Review article

Some highlights of virus research in 1990

Richard M. Elliott,* Norman E. Crook, Ulrich Desselberger, Roger Hull and Duncan J. McGeoch

1Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, 2Horticultural Research International, Worthing Road, Littlehampton, West Sussex BN17 6LP, 3Regional Virus Laboratory, East Birmingham Hospital, Birmingham B9 5ST and 4John Innes Institute, AFRC Institute of Plant Science Research, Colney Lane, Norwich NR4 7UH, U.K.

Introduction

The ‘Highlights’ review has become an annual feature of this Journal, and its aim is to present an overview of selected advances in virology made during the previous year. We stress that the choice of topics is highly subjective and does not purport to be exhaustive, but reflects the interests of the authors concerned. We start with a section dealing with newly characterized animal and plant viruses, and then move on to deal with complete viral genome sequences. Next we report on human immunodeficiency virus and the structure of CD4, advances in baculovirology, the manipulation of RNA virus genomes, novel mechanisms of resistance to virus infection in plants, and finish with a section on the use of transgenic mice in virological research.

Newly characterized viruses and novel virus types

A number of disease syndromes are well documented and of suspected viral aetiology long before the agent responsible is identified, isolated or characterized. A good example is hepatitis C virus (HCV), which we discussed last year (Elliott et al., 1990), where nucleotide sequences of cloned cDNAs derived from HCV-positive plasma suggested that the virus had a positive-strand RNA genome. More extensive nucleotide sequence data from Japanese isolates of HCV have recently been obtained (Kato et al., 1990; Takamizawa et al., 1991), and although the virus has still to be isolated, the sequence data suggest it is related to the Flaviviridae and in particular to members of the Pestivirus genus.

In 1990 a cDNA clone from the virus responsible for enterically transmitted non-A non-B hepatitis (ET-NANBH) was described (Reyes et al., 1990). ET-NANBH is transmitted by the faecal–oral route and is a major public health hazard in developing countries where it causes epidemic outbreaks of viral hepatitis. In addition it is particularly hazardous to pregnant women, with mortality reported to be as high as 20%. Reyes et al. made randomly primed cDNA to RNA extracted from gall bladder bile of an experimentally infected cynomolgus macaque, cloned this into λgt10 and identified 16 putative ET-NANBH cDNAs by differential hybridization with uninfected and infected bile-derived cDNA probes. Nucleotide sequence analysis of one clone yielded a predicted amino acid sequence with similarities to the RNA-dependent RNA polymerase motifs shared by plus-strand RNA viruses. The clone specifically hybridized to a polyadenylated 7·6 kb RNA found in infected, but not uninfected, livers of macaques, and identified cDNA sequences prepared from infected human faecal samples. The virus has been called hepatitis E virus. The data are consistent with the previous suggestion that ET-NANBH is caused by a calicivirus (Bradley & Balayan, 1988).

Jiang et al. (1990) provided evidence that Norwalk virus is also a possible calicivirus. Norwalk and Norwalk-like viruses are a major cause of acute gastroenteritis in the U.S.A., often associated with contaminated shellfish. Although first discovered in 1973, the study of Norwalk virus has been impeded by the lack of a permissive cell culture system; human faecal samples have been the only source of virus. To obtain Norwalk virus nucleic acid for cloning, purified virions were obtained from the stools of infected volunteers. Randomly primed cDNA was cloned into a plasmid vector, and positive clones were identified by their
pattern of hybridization to extracts of stool samples from pre- and post-infection volunteers. Treatment of the test samples with RNase abolished positive hybridization, indicating that the virus had an RNA genome. Further hybridization studies showed the genome to be polyadenylated, plus-stranded and about 7.5 kb in length, and limited sequence analysis identified the polymerase motifs of positive-strand RNA viruses.

A third example of a newly characterized calici-like virus is provided by haemorrhagic disease of rabbits (RHD). This disease was first reported in China in 1984 (Liu et al., 1984) and then emerged in Europe causing high mortality in commercial rabbit colonies. The identity of the agent responsible had been the subject of controversy, but two recent papers gave firm evidence that it is a calici-like virus. Both Parra & Prieto (1990) and Ohlinger et al. (1990) purified virus particles from infected rabbit livers by gradient centrifugation; the particles were unenveloped, 27 to 40 nm in diameter and morphologically similar to caliciviruses in having cup-shaped depressions on the surface. The major structural protein had an Mr of 60 000. Ohlinger et al. further demonstrated that the purified virus was infective for rabbits. Both groups extracted a single RNA species of 7.4 to 8 kb from virion preparations. Ohlinger et al. showed that this RNA produced disease when injected into rabbit liver. cDNA prepared to the virion RNA detected RNA species of 8 and 2 kb in infected rabbit liver. Taken together, these data clearly indicate that RHD is caused by a calicivirus.

At present the caliciviruses are rather poorly understood, particularly at the molecular level, and the identification of potential members of the family associated with important human diseases will no doubt be the impetus for further studies.

Borna disease (BD) is a chronic neurological condition characterized by severely abnormal behaviour patterns in naturally affected horses and sheep. Some reports suggest BD may play a role in human neuropsychiatric diseases. The disease may be transmitted to experimental animals via inoculation of filtered brain homogenates from affected animals. Because treatment of homogenates with detergents reduced the capacity to induce disease, the causative agent was suspected to be an enveloped virus. Lipkin et al. (1990) made a cDNA library to polyadenylated RNA extracted from an affected rat brain, and by differential hybridization obtained seven BD agent clones. These clones hybridized to RNA transcripts of 8.5, 2.0 and 0.8 kb specifically in affected rat brain, and one clone arrested translation in vitro of BD-specific proteins. The two small RNA transcripts were shown to be polyadenylated and complementary to the 8.5 kb RNA. In a subsequent paper, de la Torre et al. (1990) identified the same transcripts in a cell line persistently infected with BD agent; interestingly, the 8.5 kb transcript was found to be abundant in the nuclei. Although further characterization is needed, de la Torre et al. favour the notion that BD is caused by a single-stranded, negative-sense RNA virus with a nucleus-associated 8.5 kb genome encoding at least two mRNAs.

The pathogenesis of BD virus (BDV) has been studied by Richt et al. (1990), who established a CD4+ T cell line reactive to BDV antigen (extracted from BDV-affected brain). Passive transfer of these cells into healthy, immunosuppressed BDV-infected rats resulted in full-blown disease. Immunohistological examination of diseased brains revealed an abundance of macrophages and CD4+ T cells, which Richt et al. take as good evidence that a delayed hypersensitivity reaction is involved in the pathogenic mechanism of BD.

Initial characterization of a new 'family' of insect viruses together with studies on their histopathology and pathogenicity was described by B. Federici and co-workers. The group has been tentatively named 'ascoviruses' in reference to the unique virion-containing vesicles formed by disruption of host cells. The virions are large (400 × 130 nm), enveloped, and contain an inner lipid bilayer surrounding the core which contains a large dsDNA genome (Federici et al., 1990). Of the three isolates examined, those from Heliothis zea (HAV) and Trichoplusia ni (TAV) were closely related but not identical, whereas the third, from Spodoptera frugiperda (SAV), was quite distinct with a much lower G+C content in its DNA and no detectable hybridization to Southern blots of either of the other DNAs. Differences in tissue tropisms were also noted (Federici & Govindarajan, 1990); replication of HAV and TAV was observed in a number of tissues, but SAV replication was found only in the fat body. The transmission of these viruses appears to be unusual as feeding results in only low levels of infection whereas injection results in very high levels of infection (Govindarajan & Federici, 1990), suggesting that the normal mode of infection may be transmission by parasites. These viruses may be relatively common but the lack of obvious signs of disease makes them difficult to detect in the field. Now that the basic characteristics are known, more examples of this group of viruses may be isolated.

Tomato spotted wilt virus (TSWV) is of increasing importance because of its very wide host range and recent rapid spread world-wide. It forms a monotypic group with properties unique among plant viruses. It is the only virus transmitted by thrips (in the circulative manner) and has spherical enveloped particles of 80 to 110 nm diameter. These comprise four structural proteins, nucleocapsid (N, Mr 29 000), two glycoproteins (G1 and G2, Mr s 78 000 and 58 000), which are associated
with the viral membrane, and a large protein (L, Mr about 200000) which is present in low amounts. Particles contain three species of ssRNA (L, 8-2 kb; M, 5.4 kb; S, 2.9 kb) which associate with the N protein to form circular nucleocapsids. TSWV thus resembles the family Bunyaviridae in its virion and genomic structures. The determination of the S RNA sequence (de Haan et al., 1990) has reinforced suggestions that TSWV has close associations with the Bunyaviridae. S RNA comprises 2916 nucleotides with terminal complementarity of 65 to 70 nucleotides. It has two open reading frames (ORFs) in an ambisense arrangement which are expressed from subgenomic mRNAs. The ORF in the genomic RNA sense encodes a protein of Mr 52400 which is considered to be non-structural (NS); that in the complementary sense encodes the N protein (28-8K). The two ORFs are separated by a long stable hairpin structure which, it is suggested, controls the expression of the subgenomic RNAs. Thus the genome organization of TSWV S RNA closely resembles that seen in the Bunyaviridae and especially the Phlebovirus genus. Serological analyses have revealed further resemblances to phleboviruses in that the N protein appears to be the TSWV serogroup-specific antigen (Law & Moyer, 1990; de Ávila et al., 1990).

The NS protein has been expressed using a baculovirus vector (Kormelink et al., 1991) and an antiserum produced against it has confirmed that it is non-structural. In immunogold labelling of infected Nicotiana rustica leaves, NS appears in the cytoplasm as either elongated flexible filaments or large paracrystalline rod-shaped structures, depending on the virus isolate. By comparing different isolates of the virus it was shown that the amount of NS correlates with the severity of symptoms.

Another group of plant viruses, the tenuiviruses, also appears to show affinities to the Bunyaviridae. Tenuiviruses are transmitted by planthoppers and have filamentous particles which contain at least four RNA species. The full sequence of the smallest species, RNA 4, of rice stripe virus has been determined (Kakutani et al., 1990). The 2317 base RNA has two ORFs, again in an ambisense arrangement. The ORF in virion RNA sense encodes a non-structural protein of Mr 20500 which is a major virus protein in infected plants. The complementary sense ORF encodes a protein of Mr 32000 which has not yet been detected in infected tissues. The 20 bases at the 5' and 3' termini of each of the four RNA species are complementary (Kakutani et al., 1990; Takahashi et al., 1990) and also show sequence homology to termini of members of the Phlebovirus and Uukuvirus genera of the Bunyaviridae family. However, tenuiviruses differ from known Bunyaviridae in the structure of the virus particles and in the number of genomic components.

During 1990 it was observed that several plant viruses with non-enveloped bacilliform particles had dsDNA genomes (Lockhart, 1990); the name badnaviruses (BAcilliform DNA viruses) was proposed for this group. The nucleic acid sequence of the type member, Comme-lina yellow mottle virus (CoYMV), indicated that this is a plant pararetrovirus (Medberry et al., 1990). These viruses exhibit several features in common with the other plant pararetrovirus group, the caulimoviruses, but also some major differences. The common features include discontinuities at the putative priming sites for each DNA strand, the primer for minus-strand DNA synthesis being initiating tRNA^Met, and the production of a more-than-full length transcript. The major differences are in the genome organization. CoYMV has only three ORFs compared with the six of caulimoviruses. The putative products of the smaller two (Ms of 23000 and 15000) have no sequence homologies with any of the proteins of cauliflower mosaic virus. The putative product of the large CoYMV ORF (Mr 216000) has sequence motifs which suggest that it is a polypeptide comprising both the coat protein (gag analogue) and the polymerase (pol or reverse transcriptase). This is further and very strong evidence to suggest that CoYMV replicates via reverse transcription. The genome organization does, however, show a departure from those of most other reverse-transcribing viruses (both retro- and pararetroviruses) in that, although the polymerase is produced as a polypeptide with gag, its expression is controlled by either frameshifting or by readthrough of a termination codon.

The genome of vaccinia virus

Two very large virus DNA sequences were published in 1990, for vaccinia virus (Goebel et al., 1990a, b) and human cytomegalovirus (HCMV; Chee et al., 1990a), comprising 191 kbp and 229 kbp, respectively. These two sequences are the largest yet published for any virus genome or indeed, so far as we are aware, for any contiguous segment of DNA from any source. In this section we outline the results of the analysis of vaccinia virus DNA, and in the next section HCMV is described.

Vaccinia virus is the prototype of the orthopoxvirus group and is the first poxvirus to be completely sequenced. Goebel et al. (1990a) described determination of the genome sequence of the Copenhagen strain of the virus; the nucleotide sequence data are listed separately in Goebel et al. (1990b). Another vaccinia virus strain, Western Reserve (WR), is commonly used experimentally and substantial parts of its DNA sequence have been published in the last few years. The
Copenhagen strain sequence was determined predominantly by chain termination methodology using successively synthesized oligonucleotide primers to 'crawl' along the sequences of plasmid-cloned genomic fragments.

Vaccinia virus DNA is a linear, double-stranded molecule with a distinct structure at each terminus, namely a hairpin (of 101 nucleotides in the Copenhagen strain) which is incompletely base-paired, and which is linked 5' and 3' to the completely double-stranded interior of the genome; the whole viral DNA thus contains a covalently closed loop of a single strand. Adjacent to these telomere-like structures are a pair of large repeat elements in opposing orientation; in the Copenhagen strain these are 12 kbp in length. Within the 4-2 kbp next to each terminus there exist complex sets of short tandemly reiterated sequences, whereas most of the rest of the major repeat consists of protein-encoding DNA.

Goebel et al. (1990b) list the DNA sequence as a 5' to 3' single strand starting with the complete left hairpin and ending with the complete right hairpin. This listing contains 191737 residues, representing 191535 bp of completely dsDNA plus 101 nucleotides in each hairpin. The occurrence of variable copy numbers in the families of reiterated sequences in the major repeats means that genomic DNA molecules must exhibit a range of lengths, and other strains of vaccinia virus differ significantly in terms of the locations of genes and the nature of the encoded proteins. Such exercises have limitations, but vaccinia virus presents an unusually favourable case: the genome contains nine reading frames, so that 189 distinct encoded proteins are predicted.

Functions have been assigned from published work and on the basis of evaluation of amino acid sequence similarities by database searches. Together these give known or suggested roles for fewer than 50 of the genes. Many genes specifying proteins involved in transcription and in DNA replication have been analysed previously. In 1990 further characterization of genes encoding subunits of vaccinia virus RNA polymerase and transcription factors was reported (Ahn et al., 1990a, b; Broyles & Fesler, 1990; Broyles & Pennington, 1990). Many similarities in amino acid sequence have been observed between vaccinia virus-encoded proteins and cellular counterparts, and Goebel et al. described some new examples, with homologues of deoxyuridine triphosphatase (also noted by McGeoch, 1990), profilin (a cellular actin-binding protein), transcription factor 2-α, glutaredoxin, a steroid dehydrogenase and superoxide dismutase. Many genes of unassigned function, particularly near the termini of the genome, have been shown to be non-essential for growth in tissue culture (see in particular Kotwal & Moss (1988) and Perkus et al. (1991)). It is now clear that vaccinia virus encodes some examples of apparently redundant genes and also many genes whose activity is of no relevance for growth in tissue culture. The virus evidently employs a variety of evasive tactics to minimize its visibility to the host's defence mechanisms; a recent example was the demonstration by Kotwal et al. (1990) that vaccinia virus can inhibit the complement system.

Apart from the protein-encoding genes, the only characterized functional element in the vaccinia virus genome is a region adjacent to each terminal hairpin which is involved in maturing nascent DNA. Such DNA occurs as head-to-head concatemers in which adjacent copies of the major repeat (which will become parts of separate mature DNA molecules) are separated by the equivalent of a double-stranded version of the whole terminal hairpin sequence. In the WR strain, hairpins equivalent to each strand of the sequence are found; curiously, however, Goebel et al. (1990a) observed only one form in the plaque isolate of the Copenhagen strain used for sequence analysis of the genomic termini.

Merchlinsky & Moss (1986) showed that when a plasmid containing the concatemeric junction sequence was introduced into vaccinia virus-infected cells, the sequence was processed to yield a linear form of the plasmid with hairpin termini like the virus genome; they then used this plasmid assay to determine nucleotide sequences needed for the hairpin resolution reaction and demonstrated that a 20 bp sequence immediately adjacent to the hairpin was essential (Merchlinsky & Moss, 1989). In further work, Merchlinsky (1990b)
refined the analysis to show that two closely spaced A + T-rich sequences, which are highly conserved among poxviruses, form the active element. The simplest mechanism which might generate the hairpins would be observed features of the reaction, and proposed a model involving recombination at a specific site adjacent to the identified conserved region in the major repeats of two adjacent concatemeric units, followed by a directed strand migration.

Goebel et al. reported that for the most part the Copenhagen strain sequence was highly similar to published sequences of other strains. Major differences, however, were found near the termini: the WR strain terminal repeat is 2 kbp smaller than that of the Copenhagen strain and has a distinct set of reiterated families. The WR repeat also contains two genes which in the Copenhagen strain occur in unique sequence. Adjacent to the left terminal repeat, the Copenhagen strain contains over 4 kbp of sequence without any WR counterpart. These changes (and others described by Goebel et al.) clearly reflect the action of recombination events involving the repeat elements. In this context it must be remembered that the history of vaccinia virus is obscure, and none of today's strains can be taken as approximating a natural wild-type isolate.

The genome of HCMV

The complete sequence of the genomic DNA of HCMV strain AD169 was published by Chee et al. (1990a); this is the fourth herpesvirus sequence to be completely determined, after Epstein–Barr virus (EBV), varicella-zoster virus and herpes simplex virus type 1 (HSV-1). The HCMV sequence was determined by chain termination and M13 shotgun methods on plasmid-cloned fragments, and contains 229354 bp of 57.2% G+C composition.

The linear HCMV genome can be viewed as comprising two linked segments, termed L and S. The L segment consists of a unique sequence (U₁, 166-9 kbp) flanked by two copies of a repeat sequence in opposing orientation (TR₁ and IR₁, each 11 kbp). The S segment similarly consists of U₅ (35-4 kbp) with flanking repeats IR₅ and TR₅ (each 2.5 kbp). The two segments are joined at their IR₁ and IR₅ boundaries, and the genome exists as four isomers differing in the orientations of L and S. This general genome structure is seen also in certain other herpesviruses, notably HSV.

The HCMV sequence was interpreted in terms of its content of protein-encoding genes; the availability of extensive sequence data for other herpesviruses contributed significantly to this exercise. As with other herpesviruses, the unique parts of the HCMV genome appear to be largely filled with functional ORFs, whereas coding sequences in the R₁ elements are sparser. Chee et al. list a set of 208 ORFs. It is quite possible that not all of these represent functional protein-encoding sequences and, in addition, it is probable that a significant number of HCMV genes possess introns; these considerations set a limit on the precision with which the HCMV gene set can presently be catalogued. Comparisons of predicted amino acid sequences showed that about 30 of the proposed HCMV genes have clear homologues in the other herpesviruses; these correspondences allow a number of HCMV gene functions to be assigned. Approximately 10 more genes have probable homologues in other herpesviruses, based on lower level similarity or on location in the genome relative to clear homologues.

HCMV DNA is substantially larger than the other completely sequenced herpesvirus DNAs, and it now turns out that the HCMV genes with counterparts in other herpesviruses occupy the central part of the genome, roughly between nucleotides 50000 and 170000 in the prototypic orientation. These conserved genes can be assigned to six or seven large blocks, such that within each block the layout of genes is relatively well conserved among the four sequenced herpesviruses, whereas the relative positions of blocks are not conserved.

The genome of HCMV thus contains two large regions, adjacent to its extremities, which have no counterparts in the other herpesviruses. These regions are filled with genes unique to HCMV, among which the most notable features are the occurrence of multigene families and the large number of genes predicted to encode membrane glycoproteins. Multigene families were first seen in the S segment of HCMV by Weston & Barrell (1986), and nine families are now recognized. Certain of the families have widely dispersed members, occurring for example in both U₁ and U₅. In all, over 50 genes occur in families. The only other virus known even to approach this degree of redundancy is African swine fever virus, which has a poxvirus-like genome (see Almendral et al., 1990; Gonzalez et al., 1990). Many of the polypeptide sequences encoded in multigene families have characteristics of membrane glycoproteins: Chee et al. estimate there to be 54 ORFs which constitute exons or complete coding sequences for glycoproteins. Certain glycoproteins encoded in families have now been detected in virions (see Chee et al., 1990a); these are distinct from the better characterized virion glycoproteins of HCMV which correspond to HSV glycoproteins B and H, and which each have a single gene.
HCMV also possesses genes, not found in other sequenced herpesviruses, for homologues of several cellular membrane protein species. These include genes for a homologue of the major polypeptide of a major histocompatibility (MHC) class I molecule (Beck & Barrell, 1988), a possible homologue of a T cell receptor (mentioned by Chee et al., 1990a) and homologues of three G protein-coupled receptors (GCRs; Chee et al., 1990b). GCRs are multiply membrane-inserted proteins resident in the cytoplasmic membrane which act in the transduction of a wide range of extracellular signals via GTP-binding G proteins. The occurrence of HCMV counterparts to GCRs is certainly intriguing but is at present of unknown significance; Chee et al. (1990b) speculate that these HCMV proteins could be involved in the regulation of latency or reactivation, or in the induction of cellular proliferation.

Some properties of the HCMV MHC homologue were investigated by Browne et al. (1990). The occurrence of this homology had previously been connected with the fact that HCMV virions bind β2-microglobulin (β2m), whose normal role is as a component of the MHC class I complex (see Beck & Barrell, 1988). Browne et al. expressed the HCMV class I homologue and human β2m from recombinant vaccinia viruses, and observed that the two proteins formed a complex and that the transport of β2m to the cell surface was, in this situation, dependent on the HCMV protein. They also found that in HCMV-infected cells synthesis of mature, cell-surface MHC class I complex was inhibited, although mRNA levels were not depressed, and they proposed that the HCMV protein produces this effect by sequestration of nascent β2m. The infected cell would thus lose the ability to present viral antigen for cytotoxic T cell recognition and action.

A notable sequence-related advance for HCMV work in 1990 was the identification by Hamzeh et al. (1990) of a region in the HCMV genome active as an origin of DNA replication during lytic infection. The locus was identified by analysis of short DNA fragments made in the presence of a chain-terminating inhibitor of DNA synthesis. Initiation of DNA synthesis appeared to be bidirectional. The origin sequences, termed ori-Lyt, are located upstream of the gene (UL57) for the ssDNA-binding protein (a component of the DNA replicative machinery) in an interesting correspondence to the location of one of the HSV origins, oriL. However, the sequence of ori-Lyt does not resemble that of oriL, or indeed any other known herpesvirus origin. The limits of ori-Lyt are as yet poorly defined; it may be many hundreds of base pairs in length. The DNA sequence of the region is rich in repeats, palindromes and potential binding sites for transcription factors. Chee et al. (1990a) list three putative coding ORFs within the possible ori-Lyt limits (UL58, UL59 and UL60), but these are small and may not be functional.

Another aspect of HCMV genome functionality was described by Marschalek et al. (1989), who reported a locus in HCMV U6, which may be translated by RNA polymerase III to give RNA species analogous to polymerase III transcripts known for adenovirus, EBV and herpesvirus saimiri.

To end this section we note two papers on HCMV-related viruses. Lawrence et al. (1990) reported the first large scale sequence determination for the recently recognized human herpesvirus 6 (HHV-6). A region of 21 858 bp was determined containing 17 complete ORFs, most with counterparts in other herpesviruses. Analysis of relationships with other herpesviruses in terms of the degrees of similarity of corresponding ORFs and the layout of ORFs showed definitively that HHV-6 is much more closely related to the betaherpesvirus HCMV than to alpha- or gammaherpesviruses. This genomic relationship contrasts with the lymphotropic behaviour of HHV-6, which would group it with the Gammaherpesvirinae. Finally, another human herpesvirus was described in 1990, to be designated HHV-7 (Frenkel et al., 1990). This was isolated from T cells and appears to be most related to HHV-6.

The genome of rubella virus

Rubella virus is the sole member of the Rubivirus genus in the family Togaviridae, and contains a single-stranded, positive-sense RNA genome. Previous sequencing studies had provided information concerning the genes for the structural proteins (C, E2 and E1), contained within a single ORF in the 3' half of the genome (references given by Dominguez et al., 1990). The sequence has now been completed by Dominguez et al. (1990). The rubella virus genome is 9757 bases in length and has the highest G+C content, at 69.5%, of any RNA genome sequence determined to date. The RNA contains two large ORFs, the previously identified 3' ORF comprising 1063 codons and the 5' ORF, which comprises 2205 codons. The high G+C content is reflected in the codon usage in these ORFs, with 80-9% of the codons having third position G or C residues and a strong bias in favour of G+C-rich codons for isofunctional amino acids.

In gross terms, the rubella virus genome is similar in organization to those of other togaviruses, e.g. the alphavirus Sindbis, but is more compact and nearly 2000 nucleotides shorter. The 3' untranslated region and the two ORFs are smaller in rubella virus; these ORFs overlap by 149 bases and are read in different frames, whereas in Sindbis virus the ORFs do not overlap and are in the same translational frame. The structural
protein ORF in rubella virus lacks the 6K protein found between the E2 and E1 proteins of Sindbis virus. Cleavage of the rubella virus precursor polyprotein is mediated by cellular signal peptidases (Oker-Blom et al., 1990); in contrast, cleavage of the alphavirus C protein from the nascent precursor polyprotein is autocatalytic. Cleavage between the C and E2 proteins of rubella virus, which occurs after the E2 signal sequence, leaves the signal peptide as part of the C protein in virions; furthermore the C protein remains membrane-associated following translation in vitro of a synthetic mRNA encoding the C and E2 regions (Suomalainen et al., 1990).

The 5'-proximal ORF presumably encodes the rubella virus non-structural proteins, and a short region (122 amino acids) in this ORF displays significant homology with the N-terminal part of the Sindbis virus nsP3 protein. Two amino acid sequence motifs, which have been found in a diverse array of positive-strand RNA viruses associated with replicase and helicase activities, are predicted in the rubella virus 5' ORF. Intriguingly the helicase domain and nsP3 homology domain are in reverse order in the rubella virus genome compared to that in Sindbis virus. Dominguez et al. (1990) speculate that complex genome rearrangements or even interviral recombination events have occurred during the evolution of these viruses.

Research on human immunodeficiency virus (HIV): entry into cells and vaccine development

This section describes recent work on interactions between HIV and the cell surface, and outlines the present state of development of anti-HIV vaccines. Research on HIV and related viruses is extremely active and extensive, so that this section must be even more selective and lightly sketched than the other parts of our review.

We start with a description of advances in structural analysis of the CD4 protein, which has achieved notoriety as the primary cellular receptor for HIV (for a recent review on CD4 function, see Robey & Axel, 1990). The CD4 glycoprotein is found as a surface species in subpopulations of T lymphocytes, mainly helper T cells, and is recognized as a key component in the molecular machinery of the cell-mediated immune response. CD4 is part of the antigen recognition system of helper T cells; the CD4 molecule is thought to interact with the T cell receptor on the surface of the helper T cell (Chuck et al., 1990) and also with the MHC class II antigen-presenting complex on the target cell's surface (Doyle & Strominger, 1987). CD4 is also involved in transmitting the signal initiated by antigen–receptor interaction, via its association with a cytoplasmic protein kinase (Turner et al., 1990).

The CD4 polypeptide contains 433 amino acids, of which residues 1 to 371 are extracellular, and the molecule is anchored by a C-proximal transmembrane sequence. On the basis of amino acid sequence features the extracellular portion has been proposed to comprise four domains (termed V1 to V4) which each possess an immunoglobulin-like folding arrangement (Clark et al., 1987). For the N-terminal domain, V1, this assignment has been supported by analysis of binding patterns of monoclonal antibodies (MAbs) and the binding of HIV to a panel of CD4 point mutants (Peterson & Seed, 1988).

In 1990 two papers reported crystallization studies on the whole extracellular portion of CD4. Kwong et al. (1990) were able to crystallize this part of human CD4, but the crystals were not suitable for X-ray diffraction analysis. They indicated that the molecule had approximate dimensions of $25 \times 25 \times 125 \AA$, that is, it forms an elongated structure. Davis et al. (1990) described preliminary data on co-crystallization of the corresponding rat CD4 species with the Fab fragment of a MAb; the crystals obtained were judged suitable for diffraction studies, but results of this analysis are not yet available.

Two 1990 papers described high resolution structural analysis on a 182 amino acid fragment of human CD4 representing the N-terminal domains V1 and V2 (Wang et al., 1990; Ryu et al., 1990; reviewed by Travers, 1990). This part of the CD4 molecule includes the HIV binding site. The structures reported are to a resolution of 2.4 Å or better and the two results are very similar; the following description does not distinguish between the two papers.

The V1V2 fragment was found to be an all-β structure and does indeed form two domains with immunoglobulin folds. V1 is closely similar to immunoglobulin variable region structures, whereas V2 has a variant structure with a non-standard disulphide link and a distinct arrangement of β-strands. The two domains are held closely together, with a continuous β-strand running from V1 into V2 and with a large hydrophobic interface. Together they form a structure of $25 \times 35 \times 65 \AA$; that is, effectively the 'top half' of the proposed 125 Å long extracellular part of CD4.

The structure of the V1V2 fragment at present gives only limited insight into the functioning of CD4 in T cell immune interactions. The V1V2 structure is thought to bind MHC class II proteins, and the binding region is now seen to be a cavity formed by packing of the two domains; however, the structure of the MHC class II complex is not available to allow any direct evaluation of fit.

A number of studies have located the main determi-
nants for binding between CD4 and the HIV external glycoprotein gp120 moiety to a region within V1 (for instance, Peterson & Seed, 1988; Arthos et al., 1989; Ashkenazi et al., 1990; Brodsky et al., 1990). The residues of CD4 involved in gp120 binding correspond to one of the complementarity determining regions (CDR2 loop) in an immunoglobulin variable domain. The structures now available show that in CD4 this loop is longer and relatively more prominent than in immunoglobulins, and forms a distinct ridge on the molecule. Thus, the gp120 locality involved in interaction with this ridge may well prove to constitute a depression or groove on the surface of gp120, similar to poliovirus and influenza virus receptor-binding sites. Possibilities for designing inhibitors of binding based on this model will certainly be explored.

In a comparative study of the ability of chimpanzee, macaque and human CD4 proteins to promote membrane fusion in the presence of the HIV glycoprotein, Camerini & Seed (1990) found that human CD4 residue 87 (Glu) was required for fusion. This residue is now seen to lie in a surface loop corresponding to the immunoglobulin variable region CDR3, which forms a patch of negative charge on the generally positively charged surface of V1. These observations are taken to hint at the nature of processes in entry of HIV into the T cell following the initial binding event.

Although CD4 is clearly a major determinant of susceptibility to HIV, there are definite indications that it is not the only cell surface component involved. Chesebro et al. (1990) observed that in clones of HeLa cells into which a CD4 expression vector had been introduced, growth of HIV strains correlated with CD4 expression. However, when human squamous cell carcinoma or astroglial lines were treated in the same way, HIV infection failed even with high level CD4 expression; this block could be overcome by introducing the HIV genome via murine leukaemia virus pseudotypes. Similarly, Kim et al. (1990) found that HIV replication in a monocytic cell line naturally expressing CD4 was restricted at the entry stage.

The other player in the initial interaction between the HIV virion and the cell is the product of the env gene, the virion surface glycoprotein (which is cleaved into two separate chains, gp120 and gp41; the latter is the transmembrane component). Many lines of study are being pursued with these key components, only some of which are registered here. Much of the current work really awaits the determination of a three-dimensional structure for the env protein (not as yet achieved) to enable its full interpretation.

Isolates of HIV may vary widely in the properties of their virion glycoproteins. Daar et al. (1990) reported that, although laboratory strains of HIV bind to CD4 with high affinity, fresh clinical isolates may bind much less tightly, and Cheng-Mayer et al. (1989) proposed that neurological isolates of HIV may constitute a subgroup with distinct glycoprotein-associated properties. Such aspects will compound problems in arriving at an overall understanding of env protein functions.

Many papers have reported mapping exercises involving HIV-1 env. Olshevsky et al. (1990) evaluated CD4-binding of a panel of gp120 mutants (in the converse of the CD4 mutation analyses described above) and succeeded in identifying a small number of critical residues, located in several of the 'constant' regions of the protein. Coarse level mapping of properties that differ among HIV-1 strains has been obtained by construction of recombinants between distinctive infectious clones, and correlation of their biological properties and genomic structures (Cheng-Mayer et al., 1990; Liu et al., 1990; York-Higgins et al., 1990). Several variable properties, including tropism for macrophages (as opposed to T cells), cytopathogenicity and ability to alter CD4 expression were mapped to the env gene. A similar approach using recombinants between HIV-1 and simian immunodeficiency virus (SIV) may be workable, since some such chimeras can grow in human CD4+ cells (Shibata et al., 1990).

Human antisera to HIV can enhance infectivity of the virus and this could potentially be an important factor in the progression of HIV infection towards AIDS. One aspect of the phenomenon has now been described in more detail: Robinson et al. (1990a, b) showed that certain human MAbs which mediate complement-dependent enhancement of HIV-1 infectivity bind to a 35 amino acid immunodominant region in the N-terminal part of gp41. In addition, the synthetic peptide corresponding to the binding region inhibited enhancement of infectivity both by MAbs and by polyclonal human anti-HIV sera.

An important aspect of env protein analysis is the high sequence variability of HIV. The env gene is particularly affected, with five hypervariable regions (V1 to V5) within gp120 coding sequences. Analysis of variation has now become very precise, through use of polymerase chain reaction (PCR) techniques to study proviral DNA molecules or populations, from a single infected individual over a period and from persons infected from a single source. Simmonds et al. (1990), in a particularly thorough study, analysed env gene diversity in an HIV-1-infected haemophiliac cohort. They found extensive variability in encoded amino acid sequences: many hypervariable region changes were predicted to alter N-linked glycosylation of gp120. In addition to nucleotide substitutions, short deletions and duplications were common. The patterns of nucleotide substitution were consistent with the operation of a strong selection for
amino acid change in the hypervariable regions, presumably reflecting immune system pressure.

In other, similar studies, particular attention was paid to the gp120 V3 region, regarded as the principal neutralizing determinant of HIV-1. Wolfs et al. (1990) studied V3 variation in HIV-1 from six children infected from a single blood donor, and found that divergence of the V3 coding region was specific to each individual and rapid (with greater changes in hosts progressing to AIDS), and that the rate of change was independent of levels of virus antigen. LaRosa et al. (1990) related limits in the variability of V3 to probable structural requirements of this element, which is considered to form a disulphide-linked surface loop. The V3 regions of the commonly used HIV-1 reference strains, HTLV-IIIb and LAV-1, are now seen to be rare, and only a minority of donor sera react with them.

Development of vaccines against HIV represents a prime objective of research on the virus. Several 1990 papers reported production of non-infectious virus-like particles which might be useful in vaccines: Aldovini & Young (1990) described HIV cDNA clones lacking genomic packaging signals which generated virions devoid of genomic RNA on transfection; Hafler et al. (1990) expressed the gag and env genes in vaccinia virus vectors to generate HIV-like particles (some containing packaged env mRNA); and Smith et al. (1990) expressed gag-pol sequences linked to the rev-responsive element from a simian virus 40 (SV40) vector in the presence of rev protein to produce virus-like particles.

For the development of a candidate HIV vaccine, successful experiments in an animal model (e.g. the virus–host pair, SIV–macaques) seem to be a necessary prerequisite (Daniel & Desrosiers, 1989). After several years of unsuccessful trials, several groups have now shown that inactivated vaccines can protect macaques against challenge with live SIV (reviewed by Gardner & Stott, 1990). Murphy-Corb et al. (1989) showed that eight out of nine rhesus monkeys were protected by a gradient-purified, formaldehyde-treated whole SIV vaccine against intravenous (i.v.) challenge with 10 ID50 of the homologous virus; all of 17 control animals were infected by this dose. In contrast, a whole SIV vaccine (inactivated by psoralen treatment) with muramyldipeptide as adjuvant failed to protect rhesus monkeys against 102 to 103 ID50 i.v. or 1 to 10 ID50 given via the genital mucosa (Sutjipto et al., 1990).

Glutaraldehyde-fixed, SIV-infected cells with Quil A as adjuvant were successfully used as a vaccine to protect cynomolgus macaques against a challenge of 10 50% monkey infectious doses. The vaccine success was monitored by serological and PCR techniques, and by finding a strong cellular immune response (Stott et al., 1990). Post-exposure application of the successful vaccine did not clear the virus from SIV-infected monkeys nor significantly alter the course of infection (Stott et al., 1990). Recombinant gp120 seemed to protect chimpanzees against HIV-1 infection, but not gp160 (Berman et al., 1990), possibly due to enhancing antibodies binding to the transmembrane glycoprotein (as outlined above).

Thus, although various parameters of protection by vaccine remain to be worked out, work with the SIV-macaque model is making real progress.

**Baculoviruses: molecular biology and expression vectors**

Although insects are susceptible to viruses belonging to many families which also cause diseases in vertebrates, baculoviruses, which occur exclusively in arthropods, are the most extensively studied. *Autographa californica* multiply enveloped nuclear polyhedrosis virus (AcMNPV) continues to dominate molecular and *in vitro* studies on these viruses, although considerable progress has also been made in characterizing *Orgyia pseudotsugata* NPV (OpMNPV). Two comprehensive reviews were published in 1990: Blissard & Rohrmann (1990), on the molecular biology of these two NPVs, and Volkman & Keddie (1990), on AcMNPV pathogenesis both *in vivo* and *in vitro*.

Sequence information for AcMNPV published during the year brought the total sequenced up to about one-third of the 130 kbp genome. Guarino & Smith (1990) sequenced the 3-2 kbp *PstI* K fragment of AcMNPV which revealed seven ORFs, including the apparently misnamed 39K gene (predicted Mr, 31-3K from the sequence), whose promoter has been extensively used in functional mapping of viral factors which regulate delayed early gene expression (reviewed by Guarino, 1989), and V-ubi, a late gene, which has 76% similarity to eukaryotic ubiquitin (Guarino, 1990). The function of V-ubi is unknown, but the failure to isolate insertion mutants suggested that the gene was essential for virus replication.

The sequence of one of the most interesting baculovirus genes found to date was published by O'Reilly & Miller (1990). The egt gene, transcribed early in infection, encodes an ecysteoid UDP glucosyltransferase which catalyses the transfer of glucose to ecysteoids (O'Reilly & Miller, 1989). These steroid hormones are essential for the induction of moulting in insects, both between larval instars and between larva and pupa. The egt protein inactivates the hormone, and so prevents moulting and also the weight loss that occurs in larvae prior to pupation. The larger larvae are therefore able to produce a higher yield of progeny virus. This is the first example of a baculovirus gene that controls major
physiological processes within the host. It is also the first baculovirus gene product shown to be secreted and could be exploited for the design of baculovirus expression vectors that secrete foreign genes.

An interesting relationship between a baculovirus protein and an insect pox (entomopox) virus protein was reported by Vialard et al. (1990b). Spheroidin is the major occlusion body protein of entomopoxviruses and, like polyhedrin in NPVs, its gene has a very high level promoter. Yuen et al. (1990) sequenced the spheroidin gene and found no sequence similarity between spheroidin and polyhedrin. Subsequently, however, Vialard et al. (1990b) found that the spheroidin sequence was similar to that of a 37K glycoprotein encoded by another gene from AcMNPV which had been sequenced previously (Wu & Miller, 1989), and that the 37K glycoprotein cross-reacted with antibodies to spheroidin.

OpMNPV has a genome organization very similar to that of AcMNPV, although the homologous repeat regions (Cochran & Faulkner, 1983) characteristic of AcMNPV are not present in OpMNPV. The similarity between related gene products from the two viruses varies greatly, from less than 30% to around 80% at the amino acid level. Part of the HindIII G fragment was sequenced by Russell & Rohrmann (1990) and shown to contain four ORFs oriented in the same direction and expressed as a series of overlapping transcripts with a common 3' termination site. This arrangement is similar to that observed in a number of other regions of the OpMNPV and AcMNPV genomes. One of these ORFs encoded a polypeptide with 80% similarity to the basic DNA-binding protein of AcMNPV. The remainder of the HindIII G fragment, characterized by Müller et al. (1990), contained an ORF oriented in the opposite direction encoding a protein of predicted Mr 70600. Antibody raised against a bacterial fusion protein containing part of the coding sequence reacted with an 87K protein which was shown to be a component of budded virus, polyhedron-derived virus and purified empty capsids. This is the first OpMNPV protein for which a homologous protein could not be detected in AcMNPV either immunologically or by DNA hybridization at low stringency. If there is a homologous protein in AcMNPV, it must be poorly conserved and located in a different position in the genome.

Much of the interest in baculoviruses over the past few years has come from their use as expression vectors. The major protein of the occlusion body, polyhedrin, can account for as much as 20% of the total protein in infected cells, yet the polyhedrin gene is not essential for virus replication and can be replaced by foreign genes which are then also expressed at high levels (though rarely as high as polyhedrin itself) under the control of the polyhedrin promoter. A wide variety of transfer vectors have been developed for use with both the AcMNPV (Luckow & Summers, 1988) and the Bombyx mori NPV (Maeda, 1989) systems, and a very large number of gene products have been expressed, mostly from simple genes but also from intron-containing chromosomal genes (Iatrou et al., 1989). In general, the proteins are correctly processed and biologically active (Miller, 1988), although some highly glycosylated proteins have been shown to have less complex carbohydrate side-chains when expressed in baculoviruses (Kuroda et al., 1990; Vialard et al., 1990a) and other processing events may be significantly retarded (Kuroda et al., 1991). Jarvis et al. (1990) examined the role of glycosylation by using inhibitors of N-glycosylation and N-linked oligosaccharide processing, and concluded that N-glycosylation was sometimes required for the transport of recombinant glycoproteins through the secretory pathway of SF9 insect cells, whereas processing of the oligosaccharides was not.

A second gene, p10, encodes a 10K protein which, like polyhedrin, is not essential but expressed abundantly very late in the replication cycle. Foreign genes can also be expressed at high level under the control of the p10 promoter (Vlak et al., 1988; Weyer & Possee, 1989). Other baculovirus promoters have now been used for high level expression of foreign genes at earlier times post-infection (p.i.) than either the polyhedrin or p10 promoters (Hill-Perkins & Possee, 1990; Thiem & Miller, 1990). The AcMNPV 6.9K basic protein is a structural core protein closely associated with the virus DNA and vp39 is the major capsid protein. Copies of the promoters of these genes were used at the polyhedrin locus. β-Galactosidase was expressed under the control of the basic protein promoter at the same time as the viral basic protein, between 8 and 24 h p.i., with a peak synthesis at 12 to 15 h p.i. (Hill-Perkins & Possee, 1990). The vp39 promoter is more complex because there are three transcription start points. Promoters including one, two, or all three of these points expressed chloramphenicol acetyltransferase (CAT) by 12 h with increasing efficiency (Thiem & Miller, 1990). In addition, a hybrid promoter comprising the two distal transcription start points of vp39 placed upstream of the polyhedrin promoter exhibited both late and very late regulation. By comparison, expression of either β-galactosidase or CAT under the control of the polyhedrin promoter was barely detectable at 12 h p.i. It was suggested that earlier expression may be an advantage for those genes requiring extensive post-translational modification such as glycosylation, because the cells may be more able to carry out these processes earlier in infection.

One of the difficulties experienced by many workers in using baculovirus expression vectors has been the identification of recombinant virus against the back-
ground of parental viruses because the recombination frequency is usually 0.1 to 1%. The original selection method, loss of polyhedron formation, takes an experienced eye to achieve reliable results. Several papers published during 1990 have sought to address this problem, and as a result the selection of recombinants should now be very much simpler and quicker. The main approach used has been the design of vectors with two promoters, one directing expression of the gene of interest and the second directing expression of a selectable marker. Vialard et al. (1990a) constructed a polyhedrin replacement vector with a lacZ gene under the control of the p10 promoter adjacent to the polyhedrin promoter. Recombinants thus expressed both the foreign gene and β-galactosidase which allowed easy selection of plaques containing recombinant virus. A somewhat similar strategy was employed by Zuidema et al. (1990) to construct a polyhedrin-based transfer vector, pAcDZ1, with the lacZ gene under the control of the Drosophila melanogaster heat-shock promoter (hsp70), inserted downstream from the polyhedrin promoter. The use of the hsp70 promoter resulted in constitutive expression of the lacZ gene, but at a lower level than would be expected from either the p10 or polyhedrin promoters. This avoided possible complications from competition for transcription or translation factors due to high level expression of the reporter gene. The hsp70 promoter–lacZ gene–SV40 terminator cassette from pAcDZ1 was inserted into a p10-based transfer vector, pAcAS3 (Vlak et al., 1990), to provide similar easy selection of recombinants using this alternative locus for expression of foreign genes. This allows production of occluded recombinant virus which would be necessary where such recombinants had potential as genetically improved viral insecticides.

A rather different approach was used by Weyer et al. (1990), who examined the consequences of locating the polyhedrin coding sequence and promoter at heterologous positions in the AcMNPV genome. Polyhedrin was expressed at normal levels when the gene coding sequence was under the control of the p10 promoter. The p10 promoter was shown to function equally well at the polyhedrin locus as at its normal location and was used to direct expression of the lacZ gene when inserted upstream from the polyhedrin start codon. Either orientation of the gene gave similar levels of expression. Deletion of most of the lacZ gene from these constructs and replacement with a BgII cloning site resulted in two transfer vectors which could be used for expression of both polyhedrin and p10, and contained only the single additional foreign gene. Selection is on the basis of polyhedron-positive plaques using a polyhedrin-negative parental virus.

Kitts et al. (1990) showed that the proportion of recombinants among the progeny viruses could be dramatically increased by linearization of the viral DNA at a locus contained within the homologous region of the transfer vector. It was observed that AcMNPV DNA linearized at a Bsu36I site inserted at the polyhedrin locus was much less infectious than circular DNA and a further substantial reduction in infectivity took place when a fivefold excess of transfer vector DNA was cotransfected with the linearized DNA. As a consequence, although the total number of recombinants produced was reduced relative to using circular DNA, these recombinants accounted for between 14 and 37% of the total virus yield. A similar increase in the proportion of recombinants produced using a p10-based transfer vector was achieved using an analogous strategy.

Many viral antigens have been expressed in baculoviruses with a view to vaccine production. One of the more interesting recent examples (French et al., 1990) involved the simultaneous expression of four structural proteins from bluetongue virus (an orbivirus) which assembled to form double-shelled virus-like particles of the same size and appearance as authentic bluetongue virions. This was achieved by co-infection of cells with two recombinant viruses, each containing two of the bluetongue virus genes under the control of opposing polyhedrin promoters. Antibodies raised to these particles showed high neutralizing activity titres against the homologous serotype. Subsequently a fifth protein has been expressed, and shown to associate with subcore particles, in a similar system using three recombinant baculoviruses (Loudon & Roy, 1991).

Recombinant baculoviruses are also being developed as insecticides with enhanced pathogenicity relative to wild-type viruses. The δ-endotoxin genes from Bacillus thuringiensis were obvious candidates for such recombinants as the δ-endotoxin proteins provide the major insecticidal activity in preparations of the bacterium which are already widely used as biological insecticides. Two groups have now inserted δ-endotoxin genes from different serotypes of B. thuringiensis into AcMNPV-producing, non-occluded (Martens et al., 1990) and occluded recombinant viruses (Merryweather et al., 1990). Both papers demonstrated expression of toxin which was fully active when fed to insects. However, the occluded recombinant virus produced by Merryweather et al. (1990) had a twofold higher LD50 value than wild-type AcMNPV. Therefore insertion of the full-length δ-endotoxin gene did not apparently cause any increase in virulence of the virus.

Another protein with the potential to improve the effectiveness of baculovirus insecticides is juvenile hormone esterase (JHE), an enzyme which dramatically affects larval development by helping regulate levels of juvenile hormone. The gene encoding this enzyme has
been isolated from the tobacco budworm *Heliothis virescens* and inserted into AcMNPV (Hammock et al., 1990). High levels of expression were achieved in vitro, but expression in vivo was much lower and the enzyme was inactivated rapidly. The JHE-expressing recombinant profoundly reduced feeding and growth, relative to a polyhedrin-negative control, when non-occluded virus was fed to first instar larvae but did not have a significant effect on later instars.

**Systems to manipulate RNA virus genomes**

The genomes of positive-strand RNA viruses are infectious and thus, in principle, viable virus should be recoverable by transfection of RNA transcripts produced in vitro from cloned cDNA. For members of two genera of the Picornaviridae, the cardioviruses (e.g. *Mengo virus*) and the aphthoviruses (foot-and-mouth disease virus, FMDV), this has proved difficult because of the presence of a sequence containing predominantly C residues [referred to as the poly(C) tract] of variable length in the 5′ non-coding region of the genome which has hampered the construction of stable infectious cDNA clones. This problem was solved for Mengo virus (Duke & Palmenberg, 1989) by reducing the C tract to as few as eight residues to give cDNAs which were stable in bacteria and gave rise to viable virus in HeLa cells. For FMDV, cDNAs with short poly(C) tracts were non-infectious, and to overcome the stability problem in bacteria Zibert et al. (1990) first established full-length cDNAs containing 50 to 150 residues in the C tract in a 2 μm circle-derived vector in yeast cells. Although RNA produced from these clones was infectious, the low yield of 2 μm replicon rendered the system impractical for widescale use. Subsequently, the authors succeeded in obtaining a cDNA, containing a 32 C residue tract, in *Escherichia coli*, which produced infectious RNA.

Interestingly, after transfection of BHK cells viable FMDV RNA containing 60 to 80 C residues was recovered and the enlarged poly(C) tract remained stable after multiple passage of the recovered virus. Zibert et al. suggested a strand-switching mechanism of the RNA polymerase during replication elongates the poly(C) tract; the low frequency at which this occurs would account for the relatively low infectivity of recombinant FMDV RNA compared to that of other picornaviruses. The function of the poly(C) tract is unclear, but studies by Duke et al. (1990) showed that altering this region of the Mengo virus genome (using the cDNA described above) attenuated the pathogenicity of the virus in mice. Viruses were recovered from infectious cDNA clones containing different lengths of poly(C) tract, but in contrast to FMDV the virus RNA maintained precisely the same sequence as that transfected. Thus the poly(C) tract was not involved in infectivity of the RNA per se; however, whereas a recombinant virus containing the authentic Mengo virus sequence $C_{50}UC_{10}$ behaved exactly as a wild-type virus, and induced neurological illness and death in mice, those recombinants with shorter poly(C) tracts were severely attenuated. Administration of large amounts (50 μg) of live attenuated recombinant virus to mice produced high titres of neutralizing antibodies and conferred long-term protection against a lethal challenge with wild-type virus. Although Duke et al. suggest that analogous recombinant FMDV could form the basis for a new generation of attenuated vaccines, the elongation of the poly(C) tract described by Zibert et al. (1990) may present difficulties.

For negative-strand RNA viruses the ability to rescue infectious virus from cloned cDNA has only recently been realized. In our review last year (Elliott et al., 1990) we described work by P. Palese’s group which opened the way to make defined nucleotide alterations in the segmented genome of influenza virus. In 1990, this group achieved just that (Enami et al., 1990). RNA was synthesized in vitro from a cDNA clone of the neuraminidase (NA) gene of influenza virus A/WSN/33 (WSN), incubated with purified influenza virus polymerase complex to make ribonucleoprotein (RNP) (Parvin et al., 1989) and transfected into MDBK cells. The cells were superinfected with a helper virus, a reassortant virus composed of seven WSN genes plus the NA gene of the influenza virus A/HK/6/68 (HK). The WSN–HK reassortant cannot form plaques in MDBK cells unless protease is supplied in the medium to cleave the WSN haemagglutinin; in contrast WSN forms large plaques in MDBK cells in the absence of protease. Thus the use of this virus/cell combination provided strong selection for the isolation of a virus containing the WSN NA gene. Plaques were obtained and amplified, and virion RNA was extracted for analysis on polyacrylamide gels; the RNA profiles resembled those of WSN rather than WSN–HK, indicating the successful transfer of the synthetic NA gene. Further experiments used a WSN NA gene containing specific silent nucleotide substitutions, and viruses were obtained carrying these markers in their NA genes.

Measles virus has a non-segmented genome of about 16 kb which encodes six genes: nucleocapsid (N), phosphoprotein (P), matrix protein (M), fusion glycoprotein (F), haemagglutinin (H) and polymerase (L). Ballart et al. (1990) constructed full-length cDNAs of the measles virus genome under control of bacteriophage T7 or T3 promoters. These plasmid constructs could be transcribed in vitro to produce either genome or antigenome sense RNAs with a minimum of extra nucleotides at the 3′ and 5′ termini. In addition, specific silent
mutations in the envelope genes M, F and H. Supernatant fluid harvested from the cells 5 to 14 days after injection yielded a small number of plaques (1 to 18) when assayed on Vero cells. Lytic virus recovered from the supernatants could be passaged in Vero cells and generated titres approaching those reached by ‘authentic’ measles virus. Nucleotide sequence analysis of the recovered virus showed the presence of the specific substitutions made at the cDNA level, confirming the origin of the virus. This result represents a significant advance in the study of negative-strand viruses. The system relies on the non-producer persistent cell line, but could probably be adapted for use with cells which express individual functional transcription/replication proteins to propagate the injected synthetic genome.

Progress towards the propagation of negative-strand viral genomes in the absence of homologous helper virus has been described for two viruses, influenza virus and vesicular stomatitis virus (VSV). Both involve expression of replication proteins by vaccinia virus vectors. Huang et al. (1990) used a synthetic influenza virus-like RNA carrying a reporter gene (CAT) to monitor the protein requirements to replicate this molecule. The synthetic RNA was incorporated into RNP as described above (Parvin et al., 1989) and then transfected into cells which had been infected previously with recombinant vaccinia viruses expressing influenza virus proteins. The minimum set of proteins needed to replicate the synthetic RNA (as monitored by CAT activity) was found to be the three polymerase proteins and the nucleoprotein. Interestingly, neither of the non-structural proteins, NS1 or NS2, had a marked effect on replication in this system, although it had previously been speculated that they play a major role because both proteins are present in large amounts in the cell nucleus (reviewed by Krug et al., 1989). Huang et al. suggest that the non-structural proteins may be involved in the late stages of gene expression (i.e. their CAT-containing template behaves as an early influenza virus gene) or, although the authors feel it unlikely, a vaccinia virus protein may substitute for the non-structural proteins. In any event the system described provides a vehicle for detailed dissection of the functions of the different proteins involved in transcription and replication by mutagenesis of the vaccinia virus-expressed genes, and ultimately may lead to the production of infectious influenza virus, following transfection with mixtures of synthetic RNPs, without any requirement for a helper influenza virus.

A similar type of expression system has been described by Pattnaik & Wertz (1990) working with VSV, but in this case the replication proteins were expressed by the vaccinia virus–T7 polymerase system first described by Fuerst et al. (1986). Plasmids containing cDNAs for the L, N and NS genes of VSV, under control of the bacteriophage T7 promoter, were transfected into cells which had previously been infected with a recombinant vaccinia virus expressing T7 RNA polymerase. All three proteins could be synthesized if mixtures of plasmids were transfected, and the amount of each protein made could be controlled by altering the amount of DNA transfected. This system was used to study the protein requirements for the replication and amplification of the genome of a VSV defective interfering (DI) particle, which ordinarily is totally dependent on the expression of viral DNA and proteins by helper VSV. Replication of the DI RNA required expression of L, N and NS proteins by vaccinia virus–T7 polymerase; omission of any one of the proteins abolished replication. Furthermore, the degree of RNA replication depended on the relative ratios of the proteins; maximal levels were achieved at molar ratios of L : N : NS of 1 : 200 : 200, and were eight- to ten-fold higher than those achieved by wild-type VSV as a helper virus. Thus a functional analysis of these three proteins in VSV replication can be undertaken, and in the longer term this approach may provide the basis to replicate synthetic VSV genomes.

As the final topic in this section we deal with dsRNA viruses, which until 1990 had also proved refractory to genetic manipulation. Roner et al. (1990), in a paper somewhat misleadingly entitled ‘Reovirus RNA is infectious’, describe conditions under which reovirus RNA can establish a productive infection. The system is complex: the optimal transfection mixture comprised single-stranded (plus-sense) RNA and double-stranded genomic RNA of serotype 3 (ST3) as well as reticulocyte lysate in which all 10 reovirus ST3 mRNAs had been translated. Eight hours after the introduction of this mixture into L929 mouse fibroblasts by cationic lipid-mediated transfection, the cells were infected with reovirus ST2 and virus was harvested 23 h later. The use of ST3 RNAs and ST2 helper virus selects for ST3 progeny virus, in that ST2 forms plaques only after 10 to 12 days, whereas ST3 forms plaques in 5 days. The yield of ST3 virus from the transfected RNAs was 0·1 to 0·2 p.f.u./cell and contained virus with only a homogeneous ST3 genotype, i.e. no reassortants were found. However, reassortants were generated when a mixture of ST2 and ST3 dsRNAs, together with reticulocyte lysate primed with both ST2 and ST3 RNAs, were transfected. Seven of 50 plaque isolates examined were mono-reassortants,
i.e. they contained nine segments from ST3 and one from ST2. Although the roles of the various components in the transfection mixture are not fully understood or explained, this does not detract from the significant advance made in the study of reoviruses.

Non-conventional forms of resistance to plant viruses

It has been realized for several years that transformation of plants to produce viral coat protein can lead to resistance to certain viruses (for a review see Beachy et al., 1990). The experiments on coat protein-mediated resistance have been leading the development of the concept of non-conventional resistance to plant viruses, which can be conferred by virus or virus-related sequences that interfere with the virus infection cycle. Two papers published in 1990 highlight other approaches to generating resistant plants.

The geminivirus African cassava mosaic virus (ACMV) has a bipartite genome, comprising DNA A, which encodes information for virus replication and for the coat protein, and DNA B, which produces gene products essential for the movement of the virus through the plant. Stanley et al. (1990) transformed N. benthamiana plants with a tandem repeat of a subgenomic version of ACMV DNA B and showed that it ameliorated the symptoms of virus infection. The defective DNA was mobilized and amplified as a monomer in ACMV-infected plants and significantly reduced viral replication, especially after serial transfer. This result demonstrates that plants can be transformed with constructs which yield DI sequences.

Tobacco mosaic virus (TMV) encodes at its 5' end two co-initiated proteins, the 126K and 183K proteins, which are considered to be components of the viral RNA-dependent RNA polymerase. The 183K protein arises from readthrough of the stop codon of the 126K protein. It has been suggested that the readthrough portion, encoding a potential protein of Mr 54000, is also produced separately. Golemboski et al. (1990) transformed N. tabacum with a cDNA clone of the region of TMV encoding the 54K protein. Transformed plants contained the virus-specific transcript but no 54K protein was detected. These plants were resistant to infection by virus particles and by viral RNA of the strain of TMV that gave the transforming sequence. Plants expressing TMV coat protein are resistant to the virus but not to viral RNA infections. Coat protein-mediated resistance breaks down on inoculation with high concentrations of virus, at levels at which the 54K-transformed plants still showed resistance. However, the plants containing 54K DNA are resistant only to the original and very closely related strains of TMV whereas coat protein transgenic plants have some resistance to a much wider range of TMV strains. The mechanism of resistance in the 54K-transformed plants is unknown.

Other important targets for inducing non-conventional resistance are the enzymes involved in virus replication. More than 75% of plant viruses have plus-strand RNA genomes and require virus-encoded RNA polymerase for replication. Up to last year there was no reliable in vitro system in which plus-strand RNA virus products could be synthesized from plus-strand templates. The isolation and characterization of a soluble virus-specific RNA-dependent RNA polymerase from cucumber mosaic virus-infected N. tabacum leaves was described by Hayes & Buck (1990) and should provide such a system whereby the interactions involved in RNA virus replication can be studied. The polymerase preparations contained virus gene products 1a and 2a, which have motifs for helicase and polymerase respectively, together with a host polypeptide of 50K. The ratio of single-stranded to double-stranded RNA was about 5:1 which suggested that the minus-strand template was copied more than once.

Transgenic mice in virus research

The ability to introduce specific genes into the germlines of mice provides exciting methodology for the investigation of aspects of virus-host interaction, such as host range restriction, and opens up the possibility of rendering such animals resistant to viral infection by 'intracellular immunization' (Baltimore, 1988). We have selected two 1990 papers which illustrate the interesting uses of these manipulated animals.

Following the isolation of the human poliovirus receptor (PVR) gene (Mendelsohn et al., 1989), the same group has now generated transgenic mice carrying the PVR gene (Ren et al., 1990). Cosmid clones containing the PVR gene and its functional promoter were isolated from HeLa cell DNA, and used to produce transgenic mice. PVR RNA transcripts were expressed in most tissues, with the highest levels found in brain, spinal cord, lung and thymus. Homogenates of transgenic mouse tissues bound high levels of poliovirus, indicating expression of functional PVR. The transgenic mice were susceptible to infection with type 1 poliovirus (Mahoney strain) following intracerebral inoculation and developed paralytic poliomyelitis; extensive viral replication was detected in the brain and spinal cord. Non-transgenic mice did not show any sign of disease after inoculation with this virus strain. These results thus demonstrate that the ability of poliovirus to cause disease in mice is limited only by the lack of the cell surface
receptor. Inoculation of transgenic mice with the Sabin type 1 vaccine strain did not lead to clinical disease. Therefore, these mice are potentially valuable vehicles for testing attenuation and neurovirulence of new candidate poliovirus vaccine strains, which may reduce the dependence on the world's ever-decreasing supply of monkeys for testing purposes.

The dominant allele Mxl+ of the murine Mxl gene determines the resistance of mice to infection with influenza A or B viruses. The Mxl+ allele encodes a nuclear protein of M, 72000 which is inducible by α and β interferons; Mxl− mice do not resist influenza virus infection and, although the two identified Mxl− alleles are interferon-inducible, they contain mutations which leave only the coding capacity for the N-terminal half of the Mxl protein (reviewed by Staehli, 1990). Influenza virus-susceptible Mxl− cells can be rendered resistant in vitro by constitutive expression of the Mxl protein. The decisive role of the Mxl gene in conferring anti-influenza virus resistance in mice was demonstrated by introducing the gene into the germline of susceptible mice (Arnheiter et al., 1990).

Mxl DNA linked to its own promoter was injected into one of the two nuclei of one-cell Mxl− embryos, and pups containing the transgene were identified by Southern blotting. Six transgenic mouse lines were established which expressed different levels of Mxl protein after interferon induction; the time course of Mxl RNA induction in transgenic mice was similar to that in Mxl+ mice. Transgenic mice were challenged with a lethal dose of a neurotropic influenza virus; mice which were high responders to interferon treatment produced Mxl protein at the site of initial virus infection and, although the two identified Mxl− alleles are interferon-inducible, they contain mutations which leave only the coding capacity for the N-terminal half of the Mxl protein (reviewed by Staehli, 1990). Influenza virus-susceptible Mxl− cells can be rendered resistant in vitro by constitutive expression of the Mxl protein. The decisive role of the Mxl gene in conferring anti-influenza virus resistance in mice was demonstrated by introducing the gene into the germline of susceptible mice (Arnheiter et al., 1990).

Mxl DNA linked to its own promoter was injected into one of the two nuclei of one-cell Mxl− embryos, and pups containing the transgene were identified by Southern blotting. Six transgenic mouse lines were established which expressed different levels of Mxl protein after interferon induction; the time course of Mxl RNA induction in transgenic mice was similar to that in Mxl+ mice. Transgenic mice were challenged with a lethal dose of a neurotropic influenza virus; mice which were high responders to interferon treatment produced Mxl protein at the site of initial virus infection, little spread of virus ensued and the mice survived infection. In contrast, low responder transgenic mice resisted high doses of challenge virus, but succumbed to lower doses and died. However, these mice survived if interferon was co-injected with the virus. These results indicate that resistance to influenza virus is dependent on the efficient antiviral action of the Mxl protein and the extent to which it is induced by the infecting virus. Proteins related to Mxl which are active against other viruses have recently been identified (Arnheiter et al., 1990; Pavlovic et al., 1990) and thus it seems feasible to produce transgenic animals which are intrinsically protected against a range of virus infections. The precise mechanisms by which these proteins interfere with virus infection are not known.

We wish to thank M. S. Campo, A. J. Davison, R. Goldbach, M. C. Horzinek, V. Mautner, C. R. Pringle and N. D. Stow for their help in the form of information, discussion and criticism, and L. J. E. Kattenhorn for extensive help in preparing the text. R. M. Elliott is an MRC Senior Fellow.

References


We wish to thank M. S. Campo, A. J. Davison, R. Goldbach, M. C. Horzinek, V. Mautner, C. R. Pringle and N. D. Stow for their help in the form of information, discussion and criticism, and L. J. E. Kattenhorn for extensive help in preparing the text. R. M. Elliott is an MRC Senior Fellow.
AD169. Current Topics in Microbiology and Immunology 154, 125–169.


efficiency virus type 1 gp120 amino acids important for CD4 receptor binding. *Journal of Virology* 64, 5701–5707.


