REVIEW ARTICLE

Some Highlights of Animal Virus Research in 1987

By D. J. McGECH,1* A. J. DAVISON1 AND J. C. NEIL2

1MRC Virology Unit, Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR and 2Beatson Institute for Cancer Research, Wolfson Laboratory for Molecular Pathology, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, U.K.

We present here our review of highlights in animal virus research for 1987. This follows the format set by its predecessors (McGeoch et al., 1986, 1987a), particularly in that it is a personal and highly selective view of results published in 1987. Molecular virology has again been emphasized, reflecting the interests of the authors.

We start by describing current work on human immunodeficiency virus (HIV) and on papillomaviruses, two fields of particular urgency and application in modern virology. Topics relating to many viruses are then treated successively under the general categories of classification, growth, genome organization, genome replication, processes of transcription, protein modification and processing, and structural analyses. The account closes with an examination of the uses of oligopeptides as tools in recent research.

Research on HIV

Advances in the molecular biology of HIV were harder won in 1987 than in previous years, since analysis has now progressed from gene structure to gene function. Among the more significant achievements has been the analysis of the env gene products. Deletion and insertion mutations identify multiple levels at which the function of the env products can be compromised. Most notably, an essential domain for CD4 receptor binding is located towards the C terminus of gp120 (between residues 363 and 473) and consists of several conserved stretches flanked by variable regions, at least one of which can be mutated without affecting CD4 binding (Kowalski et al., 1987). Another study, involving mutagenesis and epitope mapping with monoclonal antibodies (MAbs) which block gp120-CD4 binding, also implicates a domain (residues 397 to 439) within this same region of gp120 (Lasky et al., 1987). A mutation affecting the N terminus of gp41 specifically interferes with fusion (Kowalski et al., 1987), supporting sequence-based predictions that gp41 carries this function (Gallaher, 1987). Despite the extent of env variability, the identification even of small portions of the molecule which are essential and invariant raises the hope that these can be made to function as targets for a protective immune response. The domains of CD4 involved in the gp120 interaction have not yet been mapped in the same detail, but a survey of primate lymphocyte CD4 epitopes showed that susceptibility to infection with HIV could be correlated with a conserved epitope cluster (McClure et al., 1987).

The HIV reverse transcriptase is at present the major target for antiviral chemotherapy, and it is likely that knowledge of the three-dimensional (3D) structure of the enzyme will assist in modelling new and more specific inhibitors. To this end, the gene product has been expressed in bacteria (Farmerie et al., 1987; Larder et al., 1987) and functional domains (e.g. triphosphate-and template-binding sites) have been tentatively assigned by site-directed mutagenesis (Larder et al., 1987).

The 'minor' gene products of HIV were the subject of continued controversy in 1987. The function most recently ascribed to the tat gene product is to overcome a transcriptional block which otherwise prevents the elongation of transcripts beyond the TAR (trans-acting response) sequence (Kao et al., 1987). The TAR sequence itself forms a stable stem–loop structure in the RNA form (Muesing et al., 1987) that is conceivably a protein-binding site modulated by the tat
product. The art/trs (anti-repressor of trans-activation/trans-regulator of splicing) product was also the subject of further controversy. New constructs which separate the tat and art/trs reading frames provide data which show that the function of art/trs is an absolute requirement for expression of the env product from existing mRNA. However, when co-expressed with tat, art/trs also augments the steady-state levels of env mRNA. It is not yet clear whether this effect operates at a transcriptional or post-transcriptional level (Knight et al., 1987). The continued confusion over the art/trs function highlights the problem of assaying individual regulatory elements in isolation.

The sor product, although not essential for viral replication in vitro, has been found to be required for the infectivity of extracellular virions. Spread of HIV sor mutants in cultured T cells may be limited to a direct cell–cell route (Fisher et al., 1987; Strebel et al., 1987). The mechanism of action of the sor product remains a mystery. However, for the other non-essential HIV product, 3'orf, an unexpected functional role has been outlined. It is reported that the 3'orf product bears a weak structural resemblance to ras gene products and G proteins, that it is myristylated in eukaryotic cells and that the bacterially expressed product binds and hydrolyses GTP (Guy et al., 1987). Furthermore, the same authors propose that the 3'orf protein cooperates with env protein in repressing expression of the CD4 molecule in T cells. From these and other data, it is tempting to speculate that the 3'orf product plays a role in down-regulating HIV expression and maintaining latency. A similar role may be inferred from the finding that deletion of 3'orf results in an approximately fivefold increase in viral DNA synthesis and replication (Luciw et al., 1987).

The control of HIV expression and induction from latency during T cell activation were elucidated further with the finding that the HIV long terminal repeat (LTR) enhancer region encompasses recognition sequences for a known transcription factor, NF-kB, and that mutation of the factor's recognition sequence abolishes activation of HIV LTR–chloramphenicol acetyltransferase constructs (Nabel & Baltimore, 1987). From previous studies it appears that induction of NF-kB-binding activity occurs by a post-translational mechanism following lymphocyte activation (Sen & Baltimore, 1986). A provocative finding was that herpesviruses such as herpes simplex virus type 1 (HSV-1) are capable of reactivating latent HIV infection in vitro (Mosca et al., 1987a). This may result from transcriptional trans-activation of the HIV LTR by the HSV-1 immediate early protein IE110, alone or in combination with IE175 (Mosca et al., 1987b). These findings provide a potential explanation of statistics which show an increased risk of HIV infection and disease progression in individuals with pre-existing herpesvirus infections (Quinn et al., 1987), although it must be recognized that HSV-1 is just one of many agents which may cause reactivation. As a further example, HIV-1 expression in chronically infected promyelocytic cells could be induced by cytokine mixtures or purified granulocyte–macrophage stimulating factor (Folks et al., 1987).

Confusion over the origin and identity of the human T cell lymphotropic virus designated HTLV-IV, ostensibly isolated from healthy prostitutes in West Africa, has been resolved (Kestler et al., 1988; Essex & Kanki, 1988) with the explanation that the 'new' virus resulted from laboratory contamination by the simian immune deficiency virus, SIVmac, a virus originally isolated from a captive rhesus macaque (Daniel et al., 1985). Furthermore, another 'isolate', STLV-IIIAGM, which was thought to have been derived from African green monkeys, may be explained similarly. This interpretation provides a more credible explanation for the observed identity or close similarity of the restriction maps (Kornfeld et al., 1987; Hirsch et al., 1987) and sequences of the various isolates (Chakrabarti et al., 1987; Franchini et al., 1987; Hahn et al., 1987). Nevertheless, West Africa is clearly a focus of infection with a second type of HIV, HIV-2 (originally named LAV-2), which has been cloned and sequenced (Guyader et al., 1987). Moreover, African green monkeys do harbour HIV-related viruses (Chakrabarti et al., 1987). HIV-2 and the monkey viruses appear particularly closely related in primary sequence; however, for those who wish to believe in a simian origin for the HIVs, there is still no recent antecedent for HIV-1. It may be a mistake to assume, however, that inter-species transmission has occurred in only one direction. There remains some doubt as to whether HIV-2 is as pathogenic as HIV-1. Although HIV-2 has been isolated from acquired immune deficiency
syndrome (AIDS) patients (Guyader et al., 1987). Essex and colleagues have presented serological evidence of infection with HIV-2-like viruses in apparently healthy West African populations (Kanki et al., 1986).

New and better animal models for HIV are being sought urgently, and two lentiviruses which emerged as candidates in 1987 were an old one with a new name, bovine immunodeficiency-like virus (Gonda et al., 1987; van der Maaten et al., 1972) and a new one, feline T-lymphotropic virus (Pedersen et al., 1987). Both viruses appear to be associated with immunosuppressive disease, but it remains to be seen how closely they resemble HIV in pathogenesis and cell tropism.

Apart from the well-publicized problem of virus variation, vaccine developers have been presented with further discouragement by the failure of early vaccine trials in chimpanzees. The animals had been immunized with recombinant vaccinia viruses carrying HIV env genes and had developed specific antibodies and T cell responses, but they nonetheless became viraemic after inoculation with live virus (Hu et al., 1987). Although the chimpanzee is not a perfect model for vaccine studies, the results cannot be encouraging. At present we do not know whether the human immune system is under any circumstances capable of resisting HIV infection, although it has been argued that the immune response influences the course of disease (Weber et al., 1987 a). Notwithstanding these cautions, some limited vaccine trials in human subjects went ahead in 1987. The first reported efforts were those of Zagury, who inoculated himself and a series of African volunteers with vaccinia virus recombinants and/or autologous cell extracts, infected with the vaccinia virus recombinant and containing HIV envelope glycoprotein (Zagury et al., 1987). In the U.S.A., trials of HIV env protein produced in recombinant baculovirus commenced, but these have been hampered at the outset by a slow build-up of volunteers (Ezell, 1987).

3'-Azido 3'-deoxythymidine (AZT) remains the only antiviral agent in widespread use with AIDS patients, and even this drug has severe limitations due to adverse reactions, principally through its toxicity for bone marrow haemopoietic precursor cells. Potential strategies for antiviral therapy were usefully reviewed by Mitsuya & Broder (1987). Promising targets for antiviral therapy are any of the HIV-specific functions that are not shared by the host cell. Thus, inhibitors of the protease that processes HIV gag and gag-pol precursors may be of benefit. On the basis of structural comparisons, the viral enzyme belongs to the aspartyl protease class (Pearl & Taylor, 1987) and it is encouraging to note that pepstatin A has some inhibitory effect on a range of retroviral proteases in vitro (Katoh et al., 1987). Further derivatives of this molecule, or substrate analogues based on the HIV gag precursor cleavage sites, may provide suitably potent and specific inhibition. Another new avenue of antiviral research was provided by the finding that compounds extracted from plants (e.g. castanospermine) which interfere with cellular glucosidases will inhibit HIV replication in vitro (Gruters et al., 1987). Again, the compounds involved would have to be refined and rendered more specific before they could be of clinical use. Since many drugs may be tested at first in combination with AZT, it is worth noting that some drugs show antagonistic rather than synergistic interactions with AZT in vitro (Vogt et al., 1987). A completely different approach to antiviral therapy is presented by soluble forms of the CD4 (sCD4) receptor which can inhibit virus replication in vitro (Smith et al., 1987; Deen et al., 1988; Fisher et al., 1988; Hussey et al., 1988; Trauwecker et al., 1988). Although many questions remain to be answered regarding the safety and efficacy of sCD4 in vivo, it is intriguing to consider that sCD4 may block not only virus binding but also autoantibodies against CD4, should these be present in infected individuals.

Research on papillomaviruses

The objective of this section is to give a view of current activities in research on papillomaviruses, with emphasis on viruses infecting humans. Papillomaviruses characteristically infect epithelial tissue, and for experimental purposes are distinguished by the fact that no fully permissive tissue culture system has yet been described for any member of the group. This has meant that until a decade ago characterization of these viruses was relatively limited, a situation which has now been changed completely by the application of modern methods. More
than 40 different human papillomavirus (HPV) types are currently recognized, with involvements in a wide variety of cutaneous and mucosal epithelial dysplasias and carcinomas. A number of animal papillomaviruses have also been studied experimentally, in particular bovine papillomavirus type 1 (BPV-1) and cottontail rabbit papillomavirus. Both human and animal viruses have strict host range and tissue specificities.

The present day approach to research on these viruses involves molecular cloning of genomes from samples of proliferating epithelium followed by characterization of genome organization by sequence analysis. Transcription patterns are being analysed using RNA species extracted from biopsy material or from partially permissive infected cells. Expression of virus-coded proteins by appropriate vectors in bacterial cells allows subsequent raising of specific antisera to enable the study of proteins in virus-infected tissue. In a new approach, Kreider et al. (1987) described a method for production of HPV-11 in the laboratory. This consisted of infecting epithelial fragments with material extracted from a genital wart, and grafting these beneath the renal capsule in athymic mice. After several months cysts developed and virus could be recovered.

New HPV types continue to be described. Lorincz et al. (1987) detected a new type by low stringency hybridization of HPV-16 probes to DNA from a cervical adenocarcinoma. The virus genome DNA was cloned, and designated as HPV-35. Beaudenon et al. (1987) cloned the genomes of two new types, designated HPV-39 and HPV-42, thought to belong to the low risk group and the potentially oncogenic group of genital HPVs, respectively. Naghashfar et al. (1987) described a species termed HPV-45, related to HPV-18.

Papillomavirus genomes are covalently closed dsDNA molecules of 7000 to 8000 bp. As indicated in Fig. 1, all the genes are thought to be similarly oriented. Additional genome sequence analyses were reported in 1987. Cole & Danos (1987) described the 7857 bp sequence of HPV-18 (which is associated with cervical carcinoma). Like previously sequenced DNAs, this has eight open reading frames (ORFs) that are thought to encode proteins. These authors
also carried out a phylogenetic analysis of papillomaviruses, using primarily the amino acid sequences of the large E1 ORF for comparisons. Three major families of viruses were distinguished by this approach: viruses infecting skin, genital viruses and animal viruses causing fibropapillomas. The complete DNA sequence of HPV-5 (7746 bp), associated with the potentially malignant disease epidermodysplasia verruciformis, was reported by Zachow et al. (1987). This condition is also associated with HPV types 8, 19 and 25, and Krubke et al. (1987) evaluated the relationships between members of this group by heteroduplex analysis and partial sequencing. Homologies over a large section of the genomes were estimated to be in the region of 70 to 80%. The analysis did not support recombination as a possible mechanism for the generation of HPV types. Finally, Patel et al. (1987) described the complete sequence of BPV-4 (7261 bp). This was seen as more closely related to the human and rabbit viruses that infect skin than to the well studied BPV-1, which causes fibropapillomas.

Characterization of mRNA species of several HPV types was reported in 1987. This has proceeded by the route of extracting RNA from a tissue sample, hybridizing with cloned fragments of HPV DNA and examining structures of the hybrids in the electron microscope. In the case of HPV-1 (associated with plantar warts), RNAs were also obtained from epithelial cells transiently infected with virus from warts (Chow et al., 1987c). In Fig. 1 the RNA species detected in extracts from warts are indicated; variations on this general pattern were found with infected epithelial cells (Chow et al., 1987c), and with HPV-6 and HPV-11 RNAs from genital warts (Chow et al., 1987a; Nasseri et al., 1987). Three transcription start sites and two polyadenylation sites were detected. Several different splicing modes then generated mRNAs with different 5′-proximal ORFs. Nasseri et al. (1987) carried out a detailed examination of the most abundant HPV-11 mRNA from a genital wart. This corresponds to mRNA (a) in Fig. 1 (also the most abundant HPV-1 mRNA), and joins the initiation codon plus the following four codons of ORF E1 to 85 codons of ORF E4.

Most work on the functions of the early genes of papillomaviruses has consisted of transfection experiments with BPV-1, and research in this field continued very actively in 1987. Many papers appeared describing the enhancer-binding and transcriptional activation/repression activities of BPV E2 protein (for instance, Androphy et al., 1987b; Lambert et al., 1987; Thierry & Yaniv, 1987). Mutational analyses of BPV-1 E4 and E5 genes were carried out by Neary et al. (1987) and Burkhardt et al. (1987), respectively. The function of E4 remains uncertain, and results were consistent with the protein having a role in virus maturation, as previously suggested by Doorbar et al. (1986). The E5 protein is involved in cell transformation. Burkhardt et al. (1987) showed that only the 44 residue polypeptide encoded by the 3′-terminal part of the ORF is necessary to transform cells. This protein is highly hydrophobic and is membrane-associated. The authors pointed out that this is the smallest transforming protein yet characterized and suggested that it may have a novel mechanism of action.

The first reports of the detection of HPV early proteins in papillomas, carcinomas or derived cell lines appeared only in 1986, when Doorbar et al. (1986) detected the E4 protein of HPV-1a in warts, and Smotkin & Wettstein (1986) detected the E7 protein of HPV-16 in a carcinoma cell line. More reports followed in 1987, all studies proceeding by first making antisera against viral proteins expressed in bacteria. The proteins detected, in a number of source tissues and representing several HPV types, included E1, E4, E6 and E7 (Androphy et al., 1987a; Banks et al., 1987b; Seedorf et al., 1987). Other papers reported similar results for the L1 and L2 capsid proteins (Banks et al., 1987a; Doorbar & Gallimore, 1987; Li et al., 1987).

Several 1987 papers dealt with approaches to establishing the mechanisms by which HPVs are involved in carcinogenesis. Pirisi et al. (1987) described the use of HPV-16 DNA in transforming human keratinocytes and fibroblasts to give an increased lifespan. Bedell et al. (1987) reported a system in which the E6 plus E7 ORFs of HPV-18 were sufficient to transform established rodent cell lines to anchorage-independent growth. Matlashewski et al. (1987) showed that the early region of HPV-16 DNA could transform primary rat cells to uncontrolled growth in conjunction with activated ras gene DNA, although neither was sufficient alone. Other papers (Popescu et al., 1987; Dürst et al., 1987) reported on integration sites of HPV sequences in cervical carcinomas; no simple message has yet emerged from these studies.
This section deals with some diverse topics in the classification and growth of viruses. First, 1987 saw the formal recognition of a new virus family, the Toroviridae (Horzinek et al., 1987; Weiss & Horzinek, 1986, 1987). This followed the earlier description of two viruses associated with enteric infections in horses and cattle (Berne and Breda viruses, respectively). Toroviruses are enveloped, with surface spikes. They contain an elongated nucleocapsid, tubular in cross-section. This may be straight, giving a rod-like particle, or it may more characteristically be bent into an open torus, giving a characteristic flattened version of the particle (hence the family name). The genome is a single-stranded, positive sense RNA molecule, the length of which has been estimated at about 20000 residues. However, this value was by comparison with coronavirus genome RNA, the size of which has now been revised upwards (see section on coronavirus genome structure, below). Subgenomic mRNA species have been detected in infected cells.

The Toroviridae are thus the fourth presently defined family of enveloped, positive sense RNA viruses (after the Togaviridae, Flaviviridae and Coronaviridae). The virion is said to be superficially similar to those of the Coronaviridae, but differs significantly in a number of details, in particular in the morphology of the nucleocapsid. In addition, unlike coronaviruses, the replication of torovirus in tissue culture requires the presence of a functional cell nucleus. Torovirus-related particles have been detected in faecal samples from human diarrhoea (Beards et al., 1984), and animals of many species show serological evidence of exposure to toroviruses. It is therefore expected that these will be found to be widespread agents of enteric infection.

Several papers appeared in 1987 dealing with the typing of picornavirus isolates. First, a brief report announced the extension of the numbering system for human rhinoviruses to 100 serotypes (Collaborative Report, 1987); over 90% of isolates can now be typed. This represents the culmination of the classical serological approach to typing and classification. Two papers described the next wave: large scale application of direct genomic RNA sequencing to typing, for foot-and-mouth disease virus (FMDV) and poliovirus.

Beck & Strohmaier (1987) analysed sequences of the VP1 genes (about 200 codons each) of 18 isolates of FMDV, types A and O, obtained from recent European outbreaks of the disease, plus nine older strains used for vaccines. Large numbers of single nucleotide differences were seen, many of which would be silent at the amino acid level, so that the data provided sensitive and precise information on relationships between strains. It emerged that most of the outbreaks were caused by FMDV strains closely and characteristically related to vaccine strains, while only two were thought to have been caused by exotic strains. In some cases there were close temporal links with vaccination or geographical links with vaccine production. It is the authors' unhappy but significant conclusion that most FMDV outbreaks in Europe in the last 20 years have been caused either by vaccination (with formalin-treated virus) or by escape of virus from vaccine production plants. Genome sequence analysis was advocated as a standard diagnostic method.

In a similar approach with poliovirus type 1, Rico-Hesse et al. (1987) analysed a section of 150 nucleotides within the genomes of 62 virus isolates from poliomyelitis cases world-wide. Again, many single nucleotide substitutions were found, and it was possible by systematic pairwise comparisons to discern clear relationships and groupings. Many distinct strains of the virus were distinguishable, each occupying a particular geographic locality. Transport of strains over large distances, and replacements of one strain with another over time, were also apparent. Evidence was obtained for the occurrence of recombinants in natural infection.

We close this section with a mixed set of results concerning viruses resistant to propagation in cultured cells. Milovanovic et al. (1987) reported the first successful serial propagation in culture of hepatitis B virus (HBV), in human peritoneal macrophages (obtained as an incidental product of clinical procedures). Although the virus persisted through 21 transfers, no c.p.e. was observed; this is in reasonable accord with the known pathology of HBV. The slow emergence of agents of enterically transmitted non-A, non-B hepatitis to a condition of virological tractability continued with the report of serial passage of the disease in monkeys (Bradley et al., 1987). Faecal specimens confirmed the presence of previously described virus-like particles. As mentioned above, a system was described for production of HPV-11 in the laboratory. Finally,
Yakobson et al. (1987) were able to propagate adeno-associated virus in synchronized cells in culture without the use of a helper virus, previously considered obligatory. The role of the helper, such as adenovirus or HSV, has never been precisely defined, and it now seems that it may be an indirect effect of the helper on the state of the cell that is important, rather than direct supply of functional proteins. In any case, justification for the present classification of the Paroviridae family into the autonomous Parovirus and the helper-requiring Dependovirus genera is certainly being eroded.

Coronavirus genome structure

In the general area of genome structure and organization, we describe work on coronavirus RNAs in this section and herpesvirus DNAs in the next section. One of the most striking genome structure studies published in 1987 was the report by Boursnell et al. (1987) of the completion of the genome sequence of the coronavirus, avian infectious bronchitis virus (IBV). Members of the family Coronaviridae have an enveloped particle and a large, single-stranded positive sense genomic RNA. The most commonly studied members are mouse hepatitis virus (MHV) and IBV. These two viruses are quite widely diverged, and this account is most concerned with IBV, but brings in MHV where appropriate. The true size of the genome RNA was not really appreciated until completion of the IBV sequence: the molecule contains 27600 nucleotides, by far the largest RNA genome known.

As shown in Fig. 2, coronavirus genes are expressed by a set of 3'-coterminal transcripts (reviewed by Siddell, 1987). Each of the subgenomic mRNAs has at its 5' terminus a common 'leader' of some 70 nucleotides, which is complementary to the extreme 3' terminus of full length genome complementary RNA. Synthesis of the mRNAs is thought to be accomplished by transcription of the leader followed by its dissociation from the template and reassociation at a specific sequence found in the genome at the start of each gene, where it can then act as a primer for transcription. Work in the last several years has concentrated on the smaller RNAs (A to E in Fig. 2) and it is clear that the three structural proteins of the virus (capsid, membrane and spike) are encoded by RNAs A, C and E respectively. Transcripts B and D encode non-structural proteins. Budzilowicz & Weiss (1987), working with the MHV counterpart of transcript D, showed that expression of this may be more complex, with the one RNA being translated into two polypeptides in distinct reading frames.
Boursnell et al. (1987) determined the sequence of the genome region of IBV unique to transcript F, which is thought to encode an RNA-dependent RNA polymerase. The first unexpected feature of this was its size, of over 20000 nucleotides instead of the approximately 10000 to 12000 expected. This means that essentially all detailed analysis of coronavirus gene function carried out till now has been concerned with less than 30% of the whole genome. Apart from its size, the sequence obtained was also surprising in that it contained two very large ORFs with only a small overlap, which could encode proteins of Mr 441000 and 300000, respectively. Because of a lack of sequence features associated with the 5' ends of the known mRNAs, the more distal of these seemed unlikely to be expressed via a subgenomic RNA. In a subsequent paper, Brierley et al. (1987) showed by in vitro translation experiments that the downstream ORF is expressed by a ribosomal frameshift event in the -1 direction, to generate a fusion protein. This process occurred in reticulocyte extracts with the notably high efficiency of 25 to 30%. This process was first described for retrovirus gene expression in higher eukaryotes. As an aside, a particularly notable retroviral example emerged in 1987, with demonstrations that expression of the gag, protease and pol genes of mouse mammary tumour virus involves two successive frameshifts on one messenger (Moore et al., 1987; Jacks et al., 1987; Hizi et al., 1987). Returning to IBV, this is the first authenticated instance outside retroviruses. The two enormous polypeptides which would result must at least in part provide RNA-copying activities. It is also quite possible that they could be extensively further processed.

Denison & Perlman (1986) showed that translation in vitro of MHV genome RNA, which should act as messenger for polymerase, gave rise to two protein species: a large protein of estimated Mr 220000, and a species of Mr 28000 termed p28. They then showed that the latter can be detected in infected cells (Denison & Perlman, 1987). Soe et al. (1987) obtained the sequence of 2000 residues at the 5' terminus of the MHV genome, which bears little resemblance to the corresponding IBV sequence. They identified the N-terminal part of the proposed polymerase-coding sequence and showed that p28 was encoded in this region. Clearly the structure of coronavirus polymerase and related proteins seems set to prove of considerable complexity.

Herpesvirus genomes

The genomes of herpesviruses are large and complex: that of human cytomegalovirus (HCMV), for instance, comprises some 230000 bp of linear dsDNA, contains two unique sequences each bounded by a pair of large, oppositely oriented repeats, and occurs in four sequence-orientation isomers which differ in the relative orientations of the two unique regions. Other genome arrangements are also found, in other herpesviruses, and the sizes and even base compositions of herpesvirus genomes vary widely. The characterized complexities of genome organization were further increased in 1987 with a paper by Takekoshi et al. (1987) describing a new isolate of HCMV, the Tanaka strain. The genome of this strain has an additional set of inverted repeat sequences, which gives rise to further sequence-orientation variations, generating in total eight isomers of the genome.

Davison & Taylor (1987) examined the detail of relationships between the gene layouts and encoded protein sequences for two herpesvirus genomes, those of varicella-zoster virus (VZV; 125000 bp) and Epstein–Barr virus (EBV; 172000 bp), both of which have been completely sequenced. VZV and EBV differ greatly in their biology and, as shown in Fig. 3, their genome organizations are also quite distinct. Little homology is detectable in their DNA sequences. However, comparisons of the amino acid sequences of encoded proteins, predicted from the DNA sequences, detected 29 pairs of homologous genes. A number of other pairs evidently also corresponded, from considerations of genomic location, size and orientation, but did not have any detectable similarity in their amino acid sequences. For each virus there remained a subset of genes unique to that virus. The conserved genes were located in three distinct blocks, but these blocks were arranged differently in the two genomes (Fig. 3). The two lineages have evidently thus evolved by processes of extensive mutation, rearrangement and large scale addition or deletion. These data give the most complete view yet obtained for relationships between two highly diverged herpesviruses.
The largest unit of herpesvirus DNA sequence reported in 1987 was the 20349 bp HindIII F fragment of HCMV (Kouzarides et al., 1987). This contained nine ORFs considered to represent coding regions of genes. These included amino acid sequence homologues of the DNA polymerase and glycoprotein B genes of HSV-1 and EBV, and eight in all of the HCMV ORFs were found to have EBV counterparts. These constitute the first definitive molecular examples of common genetic origins for cytomegaloviruses and other herpesviruses. The HCMV block of genes was split into two widely separate locations in EBV. In fact, the HCMV arrangement in this respect was similar to the VZV layout, and the HCMV HindIII F fragment can be seen to correspond to the adjacent portions of blocks B and C of VZV in Fig. 3. Thus, the rearrangement of blocks B and C in EBV is specific to the EBV lineage in these comparisons.

A number of papers in 1987 dealt with aspects of the 12 genes in the short unique region (Us) of HSV. McGeoch et al. (1987b) reported the sequence of most of HSV-2 Us, and showed that the genes therein were identical in layout and very similar in sequence to those previously characterized in HSV-1, with the major exception that gene US4 of HSV-2 was much larger than the equivalent HSV-1 gene. The US4 genes were shown to encode the virion surface glycoprotein species designated glycoprotein G (gG). The sequence data indicate that in HSV-2 the gG polypeptide has an Mr of 72000, while for HSV-1 gG the value is 25000; this is the most pronounced difference yet seen between corresponding genes of the two serotypes. Longnecker et al. (1987) showed that HSV-1 gene US7 encoded another virion surface glycoprotein (predicted from the sequence analysis but previously undetected), which they designated gI. HSV-1 is thus now known to possess seven surface glycoproteins [gH, gB and gC, with genes in the long unique region (UL), and gG, gD, gI, and gE, encoded in Us]. It has been known for some time that certain HSV genes are not essential for virus growth, at least in tissue culture. Remarkably, it has now emerged that 11 of the 12 Us genes also are not essential, at least in some cell types (Umene, 1986; Longnecker & Roizman, 1986, 1987; Brown & Harland, 1987; Neidhardt et al., 1987; Weber et al., 1987b). The only Us gene not in this class is US6, encoding gD. The other three glycoproteins encoded in Us [gG, gI and gE] are dispensable, and so is the protein kinase encoded by gene US3 (Frame et al., 1987; Purves et al., 1987). Furthermore, MacLean & Brown (1987) showed that two genes of unknown function at one extremity of the long unique region were dispensable: how many of the virus genes will eventually be assigned to this class?

Processes of virus DNA metabolism

This short section outlines new results in three aspects of biochemical processes of virus DNA synthesis and maturation. First, two papers reported that the large T antigen of simian virus 40 (SV40) can act to unwind SV40 duplex DNA specifically at the origin of DNA replication (Dean et al., 1987; Wold et al., 1987), in a reaction thought to represent an early step in replication of the virus DNA. The reaction requires ATP and a ssDNA-binding protein. T antigen had previously been shown to possess a helicase activity on model, ‘forked’ DNA substrates. The
new findings show that in addition it is a site-specific helicase, and capable of initiating action on an intact DNA duplex as well as at a fork structure. Deb & Tegtmeyer (1987) reported that ATP enhances binding of T antigen to the SV40 origin and extends the region of DNA protected by the protein from DNase; this presumably is a reflection of the site-specific unwinding reaction.

The gene encoding the DNA topoisomerase of vaccinia virus was mapped with seeming ease. This protein had been characterized previously, and was purified and studied in more detail by Shaffer & Traktman (1987). Physical analyses indicated that the active enzyme is a monomer with an $M_r$ of 32000. N-terminal amino acid sequence data obtained by Shuman & Moss (1987) indicated that the topoisomerase is encoded by a previously sequenced late gene in the HindIII H fragment, and a modest homology of the vaccinia virus protein with a yeast type I topoisomerase was detected.

During replication of retrovirus genomes, the genomic RNA is first, as is well known, copied in a series of events to yield a dsDNA molecule with flanking LTRs. This then becomes integrated into the host cell genome, where it can serve as template for transcription of RNA species, and where it stably remains. In a very elegant paper, Brown et al. (1987) described the development and application of a cell-free system to study integration of retroviral DNA into a target DNA.

In their scheme, the target was concatemerized phage $\lambda$ gtWES DNA, and the retrovirus was a competent murine leukaemia virus carrying the Escherichia coli $sup^F$ suppressor tRNA gene (MLV$sup^F$). Extracts of MLV$sup^F$-infected cells, containing proviral DNA, were incubated with the target DNA, and this was then deproteinized and packaged into $\lambda$ virions. Phage $\lambda$ gtWES carries three amber mutations and so does not plate on non-suppressor ($sup^-$) E. coli. However, phage carrying MLV$sup^F$ in a non-essential part of their genome expressed their own suppressor tRNA, enabling plaque formation on $sup^-$ bacteria. Thus, in this highly sensitive assay each plaque obtained represents an integration event.

It was found that the system functioned with very adequate efficiency, and a number of the resulting phage genomes were analysed. All contained the retroviral sequences, and the $\lambda$ component appeared normal. For seven of these phage, the junctions between $\lambda$ and MLV$sup^F$ were sequenced, and all were found to have structures expected for normal integration, namely a flanking 4 bp duplication of target DNA, and a deletion of 2 bp at each end of the insert DNA with respect to the proviral sequence. Thus, the target for normal integration does not need to be supercoiled or assembled into chromatin.

Previously, it had seemed likely that circularized proviral DNA with two LTRs, found only in nuclei of infected cells, was the integrating DNA species. Surprisingly, Brown et al. (1987) found that cytoplasm was the best source of integrating complexes, and that linear DNA must be active. It also appeared that the proteins necessary for integration were firmly associated with viral DNA, and a nucleoside triphosphate energy source was not necessary for integration to proceed to a covalently stable state. This paper should open a new chapter in the study of processes of retroviral infection.

Influenza virus transcription

This and the following section deal with two disparate aspects of transcription of viral genomes: first, with recent data on transcription of the segmented, negative-stranded RNAs of influenza virus, and secondly with aspects of transcription of the large DNA genomes of poxviruses.

A number of 1987 papers reported on cellular locations and complexes of influenza virus proteins, including those involved in transcription. Shapiro et al. (1987) extended previous work to show that all influenza virus RNA synthesis occurred in the nucleus. Smith et al. (1987) expressed each influenza virus protein individually from a vaccinia virus construct, and showed that the three polymerase proteins (PB1, PB2 and PA) migrate to the cell nucleus, as does the nucleocapsid protein (NP) and the non-structural species NS1. Akkina et al. (1987) and Detjen et al. (1987) demonstrated that the three P proteins form a complex both during transcription in association with nucleocapsid and when free of nucleocapsid.
Hsu et al. (1987) published an interesting account of the conformation of influenza A virion RNA species. The eight virion RNAs each contain limited regions of complementarity adjacent to their termini. This is, of course, common among virus genomic RNAs, and in some cases (with arenaviruses and bunyaviruses but not influenza virus) nucleocapsids or even free RNAs can be seen to form closed loops. Hsu et al. treated influenza virions with a psoralen cross-linking agent and found that the RNA molecules became specifically circularized by a panhandle structure at the termini. This did not happen with RNA which had previously been purified, so that the circularization must reflect the conformation of the RNA within virion nucleocapsids. They also showed that the structure could be detected inside virus-infected cells, where it was the predominant species during the period of synthesis of virus mRNAs. The characterization implies that regions of the template virion RNA involved in mRNA initiation and polyadenylation are brought into close proximity to the panhandle, and the intriguing possibility emerges that this may reflect an aspect of control between transcriptase copying (to give mRNAs) and replicative synthesis to give full-length positive strand species.

**Poxvirus transcription**

Progress, some of it particularly exciting, was made in the area of poxvirus transcriptional regulation in 1987. Yuen et al. (1987) demonstrated that vaccinia virus particles contain a factor with an Mr of 130000 which binds specifically to early promoters, and showed by DNase footprinting that the factor protects a 10 to 15 bp region centred 21 to 24 bp upstream from the initiation site. The close correspondence between the location of the binding site and sequences shown by deletion mutagenesis to be important for promoter activity (Cochran et al., 1985; Weir & Moss, 1987b; Coupar et al., 1987), and the inability of the factor to bind to mutated early promoters which fail to direct initiation, strongly suggest that the binding activity is due to a transcription factor. Thus, it is possible that initiation of early mRNA synthesis may be relatively simple, with binding of the factor to the promoter facilitating entry of the RNA polymerase complex.

Vaccinia virus early mRNAs terminate specifically in vivo and in vitro. Yuen & Moss (1987) identified the core consensus TTTTTNT (N is any nucleotide) as the termination signal in vitro. Termination occurs about 50 bp downstream from the signal, so that the signal is present (as UUUUNUNU) in the mRNA. Yuen & Moss also noted that the consensus occurs just upstream from the 3' ends of early mRNAs, but rarely in the coding sequences. Shuman et al. (1987) purified a transcription termination factor from vaccinia virions and found, surprisingly, that the factor copurified at all steps with the mRNA capping enzyme, which acts at the 5' end of the mRNA. Biochemical evidence indicated that it is the capping enzyme molecule, rather than the capping process, that is required for termination.

Transcriptional regulation of vaccinia virus late genes, which are expressed in vivo only after the onset of DNA synthesis, differs markedly from that of early genes. Late promoters differ in structure from early promoters and are not recognized by the virion-contained RNA polymerase complex responsible for transcription of early genes. The observation that for all late mRNAs so far analysed by S1 nuclease protection the 5' termini map within, or close to, a conserved TAAAT pentanucleotide suggests that this element is a crucial part of the promoter. A deletion analysis of a late promoter carried out by Weir & Moss (1987a) further indicated that at least 11 bp upstream of the TAAAT are also required for late promoter function, and that a more extensive upstream sequence is required for full activity. Since the TAAAT is followed by a G in most late promoters, so that the resulting AUG in the mRNA codes for the initiating methionine residue, it was thought until recently that late mRNAs contain few, if any, nucleotides upstream of the initiation codon.

This supposition was turned on its head by papers in 1987 which showed that late mRNAs have a poly(A) leader sequence upstream from the initiation codon, and therefore are produced by discontinuous synthesis (Bertholet et al., 1987; Schwer et al., 1987). Bertholet et al. used primer extension to demonstrate that the mRNA specified by one strongly expressed late gene contains heterogeneous lengths of sequence upstream from the 5' end defined by S1 nuclease
analysis. Analysis of cDNA clones showed that the upstream region comprises a poly(A) tract, preceded in some cases by RNA copied from sequences elsewhere in the vaccinia virus genome. These sequences seem not to be restricted by their location or orientation in the genome, or by their kinetic class. The ability of upstream sequences to protect a probe containing an appropriately located poly(T) tract from S1 nuclease digestion ruled out the possibility that they were an artefact of the use of reverse transcriptase. The heterogeneous, and often lengthy, nature of the upstream sequences was shown independently by electron microscopic examination of DNA–RNA hybrids. Bertholet et al. (1987) concluded, therefore, that vaccinia virus late mRNAs are produced by tagging protein-coding sequences on to the 3' ends of other RNAs. Schwer et al. (1987) also demonstrated the presence of upstream sequences by primer extension and S1 nuclease analysis using a poly(T)-containing probe, and showed that these RNAs are capped. They concluded, however, that the upstream sequence comprises a short tract of about 35 A residues. This was supported in a later publication by Patel & Pickup (1987), who concluded that mRNAs specified by a strong late promoter in the genome of cowpox virus commence with between five and 21 A residues. A significant step towards elucidating how late mRNAs are made was reported by Wright & Moss (1987), who described a system for synthesis in vitro of late mRNA containing a 5' poly(A) leader. The activity of a DNA template containing a minimal sequence upstream from the start of the RNA ruled out a cis-splicing mechanism for generation of the leader. Therefore, RNA polymerase 'stuttering', the use of a poly(A) primer or ligation of the poly(A) leader remain as possible models.

**Myristylation of proteins**

In this section and the next we treat findings on aspects of the modification of viral proteins, and on the assignment and sorting of proteins to different cellular compartments.

Addition of fatty acid residues to protein molecules is well known, including examples for viral proteins (reviewed by Sefton & Buss, 1987). However, the number of virus proteins known to be targets for addition of myristic acid has recently increased very noticeably, to the extent that this modification seems quite likely to become regarded as near ubiquitous. Myristic acid is n-tetradecanoic acid, straight chain 14-carbon carboxylic acid. Its occurrence in retroviral gag proteins and in the src gene product has been known for some time, and its detection in the 3'orf protein of HIV was mentioned above. It is found in proteins in an amide linkage to the N terminus, which differs from the occurrence of its 16-carbon homologue palmitic acid, which can be added to cysteine groups as a thioester near membrane-spanning regions, typically of glycoproteins.

In 1987, papers were published showing that myristic acid is found in components of picornaviruses (Chow et al., 1987; Paul et al., 1987), polyoma virus and SV40 (Streuli & Griffin, 1987), and HBV (Persing et al., 1987). In the case of picornaviruses it is the small internal protein VP4 that is myristylated (as are its precursors). A retrospective examination of the X-ray diffraction data for poliovirus showed that electron density corresponding to the myristyl group is indeed present at the N terminus of VP4, and that it interacts closely with the interior of the capsid shell (Chow et al., 1987). It is known that when poliovirus binds to a cell, VP4 is lost from the virion, and these authors speculate that the myristyl group may be involved in VP4 interacting with the cell surface. From the known sequences of myristylated species, a limited consensus for the addition of myristate can be discerned, the most notable feature of which is a terminal glycine residue. Towler et al. (1987) characterized the enzyme responsible for the addition (in yeast) as myristoyl CoA :protein N-myristoyltransferase. Wilcox et al. (1987) showed that in a muscle cell line the addition occurred cotranslationally. In some cases, but not all, the addition of myristate is associated with the acquisition of a membrane-binding capability; its function in the soluble class remains obscure. Besides addition of myristic and palmitic acids, another variety of covalent linkage of lipid to protein (with resultant membrane binding) has been found, namely attachment of a complex glycosylated phospholipid via the C terminus (Sefton & Buss, 1987; Cross, 1987). The 'prion' protein associated with scrapie is modified in this way (Stahl et al., 1987). So far, this has not been observed for a viral protein.
Targeting and sorting of proteins in the endoplasmic reticulum and Golgi apparatus

A great many virus-coded proteins, mostly destined for incorporation as virion surface components, are synthesized on membrane-bound ribosomes and concurrently transferred into the endoplasmic reticulum (ER). A variety of processing and sorting events can then occur to target a protein species to a particular cellular location. Much of the research in this area has in view the analysis of cellular processes, with virus proteins being used only as convenient models (see review by Rothman, 1987), but the findings are also potentially illuminating from a virological viewpoint.

Most membrane-translated polypeptides are not retained in the ER. Some, however, are resident there, including normal cellular processing enzymes. Munro & Pelham (1987) showed that three such cellular proteins (which are soluble species, not membrane-anchored) share the C-terminal sequence Lys-Asp-Glu-Leu, that removal of this abolishes retention in the ER, and that addition of this segment confers retention on a protein not normally resident in the ER.

Some virus proteins also reside in the ER. One of the most interesting of these is the E3/19K (or E19) adenovirus-encoded glycoprotein. Adenoviruses do not, of course, possess membranes, and this glycoprotein is not a virion structural component. Unlike the species mentioned above, E3/19K is membrane-anchored. It was previously known that in the ER it binds to nascent human histocompatibility (HLA) type I antigen (see section below on 3D structure), disrupting its glycosylation and transport. This binding involves the $\alpha_1$ and $\alpha_2$ domains, which are intimately involved with T cell recognition (Burgert & Kvist, 1987; see section on 3D structure below). Burgert et al. (1987) showed that this interaction functionally disables HLA competence of the infected cell. Paabo et al. (1987) showed that E3/19K is intrinsically resident in the ER, even in cells which do not express HLA. These workers also found that a mutant form of the protein with a truncated cytoplasmic tail became transported to the cell surface. Thus in this instance ER retention correlated with part of the protein chain on the cytoplasmic side of the ER membrane.

Another virus protein resident in the ER is the rotavirus glycoprotein VP7. The inner core of rotaviruses is assembled in cytoplasmic inclusions and then buds into the ER. After loss of the membrane acquired during this transfer, an outer shell of VP7 is added to complete the particle. These viruses are thus non-enveloped but do possess a glycosylated virion protein. Sequences available for VP7 genes of a number of rotaviruses show two possible initiation codons each followed by a hydrophobic region; each of these can act as a signal sequence (Whitfield et al., 1987). Poruchynsky et al. (1985) showed that deletion of the N-terminal region resulted in secretion of the protein and it appeared that this correlated with loss of the hydrophobic, transmembrane regions. However, Stirzaker et al. (1987) were able to demonstrate that cotranslational processing in any case removes both hydrophobic regions, so that in this instance the mature, ER-retained protein is not an integral membrane species. Nor does it possess the consensus sequence defined by Munro & Pelham (1987) for soluble cellular ER proteins.

Following the ER, the next cellular compartments involved in protein processing are the several cisternae of the Golgi apparatus. Proteins are routed to a variety of destinations from the Golgi, but some proteins are resident in this compartment. The coronavirus E1 (or M) glycoprotein is one such (see section on coronavirus genome organization above). Cloned versions of the E1 genes of both IBV and MHV can be expressed to give products which are correctly retained in the Golgi in the absence of other viral proteins (Machamer & Rose, 1987; Rottier & Rose, 1987). The E1 protein has a short, glycosylated N-terminal domain which is located inside the Golgi. This is followed by three closely spaced transmembrane sequences, and finally with a long cytoplasmic domain. Machamer & Rose (1987) investigated which parts of the molecule were required for Golgi residency. Earlier data had suggested that the elements responsible lay in the N-terminal half of the protein. Variants of the gene were constructed encoding only the first transmembrane segment (denoted $\Delta m2,3$) or only the third segment ($\Delta m1,2$) but leaving the rest of the molecule unperturbed. Both of the resulting mutant proteins were found to be inserted in the membrane. However the $\Delta m2,3$ protein was retained in the Golgi, whereas the $\Delta m1,2$ protein was found also at the cell surface. The retaining function thus appears to be within the first transmembrane segment, which is one of the regions most conserved between the IBV and MHV E1 proteins.
From the Golgi apparatus proteins are targeted to several destinations, including the cell surface, lysosomes, and other vesicles. The cell surface targeting is of particular relevance to processes of virus replication. However, it turns out that this target can be subdivided: it has been known for some time that cultured monolayers of kidney tubule epithelial cells can maintain a characteristic polarity between their two surfaces, so that different glycoproteins are routed either to the apical surface or to the basolateral surface. Influenza virus haemagglutinin (HA) becomes exposed on the apical surface, while the G glycoprotein of vesicular stomatitis virus (VSV) appears on the basolateral surface. These findings apply also to expression from isolated genes. Again, this has been pursued mainly as a topic in cellular processing studies.

Several papers in 1987 attempted to analyse the determinants of these specificities by constructing hybrids between the genes for HA and G, and expressing the resulting proteins. Puddington et al. (1987) substituted the cytoplasmic domain of G with the corresponding HA sequence, and found that the resulting protein had lost its targeting specificity in a canine kidney cell line. They concluded that either the hybrid had lost a specific signal for basolateral transport, or had gained a signal for apical transport. Roth et al. (1987), working with a monkey kidney cell line, found that polarized expression was maintained for a truncated HA lacking membrane anchor and cytoplasmic domain, as was an HA with these domains replaced by the corresponding parts of VSV G, thus suggesting that these parts of the HA molecule played no role in targeting. This result, however, was contradicted by a report that in a canine kidney cell line truncated forms of both HA and G lacking transmembrane and cytoplasmic sequences were secreted non-specifically (Gonzalez et al., 1987). The nature of the recognition processes involved is therefore not yet near complete resolution.

Three-dimensional structures

Arguably the most interesting 3D structural analysis of 1987 from a virological viewpoint was of the human class I histocompatibility antigen HLA-A2 (Bjorkman et al., 1987 a, b), rather than of any virus or viral component. The exterior portion of this important membrane protein was prepared by release from cultured lymphoblastoid cells through digestion with papain, and it was then crystallized and subjected to X-ray diffraction analysis. The molecule was previously considered to have three extracellular domains, α1, α2, and α3, and is also stably associated with β2-microglobulin (β2m). α3 and β2m are now seen from the X-ray analysis to possess immunoglobulin fold structures and to form a stalk (extending from the cell surface). α1 and α2 sit on top of the stalk as a platform which consists of a β-sheet base surmounted by two long α-helices. It is thought that the groove formed by the β-sheets and α-helices is the site of binding of fragments of foreign proteins, in the process of presentation for recognition by T lymphocytes. Most of the variable residues in the molecule are in this region, including those known to be critical for T cell interactions. Extrapolation of the broad features of the structure suggests a functionally similar architecture for class II histocompatibility antigens (Bjorkman et al., 1987 b). So, we now have detailed information on the structure of this important component of the cellular immune system, and some evocative hints on its mechanism. Just how parts of foreign proteins (for instance, viral proteins) come to be processed and presented remains a mystery.

Turning to another study on an immune system component, Colman et al. (1987) described an X-ray diffraction analysis of the structure of a complex between the Fab fragment of an immunoglobulin and the neuraminidase of influenza A virus. This is only the second such high resolution examination of an antigen–antibody complex, following the analysis by Amit et al. (1986) of a lysozyme–Fab complex. It appears that the neuraminidase–Fab complex results from perhaps 17 contacting residues. Of particular interest is the finding that in establishing this complex the conformation of the neuraminidase changes, as may that of the immunoglobulin. Thus, in the authors’ analogy, the interaction is more akin to a handshake than to a lock-and-key system.

New data continue to appear on X-ray crystallographic analyses of picornavirus particles. Fox et al. (1987) reported crystallization and preliminary diffraction data for FMDV. Luo et al. (1987) described the atomic structure at 3 Å resolution of Mengo virus (a murine cardiovirus).
This was achieved using the previously determined rhinovirus 14 structure as an initial model, without the use of isomorphous replacement forms. The overall organization of the major capsid proteins in Mengo virus was found to be very similar to that previously seen in rhinovirus and poliovirus. The surface characteristics, however, were greatly changed by insertions and deletions in superficial parts of the amino acid chains. The deep groove or 'canyon' on the surface of rhinovirus 14, proposed to contain the binding site for the cellular receptor, is in Mengo virus partially filled, to give a pit rather than a canyon.

Fuller (1987) reported on the structure of the virion of Sindbis virus (an alphavirus) to 35 Å, as deduced from computer processing of images of the particles obtained by cryo-electron microscopy. This extends the results previously obtained by Vogel et al. (1986) for the related Semliki Forest virus by similar methods. Alphaviruses possess a capsid which is closely covered by a lipid bilayer containing glycoprotein spikes. Fuller (1987) was able to describe the geometry of the outer spikes and also of the inner capsid. The lipid bilayer was visualized as polyhedral, with flat faces between the spikes. The geometry of interactions between the capsid (of 60 trimeric capsomeres) and the outer membranous shell (of 80 trimeric spikes, with each monomer consisting of one molecule each of the E1 and E2 proteins) was clarified.

This work described the Sindbis virus capsid as a smooth, fenestrated structure, and was confirmed in a subsequent direct analysis of isolated capsids using the same methods by Fuller & Argos (1987), who noted that the structure resembles that of the expanded forms to which some non-enveloped icosahedral viruses can be converted in the absence of divalent cations. With this background they then showed by sensitive, computerized comparisons of amino acid sequences that the alphavirus capsid protein sequences show distant relationships with picornavirus capsid proteins. This allowed prediction of a 3D structure for the alphavirus proteins, and led to the insight that these two virus groups may have a common evolutionary origin. As an aside, we note here that alphavirus research has now, like picornavirus work, gained the ability to 'resurrect' active virus from cDNA clones by making in vitro transcripts which are infectious, so allowing full application of recombinant DNA methods to study this virus group (Rice et al., 1987). In addition, Levis et al. (1987) showed that foreign genes can be inserted, via DNA copies, into defective interfering Sindbis virus RNAs and can be expressed there.

In another cryo-electron microscopy study, Talmon et al. (1987) analysed virions of La Crosse bunyavirus. This is also an enveloped virus, but in this case the envelope is not tightly associated with any inner capsid. The particles were nonetheless all seen to be near spherical, but with a range of diameters. None of the pleomorphic forms seen with conventional electron microscopic techniques were observed, and it now seems likely that such forms are artefactual.

Oligopeptides

It is only 8 years since it became generally realized, in a way startling at the time, that synthetic peptides could act as antigens to yield antisera active against the cognate sequence within a whole protein (Sutcliffe et al., 1980). Within that period the use of oligopeptides in virological research has increased enormously. A new phase of their application has arisen in the investigation of cell-mediated immune responses (for instance, Townsend et al., 1986; also, see section on 3D structure). Here we close this review by describing two papers which to us cogently illustrated the power that methods utilizing peptides as antigens have now attained.

First, Meloen et al. (1987) described a systematic analysis of the binding sites on FMDV VP1 of two MAbs, using synthetic oligopeptides. By means of small-scale, solid-phase methods they were able to synthesize and test a complete set of 207 overlapping hexapeptides against the 212 amino acid sequence. With the peptides still attached to a solid support, MAb-binding tests were carried out for each and quantified by ELISA. For one MAb, only two peptides, from separate parts of the sequence, were reactive (although one of these was to be later discounted). Two separate reactive groups, each containing several peptides, were found for the other MAb. Each binding site identified was then defined more fully using peptides extending in either direction from the reactive core. For each MAb this resulted in delineation of two sites, one of eight residues and one of six residues in each case. The contribution of each position within each of these sites was then evaluated by synthesis and testing of 'replacement sets'; that is, sets of
peptides with each residue in turn substituted successively by each of the 19 other amino acids. This enabled crucial binding residues to be identified.

The binding site for one MAb was convincingly located on two neighbouring localities near the C terminus of VP1, on the exterior surface of the virion as judged by comparison with rhinovirus crystallographic data. The site for the other MAb was also assigned to a surface structure; in this case, however, one of the reactive oligopeptides identified came from the interior of VP1, but was discounted as it resembled the surface site at several locations, and had closely similar patterns of activities in its replacement set. By our estimate, the work in this paper involved the synthesis and evaluation of more than 1000 oligopeptides. This work has been described here not primarily for intrinsic general interest in the results, but as a very impressive example of the degree to which the technology has developed, and of the resolution and precision attainable with synthetic peptide methods.

The second paper described in this section is by White & Wilson (1987), on analysis of the large irreversible change in conformation undergone by influenza virus HA on exposure to an acid environment. This extensively investigated process is part of the mechanism of entry of the virus into cells: influenza virus is endocytosed and, on subsequent physiological acidification of virus-containing endosomes, the HA conformation alters to enable fusion of the viral and endosome membranes and release of viral nucleocapsids into the cytoplasm (reviewed by Wharton, 1987).

White & Wilson (1987) raised antisera against 14 oligopeptides representing a variety of locations on the HA molecule. The binding of these antisera to a soluble form of trimeric HA (obtained by bromelain cleavage from virus) was then assayed, before and after exposure of the HA to pH 5 conditions. It was found that regions known from the 3D structure to be involved in contacts between monomers only became accessible to antibody after acidification. Different sites became exposed at different pH values and at different rates, so that steps in the process of conformational change could be distinguished. In the first step a loosening takes place in the shaft structure of the trimer; among other changes this exposes the so-called 'fusion peptide'. Subsequently, sites around the globular heads of the monomers became more accessible, suggesting separation of the heads. These results are consistent with other existing data, and extend knowledge of the process in the details of changes in the shaft of the trimer, and in the definition of distinct steps.

We wish to thank R. Elliott for critically reading the manuscript, and L. Kattenhorn for help in preparing this paper.

REFERENCES


Review: Animal virus research in 1987


