transformation, may now be accounted for by a lack of transactivation of the real transforming genes. More recently the E2 gene has been expressed in *E. coli* with the view to raising antibodies [27]. These should help to elucidate further the mode of action of the E2 product.

E5 open reading frame. The E5 ORF maps to the extreme 3′ end of the 69% region (co-ordinates 3714–4010) and overlaps the 3′ end of the E2 ORF [6]. Mutational analysis indicated that this region of the genome has a transforming function [20]. Furthermore, expression of the E5 ORF alone, from heterologous viral promoters, is sufficient to cause transformation of C127 and NIH 3T3 cells, which strongly suggests that E5 is indeed a transforming protein [25,28,29]. Interestingly, the E5 gene product is small, hydrophobic and only 44 amino acids long (15 of which are leucine residues) [30]. The peptide is encoded by the 3′ half of the open reading frame, in agreement with data from earlier mutational analysis [28]. Antibodies raised to a synthetic 20-amino-acid peptide corresponding to its *C*-terminus localize the E5 protein predominantly to cellular membranes [30].

E6 open reading frame. The E6 ORF maps at the far 5′ end of the 69% transforming region (co-ordinates 49–501) and overlaps with the E7 ORF [6]. Sarver *et al.* have localized a transforming function to this region by deletion mutagenesis [20]. Further studies, using either cDNAs resulting from mRNAs comprised of splicing fusions of the E6/E7 ORFs, or subgenomic fragments expressed from viral promoters in conjunction with site-directed mutagenesis, confirmed that E6 is indeed involved in transformation [23,31]. Surprisingly, there appears to be some cell type specificity in that E6, unlike E5, is only capable of transforming C127 cells and not NIH 3T3 cells [23,31]. The E6 gene has been expressed as a fusion protein in *E. coli* and antibodies have been raised to this protein [32]. The antibodies immunoprecipitated an appropriately sized protein present in C127 cells transformed either by a subgenomic fragment containing the E6 ORF or by the whole virus. The protein is located both in the nucleus and in association with cellular membranes [32]. The E6 gene product is a 15.5 kDa cysteine-rich (10.9% of the residues are cysteine), basic (17% of the residues are lysine or arginine) peptide. The cysteines are arranged in repeats of Cys-Xaa-Xaa-Cys, a sequence proposed to be characteristic of nucleic acid binding proteins [33].

E7 open reading frame. The E7 ORF overlaps both the E6 and E1 ORFs (co-ordinates 449–859) [6]. Neither a protein nor a specific transcript has been mapped to this gene. Mutations in the E7 ORF maintain the BPV genome as an extrachromosomal element in transformed cells but at a greatly reduced copy number (one to five copies per cell), the COP^- phenotype [18]. Berg and co-workers suggest that there exists a temporal need for the expression of E7 for the initial establishment phase of BPV replication. The transformation defects due to mutations in E7 may thus be an indirect consequence of a low gene dosage in the transforming genes [34].

E3, E4 and E8 open reading frames. The E3 and E4 ORFs overlap with each other and with E2; the E8 ORF overlaps with E1. No clear functions have been assigned to these ORFs to date. Moreover, the E3 ORF does not possess an in-phase methionine codon, implying that it may not even code for a polypeptide.

Transcriptional control elements

The five early transcripts are present in only very low abundance in transformed cells, due presumably to the weakness of the BPV promoters [35]. In order to assist analysis of the transcripts Yang *et al.* produced a cDNA bank of BPV-specific mRNAs [23]. From a bank of 100000 colonies they isolated 200 BPV-specific clones which were arranged into several classes. These classes represent a number of different splicing arrangements of the ORFs. Some consist of individual ORFs and others of fusions between different ORFs. The fused ORFs may either be derived from contiguous or disperse parts of the early region of the genome. All of the early region transcripts however share a common 3′ end. Sequence analysis of several cDNA clones confirmed that this was due to the polyadenylation recognition site at base 4179 [23].

Two TATAAA sequences are found at co-ordinates 7108 and 58 located upstream of the early ORFs. The latter may constitute the principle gene promoter, directing the transcription of mRNA species with 5′ ends at co-ordinate 89 [23]. Functional analysis using a promoter-less TK plasmid has confirmed the presence of two corresponding promoter elements in the non-coding region [36]. A third promoter fragment, defined by this assay, is located in the coding sequence of the E2 ORF (co-ordinates 3095–3355). However, this analysis failed to detect a promoter function which could direct the synthesis of mRNA species which have 5′ starts further downstream, possibly implying that these were derived from either truncated or spliced mRNAs [23].

The late L1 and L2 ORFs are not expressed in transformed C127 cells [13,35]. A TATATA sequence 5′ to the late ORFs can be found at co-ordinates 4072–4077. Whether this sequence forms part of the functional promoter for the late genes or whether the late gene mRNAs are a result of splicing transcripts expressed from the early region promoter is not clear. Engel and co-workers identified both large and small transcripts which mapped to the late region in a BPV-infected fibropapilloma, thus leaving the question of the late region promoter unresolved [13].

Transcriptional activity of the BPV promoters is modulated by three enhancer elements. Their positions, indicated in Fig. 1, map to: the 3′ end of the early viral gene transcripts [36]; the 900 bp non-coding region [26]; and the 31% non-transforming region [37 and see below]. The 3′ enhancer was identified initially by its

ability to enhance *tk* expression in mouse L TK$^-$ cells. The sequence responsible was further delineated to a 59-bp region 3′ to the early poly(A) addition sequence [38]. This enhancer is host specific as illustrated by the fact that it increased the level of *tk* gene expression to a greater extent in bovine cells than in murine or human cells [39]. The 5′ enhancer element located in the non-coding region acts as a typical viral transcriptional enhancer. It is located in the vicinity of the viral promoter, possesses sequence homology to the core enhancer consensus sequence and can be transactivated by the viral protein E2 [26].

Control of replication and episomality

BPV1 DNA, whether in fibropapillomas, tumours, or in transformed mouse cells, was originally reported to exist exclusively as extrachromosomal plasmids replicating autonomously at a level of 10–200 copies per cell [9,10,40,41]. However, more recent studies with some BPV-based vectors indicate that integration or complexing into multimers also occurs [42–44, and see below].

Studies by Waldeck *et al.* [45] have mapped the origin of replication to the 648 bp *Cla*I C fragment, co-ordinates 6834–7481. Replication intermediates (RI) were treated with single-cut restriction enzymes to determine the origin of the extending replication eyes by using electron microscopy. 'Cairns'-type RI molecules were seen which mapped the origin to co-ordinate 6940 ± 400 bp. A transient replication assay was then used to map accurately the origin in the *Cla*I C fragment. This region also corresponds closely to the DNAase I hypersensitive site between map units 0.88 and 0.92 detected previously in the non-coding region [46], and to a sequence mapped by Lusky & Botchan which is required for plasmid maintenance, PMS1 [15]. In fact, Lusky & Botchan found two such plasmid maintenance sequences, PMS1 and PMS2, which if deleted led to integration into the host genome. Moreover, plasmids containing either of these sequences alone (and no other viral genes) remain episomal if viral functions are supplied in *trans*. Whether PMS2 represents a normal origin of replication is unclear. It is located in the E1 ORF and has not been found to generate a 'Cairns' structure in a replication assay. A third element which has a proposed role in the replication of the BPV genome has been mapped in the 31% non-transforming region [37, and see below].

Recent studies have indicated that there are two phases of BPV replication in transformed cells, an initial amplification phase followed by a maintenance phase [17,34]. Lusky & Botchan described two *trans*-acting functions, *rep* and *cop*, encoded by the E1 and E7 ORFs respectively [18]. Mutations in E1 lead to integration of the plasmid, whereas those in E7 lead to episomal maintenance but at only one to five copies per cell. Co-transfections with a plasmid containing the E6/E7 fusion cDNA driven by a viral long terminal repeat (LTR) and a plasmid with a mutation in E7 allowed complementation of the copy number defect. However, subsequent superinfection with the cDNA clone does not lead to complementation of the E7 mutation [34]. This observation, and data derived using SV40–BPV chimaeras, has led to the postulation that BPV replication is controlled in part by a *trans*-acting repressor [16]. A recent model for the replication of BPV thus suggests that it initially replicates to a high copy number due to the expression of a positive replication factor (R) and then the amplification phase is ended by expression of a repressor factor or modulator protein (M) [18]. These two factors are encoded by the E1 ORF; the repressor factor maps to the 5′ end and the positive factor to the 3′ end [17]. In the maintenance stage of replication the modulator and replication factors act in concert to produce a stable once-per-cycle replication. Mutations in the E7 ORF appear to reduce the expression of E1, resulting in a low level of both R and M and a consequent low copy number [17].

Control of transformation

Bovine papillomaviruses are capable of inducing transformation in bovine cell cultures derived from foetal meninges, conjunctiva and palate [47,48]. Unlike many other animal papillomaviruses, BPV appears to be only partially species specific and can induce benign tumours in horse, mice and hamsters [49]. However, BPV does appear to be cell type specific, and produces tumours only in dermal fibroblasts and transforms only fibroblasts in culture [47,49]. The most extensively used host cells for BPV are either mouse C127 or NIH 3T3 fibroblasts, C127 cells are a non-transformed clonal line derived from a mammary tumour of an RIII mouse [50].

Transformed C127, and less noticably NIH 3T3 cells, produce a focus of heaped-up, criss-crossed, spindle-like cells against the normal flat fibroblastic contact-inhibited monolayer [49]. These transformed cells are capable of growth in low serum concentrations, can grow in soft agar, have lost contact inhibition, grow to a high cell density and cause tumours in nude mice [10,47,49].

As mentioned previously, the factors necessary for cellular transformation are located on a 5.4 kb *Hin*dIII–*Bam*HI fragment, the 69% transforming fragment [10]. More detailed analysis indicated two transforming functions within this region, one at the 3′ end and one at the 5′ end, which when linked to a LTR were individually capable of transforming C127 cells [19,31]. Mutational analysis defined these transforming regions as the E5 (3′ segment) and E6 (5′ segment) ORFs [25,28,29]. The protein products of these genes have now been identified and their cellular localization indicated by the use of specific antisera [30,32, and see above]. The expression of E5 and E6 are controlled by a promoter and enhancer in the 5′ non-coding region. It is not clear whether both genes are required. However, when both genes are expressed there is a synergistic effect [23,31]. Both are apparently transcribed in fibropapillomas and transformed mouse cells [8,13,35].

Interestingly, transformation efficiency of the 69% *Bam*HI–*Hin*dIII is slightly less than that obtained when the full genome is used. If bacterial sequences are linked to the 69% fragment then transformation efficiency falls substantially [37,43]. Transformation efficiency is however restored by the addition of certain eukaryotic genes which have 'stimulatory properties' (e.g. β-globin [51], rat pre-proinsulin [52], and rat growth hormone [53]). These initially puzzling observations are now best explained by the work of Lusky & Botchan who mapped a BPV enhancer in the 31% non-transforming region [37]. This enhancer is implicated in the replication of the virus. It is presumed to activate transcription at the origin of replication. Removal of the enhancer leads to reduced replication of the genome and a consequent

decrease in transformation efficiency. This defect can apparently be overcome by alternative enhancers in the genome. However, if there are intervening bacterial sequences present this compensatory activity is lost, leading to decreased transformation efficiency [37,54]. Genes like β-globin and pre-proinsulin must therefore possess some sort of enhancer function which can replace the enhancer function in the 31% non-transforming region. The transformed phenotype is controlled by the level of expression of the E5 and E6 transforming genes. Mutations in these genes, or mutations which lead to a reduced replication of the BPV genome, and a consequent reduction in dosage of these genes, therefore both lead to a reduced transformation efficiency.

BPV AS A VECTOR

The development of the vector-system

In the previous sections we have indicated the complexities encountered when recombinants were first made linking bacterial sequences to BPV sequences. Early BPV shuttle vectors often used the 69% transforming region linked to pBR322 sequences, resulting in a decrease in the transformation efficiency [43,51]. This decrease in efficiency was the same whether the complete pBR322 plasmid or a 'poison minus' derivative, pML, was used (in contrast to analogous experiments with SV40 [55]). An interesting observation in this context was reported by Campo *et al.* [56]. They linked plasmid pAT153, a derivative of pBR322 containing a deletion which is not as extensive as pML [57], to the 69% transforming region and achieved normal levels of transformation. The block in BPV function could be relieved by the presence of 'stimulatory elements' present in some foreign genes such as β-globin and pre-proinsulin [51–53,58]. If however the whole BPV genome is used different results are obtained. Thus when bacterial sequences such as pML or pAT153 are linked to the whole BPV genome normal transformation efficiencies are obtained, but linkage of the complete pBR322 moiety resulted in a 200-fold decrease in transformation efficiency [43,56,59]. These results exemplify three basic types of BPV plasmid vectors, the essential characteristics of which are illustrated in Fig. 2. The first type consists of the BPV sequence, usually the 69% transforming segments, no bacterial sequences and a eukaryotic expression cassette. This type of construction has several disadvantages: the complexity of the constructions is limited by the requirement for specific restriction enzyme sites which allow the removal of the bacterial sequences prior to transfection; transfections were often performed using linear DNA fragments which frequently resulted in complicated rearrangements [12,52,60]; and vectors cannot be shuttled back into bacterial cells. The second type of vector overcomes some of these disadvantages. They consist of: the 69% transforming region of BPV; bacterial sequences, including an origin of replication and a selectable marker; a 'stimulatory segment' containing a foreign gene, which could overcome the inhibitory effect of the bacterial sequences; and the gene of interest [51,53,61–65]. The third type of vector uses the whole BPV genome, bacterial sequence and the gene of interest [66–68]. All three types of vector have in fact been used successfully for the expression of foreign genes with little or no difference in the levels of expression (see Table 2).

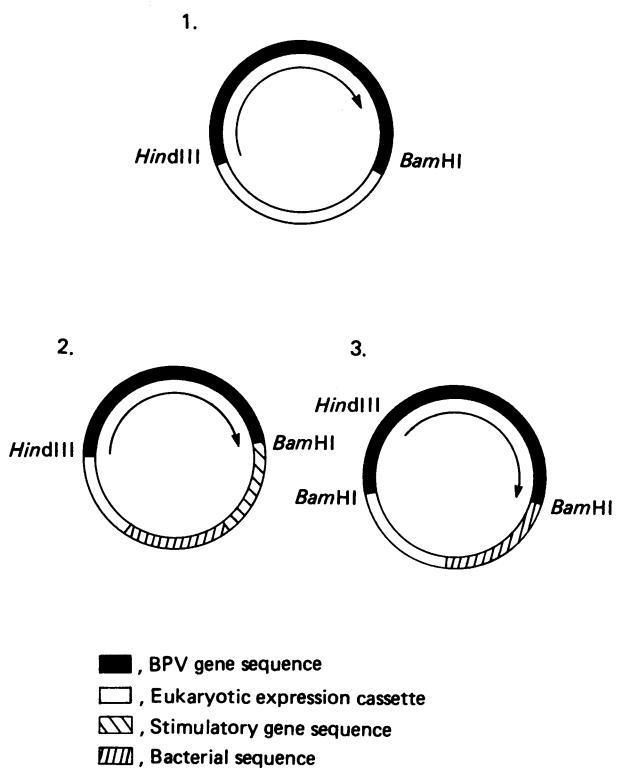

Fig. 2. Illustration of the three basic BPV vectors

The characteristics of the three basic BPV vectors described in the text are shown. Number 1 is the basic 69% transforming fragment plus the gene of interest. Number 2 consists of the 69% fragment, a stimulating gene sequence, bacterial sequences and the gene of interest. Number 3 contains the complete BPV genome, bacterial sequences and the gene of interest. The position of relevant restriction enzyme site are indicated. The arrow → indicates the direction of BPV transcription.

In addition to focus formation four different types of selection for transformants have been exploited, namely: thymidine kinase, *tk* [69]; xanthine–guanine phosphoribosyltransferase, *gpt* [44,70]; aminoglycoside phosphotransferase, *neo* [41,71,72]; and metallothionein, MT [58,63,73,74]. These selective markers were used in an attempt to increase the host range of BPV beyond that of the normally transformed host. However, these experiments have by and large been unsuccessful.

The insertion of the *tk* gene into the BPV genome did allow the selection of a TK^+ phenotype in TK^- mouse L and Syrian hamster BHK21 cell lines. However, transformation frequencies to the TK^+ phenotype were no greater than that achieved by non-BPV-containing plasmids and the plasmid copy number in the TK^+ cells was also very low [69]. For reasons which are unclear *gpt* selection leads to unusually high levels of genomic rearrangements, concatermerization and possible integration of the BPV DNA in C127 cells [44]. Consequently this selective marker is now rarely used in this system. Resistance to the drug G418 imparted by the *neo* gene has proved more successful, and several groups have been able to isolate G418-resistant colonies without evidence of plasmid rearrangement [44,71,72]. However, these results were only obtained in cell lines which would normally be transformed by BPV, thus not allowing an

extension of the host range. Similarly, resistance to heavy metals encoded by the metallothionein genes has routinely allowed the selection of cadmium-resistant C127 colonies [58,63,73,74]. Karin *et al.* have also successfully used Cd^{2+} selection in Rat 2 and F9 mouse teratocarcinoma cells to isolate stable cell lines. However, these contain only 10–15 copies of the BPV vector per cell [63]. In both of these cases (i.e. G418 or heavy metal resistance) selection for the dominant marker results in at least a 10-fold increase in the number of selectable colonies over what would be expected from a focus-forming assay [58,71]. The resistant colonies initially appear to be flat and untransformed. However, with time they do adopt a transformed morphology [44]. The reason for the reduced numbers of foci compared with resistant colonies and the delay in appearance of the transformed phenotype is assumed to be due to a requirement for a threshold level of expression of transformation functions [44]. The selection of cadmium-resistant foci has been used to isolate transformed C127 cell lines producing a number of foreign gene products, including tissue plasminogen activator (tPA) and tissue inhibitor of metalloproteinase (TIMP) [73–75].

The use of BPV vectors in the study of cellular regulation

The normally episomal nature of the BPV vector system makes it a useful tool for looking at a whole range of biological phenomena relating to gene replication and regulation, without the interference of the host chromosomes (i.e. position effects). These have included: the induction of metallothionein gene expression by heavy metals and glucocorticoids; the regulation of human β-interferon; the cell cycle dependent expression of histone genes; and the mutation frequency of transfected DNA.

Pavlakis & Hamer have used the mouse metallothionein I (mMTI) gene promoter to express various foreign genes in the BPV system [76]. They noted that the level of expression of the foreign gene was inducible by the heavy metal cadmium but not by zinc or by the glucocorticoid dexamethasone. The level of induction was low, only 2–10-fold, whereas the endogenous mMTI gene was induced 10–20-fold [76]. A similar relatively low level of induction has also been seen in our laboratory and by other groups [P. E. Stephens & G. C. G. Hentschel, unpublished work; 66,67]. The reason for the difference in the level of induction is unclear but could be due to altered chromatin structure of the cloned gene in BPV, insufficient quantities of regulatory factors or differences in mRNA stability. Karin and co-workers have studied the human metallothionein (hMT) gene family and most specifically the hMTIa and hMTIIa genes [58,63]. In gene fusion experiments using the *tk* gene, they also found that the hMTIa promoter was only responsive to Cd^{2+}, whereas the hMTIIa promoter could be induced by Cd^{2+}, Zn^{2+} and glucocorticoids. The loss of glucocorticoid induction of the metallothionein I promoter has also been seen with other eukaryotic expression systems [77,78]. It is possible that this loss of regulation results from a changed chromosome environment around the metallothionein gene [77–79]. However, in two other examples in which glucocorticoid-responsive promoters were used with BPV vectors, induction by dexamethasone was retained [80,81].

Studies using the human β-interferon gene indicated that when linked to its own promoter on an extrachromasomally replicating BPV plasmid the gene is under normal regulatory control. Expression can be induced both by double-stranded RNA and inactivated Newcastle disease virus [60,82,83]. The constitutive levels of expression are somewhat higher than seen in human cells, possibly due to additional upstream start sites for the transcripts [60,83]. In initial experiments Zinn *et al.* [60] transformed host cells with a linear BPV plasmid from which the bacterial sequences had been removed. The resultant cell lines showed a large variability in the level of β-interferon expression and the extent of induction. This coincided with the fact that the BPV DNA in these cell lines appeared to be rearranged and in some cases complexed into high-M_r forms [60]. A second generation of BPV vectors incorporating the β-globin gene and bacterial sequences which did not require cleavage gave much more consistent values for β-interferon expression [62]. Using this vector Zinn *et al.* mapped two sequences in the 5′ flanking region of the β-interferon gene which regulated its constitutive and inducible expression. Not all aspects of the control of β-interferon expression appear to be the same, when linked to BPV, compared with its normal expression in human cells. With endogenous expression cycloheximide has the effect of superinducing the levels of β-interferon produced after stimulation by double-stranded RNA. However, in the BPV system it appears that cycloheximide can have its effect in the absence of double-stranded RNA [82].

The use of the BPV system was extended by Green *et al.* to the study of the regulation of the human histone gene H_4 during the cell cycle [84]. The histone gene, containing 650 nucleotides of 5′ and 900 nucleotides of 3′ flanking sequences, when linked to the 69% transforming segment of BPV was maintained as an episome at 20 copies per cell. When the cells were synchronized, by two cycles of thymidine block, the human histone gene was regulated co-ordinately with DNA replication. There was a clear preferential expression during the S phase of the cell cycle followed by selective turnover after the end of S phase. Thus, the episomal maintenance of the histone gene by the BPV genome allows the study of cell-cycle-regulated gene expression without the influence of the site of chromosomal integration.

In addition to the study of gene regulation the BPV system has been used to analyse the mutation frequency in transfected genes [85]. Ashman & Davidson constructed a BPV vector containing the entire BPV genome, pML, and the *gpt* gene. They rescued the episomal plasmid back into *E. coli* and measured the mutation frequency in the *gpt* gene by differential selection. This indicated a high mutation frequency of $(3–16) \times 10^{-3}$ and restriction enzyme analysis of the rescued plasmids confirmed the frequent presence of gross DNA rearrangements. To what extent this high rate was a consequence of *gpt*'s intrinsic instability in the BPV system [44] is not clear. It certainly does not appear to be a general property of BPV vector systems since some cell lines exist with apparently unmutated BPV vectors even after hundreds of generations of continuous culture.

The use of the BPV vector system for high-level expression of foreign genes

To date, BPV-based vectors have been used to create stable cell lines producing at least 26 different proteins. These range from the bacterial proteins for chlor-

amphenicol acetyltransferase [80] and aminoglycoside phosphotransferase [44,71,72] to the human proteins growth hormone [76], tissue plasminogen activator [86] and α, β and γ interferon [60,65,82]. They also include the viral glycoproteins hepatitis B surface antigen, HBSag [12,66,87,88], influenza virus haemagglutinin and cap-recognizing proteins [61,68] and type 1 human T-cell leukaemia virus small envelope protein [89]. The system has in addition been used to express normal cell surface markers such as the heavy chain of an HLA human histocompatability antigen [64] and the human interleukin 2 receptor [90]. Table 2 gives a list of proteins for which information on specific production rates is available when produced in the BPV system.

Several examples in Table 2 illustrate the ability of the BPV–C127 system to process foreign proteins correctly, allowing normal post-translational modification and secretion. A good example of this is the production of hepatitis surface antigen (HBSag) [12,66,87,88]. The HBSag is secreted into the medium as a 22 nm particle with a buoyant density of 1.2 g/ml. These particles have the same lipoprotein and polypeptide composition and the same electron microscopic appearance as 22 nm particles purified from the serum of chronic hepatitis carriers. Similarly, the expression of tPA in C127 cells is another example of the ability of the system to produce large complex molecules in an active form. tPA is 563 amino acids long with a hydrophobic 'pro' sequence of 12–15 amino acids. Mature tPA has a molecular mass of 58 kDa, has 32 cysteine residues capable of forming disulphide bonds and four potential glycosylation sites [91]. tPA produced in the BPV system appears to be correctly processed and glycosylated and is active as measured in fibrinolysis assays [74].

In addition to the production of β-interferon, both α5 and γ interferon have been expressed in a BPV system from the SV40 early promoter. All three interferons were correctly processed and secreted; interferon γ was also glycosylated. All the interferons had antiviral activity and α5 and γ were neutralized by antibodies raised to natural human interferons [65].

The system is not only capable of secreting correctly processed proteins but can also be used to analyse factors anchoring proteins to the cell surface. For example, 5×10^6 molecules of native influenza virus haemagglutinin accumulate in the cell membrane, but if the anchorage sequences are removed it is secreted [61]. A similar pattern of cell surface expression has been seen for the vesicular stomatitis virus glycoprotein [92,93]. Other cell surface markers include the human heavy chain of an HLA human histocompatibility antigen [64], the transmembrane envelope protein p21e of type 1 human T-cell leukaemia virus [89] and the interleukin 2 receptor [90]. While such studies indicate the utility of BPV expression vector in studying membrane events, the host cell range may limit the potential in this area. For example, only a low-affinity interleukin 2 receptor was seen to be expressed in C127 cells in contrast with the T-cell lymphoma HUT-102 cell line, from which the gene was cloned, which has both high and low affinity receptors [90].

There appears to be little intrinsic difference in the levels of expression in the differing vectors shown in Fig. 2; all are capable of producing up to 10^8 molecules per cell per day [61,66,76]. The principal factor determining the level of expression appears to be the choice of promoter. Apart from homologous promoters the mouse metallothionein (mMTI) or the SV40 early promoters have been most widely used. The mMTI promoter is reported to be considerably stronger than the SV40 early promoter [62]. Observations from this laboratory also suggest that the metallothionein promoter is better than viral LTRs from Rous sarcoma virus or Moloney murine leukaemia virus [86]. Admittedly the choice of promoter has been somewhat limited so far and it is possible that higher levels of expression could be achieved with stronger promoters.

It is clear from the data presented in Table 2 and from other reports describing the use of BPV as a vector that the episomal nature of the viral genome in transformed cells is not as universal as first thought. The viral DNA can exist in the host in one of several forms; episomal monomers, episomal concatomers, concatenates, integrated concatomers or multiple integrated monomers [52,60,61,63,64,40,80]. What determines the form of the vector DNA is not clear, but probably reflects the complex interaction of the viral genes and the inserted genes. Lusky & Botchan have recently reported the presence of an enhancer element in the 31% non-transforming region which is believed to play a role in the episomality and replication of the viral genome [37, and see the previous section]. Vectors which have this enhancer removed (i.e. vectors using just the 69% region) appear to be much more prone to integration and rearrangement [60,83,88]. Furthermore, if bacterial sequences are linked to the 69% region then this outcome is even more likely [61,84]. Removal of these bacterial sequences prior to transfection tends to restore the episomality of the vector and leads to less rearrangement provided that the linear plasmid is recircularized before transfection [76,80,81]. The lost enhancer function could to a certain extent, be restored by 'stimulatory segments' [51,53,63] and vectors containing these sequences tended to remain episomal [62,63,65]. There are however instances where this is not the case [64]. Linkage of bacterial sequences to the whole genome appears to have a less profound effect on the vector DNA in the transformed cell lines, and vectors using the whole genome have been used with some success [44,66,89]. However, this is not always the case and some integration and rearrangements have been reported [71,94]. The replication of the viral genome is controlled by the E1 and E7 gene products in conjunction with an enhancer in the 31% region (see the previous section). Presumably any factor affecting these functions, for example the insertion of a foreign transcription unit, affects viral replication. This is most clearly illustrated by Matthias *et al.* and Schenborn *et al.* [71,94]. They found that the insertion of transcription units in different positions and orientations in their vectors resulted in radically different fates of their plasmid DNA.

Gething *et al.* [61] have used fluorescence-activated cell sorting to analyse the variation in the levels of production of influenza virus haemagglutinin in cloned cell lines. The surface fluorescence varied by 30-fold in a unimodal fashion. They proposed that this variation could be due to the state of growth of the cells or to a variation in expression as the cell passes through different phases in the cell cycle [61]. Cosman *et al.* [90] have also used an activated cell sorter to fractionate cells on the basis of expression of a cell surface marker, the interleukin 2 receptor. Using this method cell lines highly enriched for

Table 2. Specific production rates of heterologous proteins produced using BPV vectors

The specific production rates are calculated by using the following assumptions. Interferon γ, M_r 19.5 kDa and specific activity 1×10^8 units/mg [99]; HBSag, M_r 22 kDa; rat proinsulin, M_r 8.4 kDa; bovine growth hormone, M_r 22 kDa [91]; interferon β, M_r 22 kDa and specific activity 1×10^8 units/mg [100]; interferon $\alpha 5$, M_r 19.5 kDa and specific activity $= 2 \times 10^8$ units/mg [99]; TIMP, M_r 28 kDa and 10^6 cells/ml; tPA, M_r 68 kDa and 10^6 cells/ml. Specific production rates are expressed as the number of protein molecules produced/cell per day. Vector category refers to the three vectors shown in Fig. 2. Abbreviations used: HBSag, hepatitis B surface antigen; MMT, mouse metallothionein-I; SV40E, early promoter from simian virus 40; HHSP70, human heat shock protein 70; NA, not available.

Gene	Promoter	Copy number	Specific production rate (molecules/cell per day)	Vector category	Characteristics	Notes	Reference
Interferon β (human)	Own	30–50	2.5×10^5	1	69% no bacterial sequence	Some rearrangement; upstream start of mRNA; 130-fold induction	83
Interferon β (human)	Own	30–50	1.9×10^6	1	69% no bacterial sequence	Cycloheximide induction 8-fold	82
Interferon β (human)	Own	30–60	1.0×10^5	1	69% no bacterial sequence	Extrachromosomal multimers; 16-fold induction	60
Interferon β (human)	Own	20–30	6.3×10^5	2	69% + pBRD + β-globin	Extrachromosomal DNA; little cell-to-cell variability; deletion analysis of promoter	62
HBSag	Own	50–100	1.6×10^7	1+2	69% + pBR322; no β-globin sequences	High-M_r; DNA not integrated; NIH 3T3 host cells	87
HBSag	Own	20–200	5.5×10^6	1	100% BPV; no bacterial sequence	Rearranged but remains extrachromosomal	12
HBSag	Own	50–100	2.4×10^6	1	69% BPV; no bacterial sequence	Expression appears to be orientation-dependent	88
HBSag	mMT	5–180	5.5×10^7	3	100% BPV + pML/MMT sandwich	Stable for 85 days; no induction	66
Proinsulin	Own	60–80	1.4×10^3	1	69% of BPV; no bacterial sequence	The majority of DNA is unintegrated; some evidence for high-M_r DNA and rearrangements	52
Growth hormone (human)	mMT	10–100	6×10^8	1	69% of BPV; no bacterial sequence	Extrachromosomal DNA; cell line stable for 10 months; induced 2–5-fold by Cd^{2+}	76
Growth hormone (bovine)	mMT	NA	1.4×10^8	1	69% of BPV; no bacterial sequence	1.5-fold induction by Cd^{2+}	95
Influenza virus haemagglutinin	SV40E or mMT	30–200	10^6–10^7	2	69% of BPV + β-globin + pBR322	mMT promoter 10-fold better than SV40 early promoter; no induction by heavy metals	61
Interferon γ (human)	SV40E	30–50	13.6×10^6	2+1	69% BPV + β-globin; no pBR322	Normal glycosylation; unrearranged extra chromosomal DNA; loss of initial cleavage site on recircularization	65
Interferon $\alpha 5$ (human)	SV40E	30–50	13.6×10^6	2+1	69% BPV + β-globin; no pBR322		
Chorionic gonadotrophin α subunit (human)	mMT	NA	8.2×10^6	3	100% BPV + pML + mMT sandwich	Extrachromosomal DNA; glycosylation pattern similar to that seen in ectopically produced protein; 2–5-fold induction	67
Tissue inhibitor of metalloproteinase (TIMP) (human)	mMT	NA	1.1×10^7	3	100% BPV + pAT153 + functional mMT gene	Normal glycosylation pattern; initial selection for heavy-metal-resistant foci	75

Table 2. (*cont.*)

Gene	Promoter	Copy number	Specific production rate (molecules/cell per day)	Vector category	Characteristics	Notes	Reference
Tissue plasminogen activator (tPA) (human)	HHSP70	30–90	1.1×10^7	3	100% BPV + pAT153 + functional MMT gene	High-M_r unrearranged; up to 40-fold induction after heat shock; initial selection for heavy-metal-resistant foci	74
Tissue plasminogen activator (tPA) (human)	MMT	10–150	2.7×10^7	3	100% BPV + pAT153 + functional MMT gene	High-M_r DNA unrearranged; TPA correctly processed; initial selection of heavy-metal-resistant foci	86

interleukin 2 receptors could be isolated [90]. The reason for the clonal variation is not always clear but probably reflects position effects due to integration of the BPV vectors. It is also true that establishment of cell lines with a linear transforming molecule leads to a greater variability in expression levels due to rearrangements [60,65,83,95]. From the point of view of producing large amounts of a desired protein this clonal variability is not a problem but a benefit since it allows selection for rare high-producing cell lines. Moreover, in these cases integration may also be beneficial in producing more stable cell lines.

Few definitive examples exist which allow the direct comparison of the efficiency of gene expression in different systems. Those which do exist have been reviewed recently by Bebbington & Hentschel [96]. When compared with two commonly used expression systems, namely the Chinese hamster ovary (CHO) cell and co-amplification system or the COS cell transient expression system, the BPV system does not always appear particularly productive. The CHO and COS cell systems are, for example, capable of producing $(1–3.3) \times 10^8$ molecules per cell per day of hepatitis B surface antigen from its own promoter, whereas the BPV system gives only $(3–24) \times 10^6$ molecules per cell day. However, the CHO and COS cell systems do have major drawbacks which do not occur in the BPV system. The high levels of expression seen in CHO cells are only apparent after extensive periods of gene amplification, whereas the COS cell system is transient with the transformed cells only remaining viable for 2–6 days before they are killed.

FUTURE PROSPECTS AND CONCLUSIONS

As the molecular biology of the virus becomes better understood BPV vectors can be designed to achieve a number of desirable characteristics. It is, for example, clear that the transforming genes and the genes which regulate replication can be separated and this might allow the establishment of high copy number BPV vectors expressing foreign protein in a non-transformed host. Furthermore, as our understanding of the mechanisms of control of replication increase two other possibilities arise. Firstly, putting these functions under different control sequences might allow their expression in so far untested novel cell lines, thus extending the host range. Secondly, placing key regulatory genes under a controllable promoter should allow the copy number and consequently the level of expression to be controlled. Examples of this approach already exist for both polyomavirus and SV40 vectors [97,98]. The construction of streamlined vectors with minimal viral DNA is also possible. It has already been demonstrated that the possession of a plasmid maintenance sequence (PMS) on a vector is all that is required to get extrachromosomal maintenance of the vector in a host which has already been transformed with wild-type BPV [15]. If stable cell lines are produced which possess all of the necessary *trans*-acting factors stably integrated into the genome they should be able to maintain episomal replication of streamlined BPV plasmids, possessing only the *cis* required sequences. This sort of system would clearly facilitate the construction of vectors.

It seems probable that the true versatility of the BPV system as a shuttle vector has not been fully utilized as yet. Only one example of shuttling between animal and bacterial cells to examine a biological phenomenon has been reported to date [85]. The ability of the BPV system to allow plasmids to be shuttled between bacteria and animal cells could be further utilized in two ways. Firstly, genes which affect cellular phenotypes can be analysed by passaging mutants between animal and bacterial cells. Secondly, gene libraries made in BPV expression vectors could be introduced into animal cells and cloned cell lines could be isolated which produce the desired phenotype or biological activity. The plasmid DNA could then be rescued from these cells and the desired gene isolated.

The versatility of the BPV as a vector system has been clearly demonstrated in this Review. It has been used to analyse several questions concerning biological phenomena, and as our understanding of the virus increases it is clear that its use here will be extended. It is also evident that the BPV system is capable of producing large amounts of protein per cell. Unfortunately, all of the host cells used to date are attachment cells. This potential drawback for large scale culture could possibly be overcome in the future by the development of suspension variants of C127 cells or conversely by expanding the host range to a genuine suspension cell.

We would like to thank Gill Sedgewick and Paula Palethorpe for typing the manuscript, and Alan Lyons for his assistance in the preparation of the Figures.

REFERENCES

1. Harris, T. J. R. (1983) in Genetic Engineering (Williamson, R., ed.), vol. 4, pp 127–185, Academic Press, London
2. Marston, F. A. O. (1986) Biochem. J. **240**, 1–12
3. Giri, I. & Danos, O. (1986) Trends Genet. **2**, 227–232
4. Lancaster, W. D. & Olson, C. (1982) Microbiol. Rev. **46**, 191–207
5. Pfister, H. (1984) Rev. Physiol. Biochem. Pharmacol. **99**, 111–181
6. Chen, E. Y., Howley, P. M., Levinson, A. D. & Seeburg, P. H. (1982) Nature (London) **299**, 529–534
7. Stenlund, A., Zabielski, J., Ahola, H., Moreno-Lopez, J. & Pettersson, U. (1985) J. Mol. Biol. **182**, 541–554
8. Amtmann, E. & Sauer, G. (1982) J. Virol. **43**, 59–66
9. Lancaster, W. D. (1981) Virology **108**, 251–255
10. Lowy, D. R., Dvoretzky, I., Shober, R. L. & Howley, P. M. (1980) Nature (London) **287**, 72–74
11. Danos, O., Engel, L. W., Chen, E. Y., Yaniv, M. & Howley, P. M. (1983) J. Virol. **46**, 557–566
12. Stenlund, A., Lowy, D., Moreno-Lopez, J., Ahola, M., Pettersson, H. & Tiollas, P. (1983) EMBO J. **2**, 669–673
13. Engel, L. W., Heilman, C. A. & Howley, P. M. (1983) J. Virol. **47**, 516–528
14. Pilacinski, W. P., Glassman, D. L., Krzyzek, R. A., Sadowski, P. L. & Robbins, A. K. (1984) Bio/Technology **1**, 356–360
15. Lusky, M. & Botchan, M. R. (1984) Cell **36**, 391–401
16. Roberts, J. M. & Weintraub, H. (1986) Cell **46**, 741–752
17. Berg, L., Lusky, M., Stenlund, A. & Botchan, M. R. (1986) Cell **46**, 753–762
18. Lusky, M. & Botchan, M. R. (1985) J. Virol. **53**, 955–965
19. Nakabayashi, Y., Chattopadhay, S. K. & Lowy, D. R. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 5832–5836
20. Sarver, N., Rabson, M. S., Yang, Y.-C., Byrne, J. C. & Howley, P. M. (1984) J. Virol. **52**, 377–388
21. Seif, I. (1984) Virology **138**, 347–352
22. Clertant, P. & Seif, I. (1984) Nature (London) **311**, 276–279
23. Yang, Y.-C., Okayama, H. & Howley, H. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1030–1034
24. DiMaio, D. (1986) J. Virol. **57**, 475–480
25. Yang, Y.-C., Spalholz, B. A., Rabson, M. S. & Howley, P. M. (1985) Nature (London) **318**, 475–577
26. Spalholz, B. A., Yang, Y.-C. & Howley, P. M. (1985) Cell **42**, 183–191
27. Mallon, R. G., Sisk, W. & Defendi, V. (1986) Gene **42**, 241–251
28. DiMaio, D., Guralski, D. & Schiller, J. T. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 1791–1801
29. Schiller, J. T., Vass, W. C., Vousden, K. H. & Lowy, D. R. (1986) J. Virol. **57**, 1–6
30. Schlegel, R., Wade-Glass, M., Robson, M. S. & Yang, Y.-C. (1986) Science **233**, 464–467
31. Schiller, J. T., Vass, W. C. & Lowy, D. R. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 7880–7884
32. Androphy, E. J., Schiller, J. T. & Lowy, D. R. (1985) Science **230**, 442–445
33. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. (1985) Cell **40**, 9–17
34. Berg, L. J., Singh, K. & Botchan, M. R. (1986) Mol. Cell. Biol. **6**, 859–869
35. Heilman, C. A., Engel, L., Lowy, D. R. & Howley, P. M. (1982) Virology **119**, 22–34
36. Campo, M. S., Spandidos, D. A., Lang, J. & Wilkie, N. M. (1983) Nature (London) **305**, 77–80
37. Lusky, M. & Botchan, M. R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 3609–3613
38. Lusky, M., Berg, L., Weiher, H. & Botchan, M. (1983) Mol. Cell. Biol. **3**, 1108–1122
39. Spandidos, D. A. & Wilkie, A. M. (1983) EMBO J. **2**, 1193–1199
40. Law, M.-F., Lowy, D. R., Dvoretzky, I. & Howley, P. M. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 2727–2731
41. Antmann, E., Muller, H. & Sauer, G. (1980) J. Virol. **35**, 962–964
42. Allshire, R. C. & Bostock, C. J. (1986) J. Mol. Biol. **188**, 1–13
43. Sarver, N., Byrne, J. C. & Howley, P. M. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 7147–7151
44. Law, M.-F., Byrne, J. C. & Howley, P. M. (1983) Mol. Cell. Biol. **3**, 2110–2115
45. Waldeck, W., Rosl, F. & Zentgraf, H. (1984) EMBO J. **3**, 2173–2178
46. Rosl, F., Waldeck, W. & Sauer, G. (1983) J. Virol. **42**, 567–574
47. Meischke, H. R. C. (1979) J. Gen. Virol. **43**, 473–487
48. Olson, C., Gordon, D. E., Robl, M. G. & Lee, K. P. (1969) Arch. Environ. Health **19**, 827–837
49. Dvoretzky, I., Shober, R., Chattopadhay, S. K. & Lowy, D. R. (1980) Virology **103**, 369–375
50. Lowy, D. R., Rands, E. & Scolnick, E. M. (1978) J. Virol. **26**, 291–298
51. DiMaio, D., Treisman, R. & Maniatis, T. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 4030–4034
52. Sarver, N., Gruss, P., Law, M.-F., Khoury, G. & Howley, P. M. (1981) Mol. Cell. Biol. **1**, 486–496
53. Kushner, P. J., Levinson, B. B. & Goodman, H. M. (1982) J. Mol. Appl. Genet. **1**, 527–538
54. Howley, P. M., Schenborn, E. T., Lund, E., Byrne, J. C. & Dahlberg, J. E. (1985) Mol. Cell. Biol. **5**, 3310–3315
55. Lusky, M. & Botchan, M. R. (1981) Nature (London) **293**, 79–81
56. Campo, M. S. & Spandidos, D. A. (1983) J. Gen. Virol. **64**, 549–557
57. Twigg, A. J. & Sherratt, D. (1980) Nature (London) **283**, 216–218
58. Richards, R. I., Heguy, A. & Karin, M. (1984) Cell **37**, 263–272
59. Binetruy, B., Meneguzzi, G., Breathnach, R. & Cuzin, F. (1982) EMBO J. **1**, 621–628
60. Zinn, K., Mellon, P., Ptashne, M. & Maniatis, T. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 4897–4901
61. Sambrook, J., Rodgers, L., White, J. & Gething, M. J. (1985) EMBO J. **4**, 91–103
62. Zinn, K., DiMaio, D. & Maniatis, T. (1983) Cell **34**, 865–879
63. Karin, M., Cathala, G. & Chi Nguyen-Huu, M. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 4040–4044
64. DiMaio, D., Corbin, V., Sibley, E. & Maniatis, T. (1984) Mol. Cell. Biol. **4**, 340–350
65. Fukunaga, R., Sokawa, Y. & Nagata, S. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 5086–5090
66. Hsiung, N., Fitts, R., Wilson, S., Milne, A. & Hamer, D. (1984) J. Mol. Appl. Genet. **2**, 497–506
67. Ramabhadran, T. V., Reitz, B. A. & Teimeier, D. C. (1984) Proc. Natl. Acad. Sci. U.S.A. **81**, 701–705
68. Braam-Markson, J., Jaudon, C. & Krug, R. M. (1985) Proc. Natl Acad. Sci. U.S.A. **82**, 4326–4330
69. Sekiguchi, T., Nishimoto, T., Kai, R. & Sekiguchi, M. (1983) Gene **21**, 267–272
70. Giri, I., Jonanneau, J. & Yaniv, M. (1983) Virology **127**, 385–396
71. Matthias, P. D., Bernard, H. U., Scott, A., Brady, G., Hashimoto-Gotoh, T. & Shutz, G. (1983) EMBO J. **2**, 1487–1492

72. Meneguzzi, G., Binetruy B., Grisoni, M. & Cuzin, F. (1984) EMBO J. **3**, 365–371
73. Pavlakis, G. N. & Hamer, D. H. (1983) Recent Progr. Horm. Res. **39**, 353–385
74. Bendig, M. M., Stephens, P. E., Cockett, M. I. & Hentschel, C. C. G. (1987) DNA, in the press
75. Docherty, A. J. P., Lyons, A., Smith, B. J., Wright, E. M., Stephens, P. E. & Harris, T. J. R. (1985) Nature (London) **318**, 66–69
76. Pavlakis, G. N. & Hamer, D. H. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 397–401
77. Mayo, K. E. & Palmiter, R. D. (1982) J. Biol. Chem. **257**, 3061–3067
78. Mayo, K. E., Warren, R. & Palmiter, R. D. (1982) Cell **29**, 99–108
79. Searle, P. F., Davidson, B. L., Stuart, G. W., Thomas, M. W., Norstedt, G. & Palmiter, R. D. (1984) Mol. Cell. Biol. **4**, 1221–1230
80. Matthias, P., Boeger, U., Danesch, U., Schutz, G. & Bernard, H.-U. (1986) J. Mol. Biol. **187**, 557–568
81. Ostrowski, M. C., Richard-Foy, H., Wolford, R. G., Berard, D. S. & Hager, G. L. (1983) Mol. Cell. Biol. **3**, 2045–2057
82. Maroteaux, L., Chen, L., Mitrani-Rosenbaum, S., Howley, P. M. & Revel, M. (1983) J. Virol. **47**, 89–95
83. Mitrani-Rosenbaum, S., Maroteaux, L., Mory, Y., Revel, M. & Howley, P. M. (1983) Mol. Cell. Biol. **3**, 233–240
84. Green, L., Schlaffer, I., Wright, K., Moreno, M. L., Berand, D., Hager, G., Stein, J. & Stein, G. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2315–2319
85. Ashman, C. R. & Davidson, R. L. (1985) Somatic Cell Mol. Genet. **11**, 499–504
86. Stephens, P. E., Bendig, M. M. & Hentschel, C. C. G. (1985) J. Cell. Biochem. Suppl. **9C**, 95
87. Wang, Y., Stratowa, C., Schaefer-Ridder, M., Doehmer, J. & Hofschneider, P. H. (1983) Mol. Cell. Biol. **3**, 1032–1039
88. Denniston, K. J., Yoneyama, T., Hoyer, B. H. & Gerin, J. L. (1984) Gene **32**, 357–368
89. Eiden, M., Newman, M., Fisher, A. G., Mann, D. L., Howley, P. M. & Reitz, M. S. (1985) Mol. Cell. Biol. **5**, 3320–3324
90. Cosman, D., Wignall, J., Lewis, A., Alpert, A., Cernetti, D. P., Park, L., Dower, S. K., Gillis, S. & Urdal, D. L. (1986) Mol. Immunol. **23**, 935–941
91. Vehar, G. A., Kohr, W. J., Bennett, W. F., Pennica, D., Ward, C. A., Harkins, R. N. & Collen, D. (1984) Bio/Technology **2**, 1051–1057
92. Florkiewicz, R. Z., Smith, A., Bergmann, J. E. & Rose, J. K. (1983) J. Cell. Biol. **97**, 1381–1388
93. Florkiewicz, R. Z. & Rose, J. K. (1983) Science **225**, 721–723
94. Schenborn, E. T., Lund, E., Mitchen, T. L. & Dahlberg, T. E. (1985) Mol. Cell. Biol. **5**, 1318–1326
95. Ramabhadran, T. V., Reitz, B. A. & Shah, D. H. (1985) Gene **38**, 111–118
96. Bebbington, C. R. & Hentschel, C. C. G. (1985) Trends Biotechnol. **3**, 314–317
97. Kern, F. G. & Basilico, C. (1986) Gene **43**, 237–245
98. Rio, D. C., Clart, S. G. & Tijan, R. (1985) Science **227**, 23–28
99. Svedersky, L. P., Benton, C. V., Berger, W. H., Rinderknecht, E., Harkins, R. N. & Palladino, M. A. (1984) J. Exp. Med. **59**, 812–827
100. Devos, R., Opsomer, C., Scahill, S. J., Van Der Heyden, J. & Fiers, W. (1984) J. Interferon Res. **4**, 461–468