STRUCTURE OF THE MAJOR CAPSID PROTEIN (HEXON) OF ADENOVIRUS TYPES 40 AND 41, AND THE USE OF HEXON-DERIVED OLIGONUCLEOTIDE PROBES FOR DIAGNOSIS.

By Celia Toogood.

Department of Biochemistry and Microbiology, University of St. Andrews.

This thesis is submitted for the degree of Master of Science.

ABSTRACT.

The genes encoding the major capsid proteins (hexons) of human adenovirus types 40 and 41 were isolated and sequenced. Comparison of the predicted amino acid sequences with the hexons of adenovirus types 2 and 5 revealed regions of high homology, interspersed with regions of extreme variability across the four serotypes. Fitting of the Ad40 and 41 hexons to the known three-dimensional structure of the Ad2 protein, reveal that the majority of changes are confined to the loop domains which form the surface of the virion, while the P1 and P2 ß-barrels which comprise the base are well conserved. A major exception to this is the absence in both Ad 40 and 41 of 32 consecutive amino acids present in adenovirus type 2. In Ad2 this sequencee extends from the top of the I_1 loop down into the D-strand of the P1 domain, and includes a highly acidic domain which may be responsible for pH-induced conformation changes at the surface of the virion. Molecular-modelling suggests that, despite the absence of these amino acids, residues in both the Ad40 and 41 hexons can be accommodated into the P1 domain to form an alternative D-strand, hence maintaining the integrity of the base. However, it is predicted that the Ad 40 and 41 D-strands would be shorter than their type 2 counterpart. Thus it is likely that the hexons of the enteric adenoviruses closely resemble the Ad2 protein in the basal domains, but differ significantly in the architecture of their surface towers.

Unique sequences identified in the Ad40 and 41 hexon genes were used to design type-specific radioactive oligonucleotide probes which were

tested for their usefulness as tools for the diagnosis of enteric adenovirus infection. These oligonucleotides were successful in demonstrating the presence of enteric adenovirus DNA in faecal and infected cell extracts using a dot-blot hybridisation assay.

DECLARATION.

I, Celia Toogood, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

signed date

COPYRIGHT.

ACKNOWLEDGEMENTS.

I would like to thank my supervisor, Dr R.T. Hay, for his guidance throughout the course of this work, and for his advice during the production of my thesis. I also wish to thank our collaborators, Dr R. Burnett and Dr R. Murali at the Wistar Institute, Philadelphia, for the molecular-modelling data. I am indebted to the Bute Photographic Department for photography, and to Prof. U. Desselberger, Dr A.H. Kidd, and Barbara O'Donnell for supplying the clinical samples.

PUBLICATIONS RESULTING FROM THIS WORK.

Toogood, C. I. A. and Hay, R. T. (1988). DNA sequence of the adenovirus type 41 hexon gene and the predicted structure of the protein. Journal of General Virology **69**, 2291-2301.

Toogood, C. I. A.; Murali, R.; Burnett, R. M. and Hay, R. T. (1989). The adenovirus type 40 hexon: sequence, predicted structure and relationship to other hexons. Journal of General Virology **70**, 3203-3214.

This thesis is dedicated with love to my husband Dan, and to my parents.

CONTENTS.

experimental procedures cont.

INTRODUCTION

1. THE ENTERIC ADENOVIRUSES:

CHARACTERISTICS, PROPERTIES AND OCCURRENCE.

Adenoviruses are a group of non-enveloped, icosahedral viruses, all having a linear double-stranded DNA genome of approximately 35 kilobase pairs. They infect a broad range of mammalian species causing a variety of illnesses, primarily ocular, respiratory and gastric. The human adenoviruses which, to date, number forty-one distinct serotypes, share many common features including genome size, genome organisation, polypeptide composition and overall architecture. They have however, been subdivided into six groups (A-F), on the basis of shared immunological characteristics and on biological, chemical and structural properties. Adenovirus types 40 and 41 (Ad40 and 41) comprising subgroup F, were first identified in stool samples of children suffering from acute gastroenteritis (Flewett *et al*, 1975) and are commonly known as the enteric adenoviruses. This term is somewhat misleading since other serotypes have been isolated from the gut. For some years, confirmation of the role of the enteric adenoviruses in the aetiology of infantile diarrhoea was difficult to establish due to their fastidious growth characteristics. Both viruses proved non-cultivatable in cell lines conventionally used for the propagation of adenoviruses (KB, HeLa, human embryonic kidney and human amnion), thus the only method available for distinguishing Ad40 and 41 from other serotypes, was by electron microscopy followed by the inability to culture them. Cell lines have subsequently been identified which are susceptible to infection by Ad40 and 41 (Takiff *et al*, 1981; de Jong *et al*, 1983). These include Chang conjunctival cells, tertiary monkey kidney cells

and in particular, the Graham 293 cell line (Graham *et al*, 1977) which is a human embryonic cell line transformed by Ad5 and which expresses the Ad5 early E1 genes.

Techniques (discussed later) have now been developed for diagnosing infection by Ad40 and 41, and several studies have confirmed these enteric adenoviruses as major pathogens of childhood gastroenteritis (Whitelaw *et al*, 1977; Gary *et al*, 1979; Johannson *et al*, 1980; Hammond *et al*, 1987). Enteric adenovirus infection is characterised by diarrhoea and may be associated with fever, vomitting, and respiratory symptoms. Shedding of virus can be up to 10^{11} particles per gram of faeces at the acute stage of infection, and the level of virus excreted directly correlates with the progression of symptoms (Retter *et al*, 1979). There have been a number of studies on outbreaks of enteric adenovirus infection both in industrialised countries (Madeley el al, 1977; Brandt *et al*, 1985) and in developing countries (Dowling *et al*, 1981; Leite *et al*, 1985), and these suggest that Ad40 and 41 are second only to rotavirus as the aetiological agents of viral diarrhoea. In the West infection is not usually severe (Flewett *et al*, 1975), although occassional fatal cases have been reported (Retter *et al*, 1979). Outbreaks tend to be sporadic and fairly minor although hospitalisation is frequently required for rehydration. In Third-World countries however, where both sanitation facilities and public health care are poorly developed, dehydration due to gastroenteritis is a frequent cause of infant mortality. It seems likely therefore, that enteric adenoviruses may account for a significant portion of these cases. Electron microscopy of the gastrointestinal tract of a fatal case of infantile diarrhoea revealed crystalline arrays of

adenovirus particles within the nuclei of cells of the small intestinal mucosa (Whitelaw *et al*, 1977), and this, along with the high number of virus particles which can be isolated from the gut, strongly suggests that both Ad40 and 41 replicate within the gastrointestinal tract. The occurrence of respiratory infection in some children implies that the respiratory tract may also support replication, however no evidence for this has yet been found. A follow-up study on children hospitalised as a result of infection revealed that in some children, malabsorption of certain nutrients persisted for several months after the disappearance of symptoms (Uhnoo *et al*, 1984), implying that malnutrition following Ad40 or 41 infection may be an important secondary effect of these viruses, particularly where there is little or no after-care provision.

2. DIAGNOSIS.

In 1978, a programme aimed at the prevention and control of childhood diarrhoea was instigated by the World Health Organisation, and as a result the roles of various pathogens world-wide have been recognised. In elucidating the precise role of Ad40 and 41 in infantile diarrhoea, the development of a sensitive diagnostic test that is serotype- and preferably strain-specific is of prime importance. This would enable the epidemiology to be studied in detail, along with the spatial and temporal distribution of each serotype. A wide range of methods for identifying these viruses have now been described which vary in their usefulness as diagnostic tools. These methods are summarised below.

2.1 Electron microscopy (EM). Adenovirus particles are present in sufficient numbers in stools to allow easy detection by direct EM (Brandt *et al*,

1984). However, this approach is neither specific nor sensitive, relying on high particle concentration as an indication of enteric adenovirus infection. Specificity is improved dramatically by solid-phase immune-EM using antisera against Ad40 and 41 (Leite *et al*, 1985, Wood *et al*, 1989). Both these techniques rely on expensive equipment and skilled operators and are therefore not generally available to routine diagnostic laboratories.

2.2 Restriction enzyme cleavage profiles. Viral DNA isolated either directly from stool or from virus propagated in tissue culture, is cleaved with restriction enzymes and the resulting fragments resolved by electrophoresis. DNA is visualised either with ethidium bromide (Kidd, 1984; Kidd *et al*, 1984; Takiff *et al* 1984; Hammond *et al*, 1985), or by silver-staining (Allard *et al*, 1990). This technique has been used successfully to type clinical isolates and to detect strain variation, and could therefore be used to study distribution of the two serotypes and their variants. However, despite the apparent specificity of this method, virus levels are usually not high enough in stools to allow direct analysis, so culturing of samples would generally be recommended to ensure detection of all positives. This means that the processing of samples would be quite labour intensive, require tissue culture facilities, and that a number of positives may be missed due to a failure to grow. A good example of this is in the study by Allard *et al* (1990), where 60 stool samples grown in tissue culture were examined by restriction enzyme analysis revealing 20 positives. When the same samples were analysed by the more sensitive polymerase chain reaction, 51 proved positive.

2.3 Cell culture. This relies on differential cell culture growth characteristics,

for example virus which can be cultivated in 293 cells but fails to grow in HeLa cells is likely to be either Ad40 or 41. This method lacks specificity and sensitivity as virus multiplication can vary between different laboratories and different batches of cells, and infection may not always produce a visible cytopathic effect (de Jong *et al*, 1983; Brown *et al*, 1984). In addition it does not differentiate between the two enteric serotypes. The presence of more than one serotype in clinical isolates has been noted, indeed some adenoviruses can be shed in minute amounts for up to 900 days post-infection. Due to the poor growth of the enteric adenoviruses it is likely that other pathogens would outgrow them, thereby leading to a wrong diagnosis.

2.4 Immunological methods. Since the production of type- and strain-specific monoclonal antibodies, this avenue of approach has become the most successful, providing techniques which are simple and sensitive (Singh-Naz and Naz, 1986; Herrman *et al*, 1987). The variety of methods include multi-layer radioimmunoassay (Halonen *et al*, 1980; Cepko *et al*, 1983), indirect immunofluorescence microscopy (Retter *et al*, 1979; Cepko *et al*, 1983), neutralisation assay (de Jong *et al*, 1983; Singh-Naz and Naz, 1986), immunoelectro-osmophoresis (Jacobsson *et al*, 1979), enzyme-linked immunosorbent assay (ELISA) (Johansson *et al*, 1985; Uhnoo *et al*, 1984; Singh-Naz and Naz, 1986; Herrman *et al*, 1987) and latex agglutination (Grandlen *et al*, 1987). The ELISA method lends itself most readily to routine diagnostic work, being simple, sensitive and not requiring specialised equipment or training. However, it is dependent on the availability of specific monoclonal antibodies and furthermore, antigenic drift may result in certain

strains of Ad40 and 41 avoiding detection.

2.5 Polymerase Chain reaction. The recently developed polymerase chain reaction (PCR) (Saiki *et al*, 1988) has proved a powerful tool in numerous areas, not least in diagnosis. The method involves the *in vitro* enzymatic amplification of specific DNA sequences using selected primers, followed by visualisation of the amplified product on agarose gels stained with ethidium bromide. Only a few copies of target DNA are required to produce sufficient DNA for visualisation. This method has recently been used very successfully for the detection and typing of adenoviruses (Allard *et al*, 1990). In this study different primers were developed which were either group-specific, subgroup F-specific, or which could distinguish between types 40 and 41. PCR performed directly on stool samples gave a positive result for 13 out of 60 samples, and it was noted that the reaction could be inhibited by the use of too much sample. PCR performed following growth in tissue culture of the same samples gave 51 out of 60 positives, a finding which highlights the importance of tissue culture in this method. However this increases the chance of minor pathogens outgrowing the enteric serotypes, which though unlikely, may have the potential for being amplified hence masking the true cause of the symptoms. In addition, PCR is so sensitive that, not only will it amplify levels of virus which are too low to be of any clinical significance, it is extremely easy to contaminate the sample with foreign DNA which may lead to false positives. In its favour, PCR is cheap, rapid, extremely sensitive and suitable for routine diagnosis.

2.6 Dot-blot Hybridisation. This technique has been successfully used to

demonstrate the presence of both enteric adenoviruses in faecal samples without the need for tissue culture (Kidd *et al*, 1985; Hammond *et al*, 1987). Denatured DNA obtained from stools or from infected cell extracts is immobilised by spotting onto nitrocellulose or nylon filters and probed with radiolabelled fragments of homologous DNA. The probe hybridises to complementary DNA on the filters, the stability of the duplex being directly related to the degree of homology. Subsequent washing under conditions of high temperature and/or low ionic strength, eliminates unstable duplexes allowing only specific interactions to remain. A positive signal following autoradiography reveals the presence of the DNA of interest. The specificity of the probe is established by the washing step and can be adjusted to allow the test to be group-, subgroup- or type-specific. For example, a probe which is exactly homologous to Ad41 may possess 80% homology to Ad40, and 60% to other serotypes. Thus under certain washing conditions the probe can be used in a group-specific manner to differentiate adenovirus from, say, rotavirus, or under more stringent conditions it may be subgroup-specific or type-specific, should a more definitive diagnosis be required. Dot-blot hybridisation is a simple and specific method, however as the probes are relatively large, they cannot be used to detect the subtle changes in sequence found between strain variants. In contrast to large DNA fragments, synthetic oligonucleotide probes have the advantage of being cheap and easy to prepare in large quantities, and are sufficiently specific to detect single base-pair changes between target sequences. They can be radiolabelled to a high specific-activity using polynucleotide kinase and $\gamma-^{32}P$ ATP to attach a radioactive phosphate to the

5' end. However, the major disadvantage with oligonucleotide probes is that while the specific activity may be higher than that of nick-translated DNA probes on a 'per μg' basis, the actual amount of radioactivity that hybridises per copy of target DNA is much lower, resulting in an overall reduction in sensitivity. End-labelled probes are less prone to internal degradation than large nick-translated fragments and are thus marginally more stable. A major consideration when using both DNA and oligonucleotide systems is that the most commonly used radio-isotope, ^{32}P, has a half-life of only 14 days, so probes have to be made frequently to ensure maximum sensitivity, thus making the technique expensive. This problem has somewhat been alleviated by the use of ^{35}S and ^{125}I labelled probes (Pettersson and Hyppia, 1985; Matthews and Kricka, 1988), however this does not overcome the inherent instability and potential biohazard of radioactive probes.

More recently, a lot of effort has been put into the development of non-radioactive labelling strategies to circumvent the drawbacks associated with radioactive probes (reviewed by Matthews and Kricka, 1988). A vast array of methods are now available, the majority being based on the principle of the direct or indirect incorporation into the probe of a fluorescent tag, or an enzyme capable of producing an insoluble coloured product on the support matrix. Enzymes such as alkaline phosphatase, can be directly conjugated to a probe, either singly or as an oligomer. This approach has limited potential for signal amplification, and since the major drawback of non-radioactive systems is a lack of sensitivity, other methods involving multiple layers, have become more

common. The most widely used assays involve the incorporation of a hapten into the probe, which is detected by a reporter molecule. For double-stranded probes, methods developed for radioisotope incorporation can be employed for introducing the hapten, such as nick-translation and primer-extension. For oligonucleotides, haptens are generally attached to the 5' termini. The most commonly used hapten is biotin (vitamin H), which can be incorporated by a number of means. One approach involves nick-translation using nucleoside analogues, such as biotin-11-dUTP which consists of a biotin molecule linked to the nucleotide by an 11 carbon spacer arm. This system is reported to allow detection down to 1-5pg DNA. The spacing of the hapten from the DNA ensures that there is no steric interference to hybridisation. Another procedure utilises a photoactivatable analogue of biotin (Forster *et al*, 1985) which, in the presence of strong visible light, forms stable linkages with DNA, with approximately one biotin molecule being incorporated per 100-150 nucleotides. The chemical attachment of biotin to the 5' end of DNA via a multi-functional linker is particularly applicable to oligonucleotide probes (Agrawal *et al*, 1986; Tang and Agrawal, 1990). The advantage of biotin as a hapten is that it binds with very high affinity to the egg-white protein avidin and the bacteria-produced streptavidin. Both proteins possess four binding sites for biotin and can therefore be used as a bridge between a biotinylated probe and a biotinylated reporter molecule. Alkaline phosphatase and horseradish peroxidase are both available in a biotinylated form. This three-layered technique has the potential for signal amplification by the formation of large streptavidin-enzyme complexes, or by the synthesis of long polymers of

enzyme coupled to biotin. The assay can be simplified to a two-layer process by the reporter being directly linked to avidin. A point to be considered when using biotin in conjunction with stool sample analysis is that it is endogenous to many foodstuffs and may already be present in the sample. If present in sufficient quantity it may interfere with the assay, and this should be taken into account when interpreting results. Other haptens include dinitrophenol (DNP) which may be detected with an anti-DNP antibody, followed up with an anti-rabbit IgG - alkaline phosphate conjugate. The most common reporter enzymes to be used are alkaline phosphatase which utilises nitroblue tetrazolium and bromo-4-chloro-3-indolyl phoshate producing a blue precipitate, and horseradish peroxidase which forms a brown precipitate in the presence of hydrogen peroxide and diaminobenzidine.

A major disadvantage of all these methods is the problem of obtaining a lasting copy of the results. Filters are inherently difficult to photograph and signals tend to fade with time. This has recently been overcome with the introduction of enhanced-chemiluminescence (ECL). The method uses probes conjugated to horseradish peroxidase, and is based on the generation of light by the peroxidase-catalysed oxidation of luminol in the presence of an enhancer. The emitted light signal is captured by autoradiography, generally in less than 1hr, and a sensitivity of 1pg has been reported. ECL has many advantages over radioactivity particularly in the speed of obtaining results, and is certainly comparable with regard to sensitivity. ECL kits are available from commercial sources, however the technique is relatively expensive and

therefore not suitable for routine diagnosis.

Biotinylated probes produced from restriction fragments have successfully been used to detect and distinguish between Ad40 and 41 in clinical samples (Niel *et al*, 1986). It would seem likely that similar success could be obtained with short oligonucleotides, and it is the development of oligonucleotide probes as diagnostic tools for the detection of Ad40 and 41 which forms part of the study reported here. The selection of suitable oligonucleotides is of prime importance, namely, to identify sequences which will be unique to each virus. The Ad5 hexon gene (Kinloch *et al*, 1984) contains areas which are very highly conserved with respect to the Ad2 hexon gene, and these are interspersed with regions of extreme diversity. In particular, a fragment from the 5' end of the Ad2 gene has been shown to be highly homologous not only to Ad5, but to bovine adenovirus type 3 (Hu *et al*, 1984), suggesting that this region may be well conserved across many mammalian serotypes. Using this fragment as a probe, it was possible to identify and isolate homologous hexon gene fragments from Ad40 and 41. By determining the nucleotide sequences of these fragments, the information to design specific probes was obtained. Following radioactive labelling, these probes were tested for their ability to detect Ad40 and 41 in stool samples and infected cell extracts using a dot-blot hybridisation assay. The potential for further development of the oligonucleotides as non-radioactive probes is discussed.

3. ADENOVIRUS ARCHITECTURE.

The organisation of the structural proteins within the adenovirus has been the subject of investigation for more than twenty years. Initial studies using electron microscopy showed the virion to consist of two major structural complexes: the capsid (an icosahedral protein shell), and the core (comprised of the viral genome in close association with at least two proteins). Fibres of variable length depending on the serotype, protrude from each of the 12 vertices. All adenovirus serotypes possess virions of approximately equal size which, for Ad2, is 900Å in diameter calculated from a sphere touching all facets of the icosahedron (Burnett, 1985). To date, at least ten structural proteins have been identified within mature virions, and for the majority, a location and function have, to a greater or lesser extent, been defined. The major components are virus polypeptides (VP) II, III, IIIa, IV, V, VI, VII, VIII and IX, named by virtue of their relative order of mobility on SDS polyacrylamide gels (figure i). A small number of virion components remain poorly characterised, and these include VP X, XI and XII which are thought to be small molecular weight cleavage products of precursor polypeptides.

3.1. The virus capsid.

Four proteins have been identified as the major components of the adenovirus capsid: hexon (VP II), penton base (VP III), fibre (VP IV) and a small protein, VP IX. Hexon is the most abundant, accounting for 62% of the virion protein. The capsid is comprised of 252 morphological units or capsomers (Horne *et al*, 1959) of which 240 are hexons and the remaining 12 are pentons. Hexons

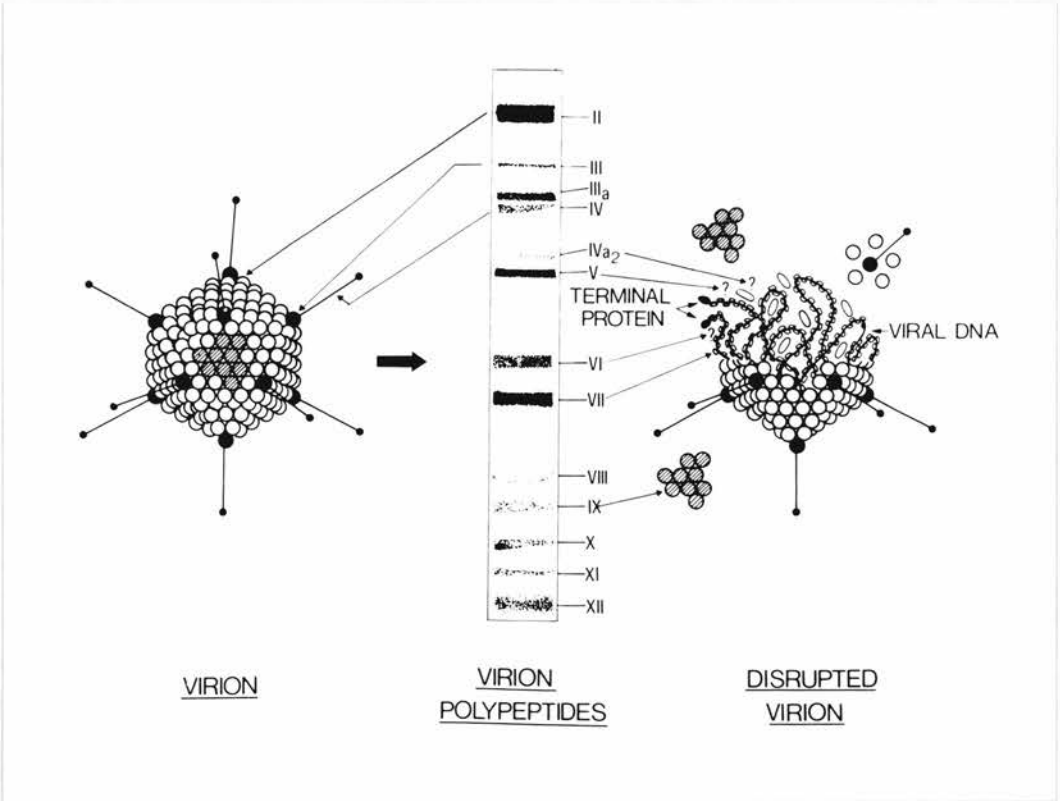


Figure i. Structure and polypeptides of the adenovirion. Left-hand side: the icosahedral form of the capsid showing the location of the hexon (II), penton (III) and fibre (IV). Centre: polyacrylamide gel electrophoresis showing the relative mobilities of the different polypeptides. Right-hand side: Upon dissociation of the virus groups-of-nine hexons, the peripentonal hexons and the penton are released revealing the double-stranded genome with attached terminal protein, and the core proteins (V and VII). The exact locations of VP VI, VIII, X, XI and XII within the virus have not been determined. **(From Ginsberg, H.S. 1984 p.207).**

make up the 20 facets of the capsid while pentons, complexes of penton base and fibre, form the vertices and give the adenovirus its characteristic shape (figure i,). The amino acid sequences of the Ad2 (Akusjarvi *et al*, 1984) and the Ad5 (Kinloch *et al*, 1984) hexons have been determined from the nucleotide sequences of the respective genes, which predict the proteins to have molecular weights of 109kD and 103kD respectively. Hexon is a trimer of three identical polypeptide chains of 967 amino acids in Ad2, and 951 in Ad5, which combine in a pseudo-hexagonal conformation (Grütter and Franklin, 1974). The three-dimensional structure of the Ad2 hexon has been determined by X-ray crystallography to a resolution of 2.9Å (Roberts *et al*, 1986) revealing a protein of great complexity (figures ii, iii). Each hexon monomer is comprised of two topologically identical eight-stranded ß-barrel pedestal domains (P1 and P2), the three copies being arranged in an alternating fashion around a pseudo-sixfold axis. The ß-barrels form each of the six corners of the hexagon. (figure iiia). Superimposed upon the base is a triangular tower made up of three T domains. The T domains are formed from loops that arise from the six ß-barrels, I_1 and I_2 from each of the three P1 domains and I_4 from the P2 domains. Thus there are three copies of each loop per hexon trimer. The arrangement of these loops is particularly complex, each T domain being composed of one copy of I_1, I_2 and I_4, and each of these arising from different monomeric units of the trimer (figure iiib). The resulting high degree of interweaving confers great stability on the hexon. The arrangement of the basal domains with respect to the towers results in two distinct vertical faces A and B, A lying directly below the apices of the triangle (figure iv).

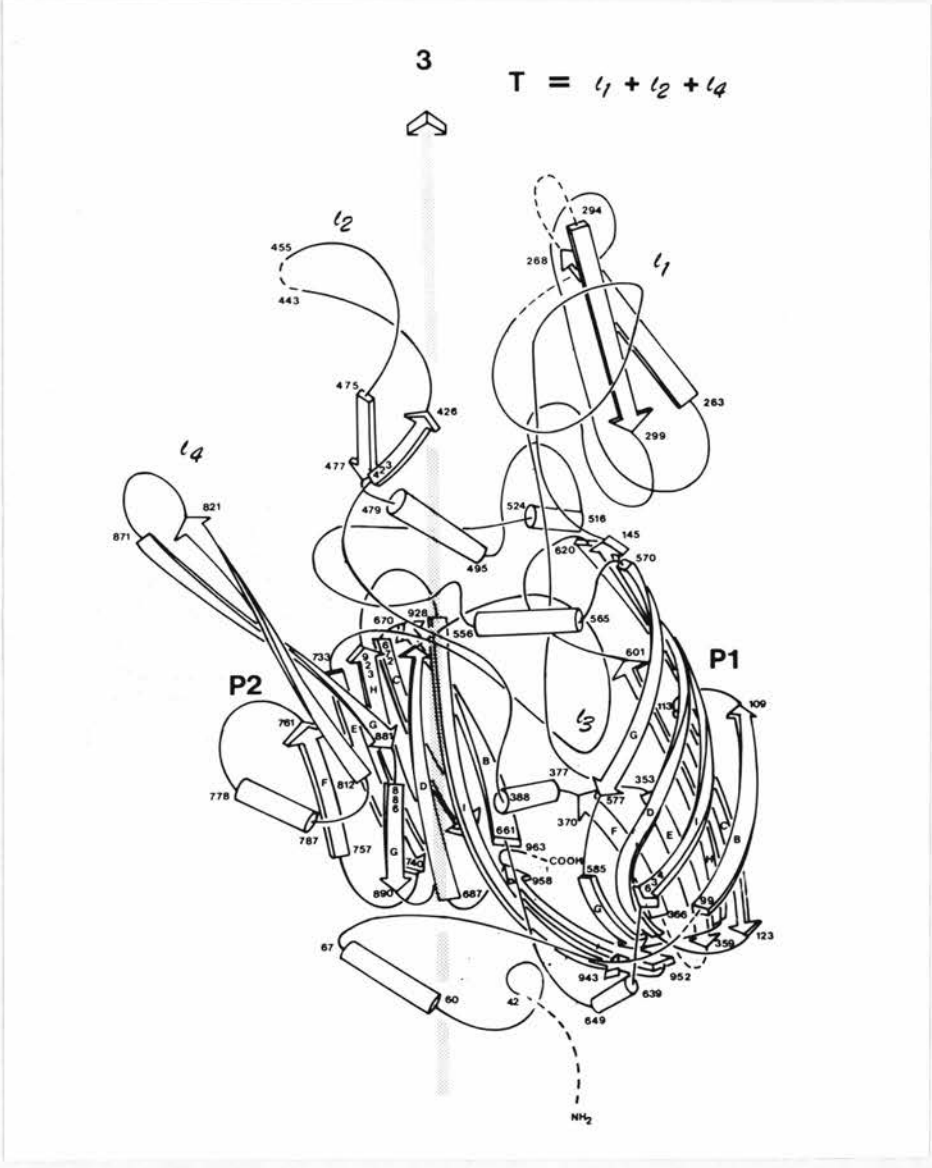
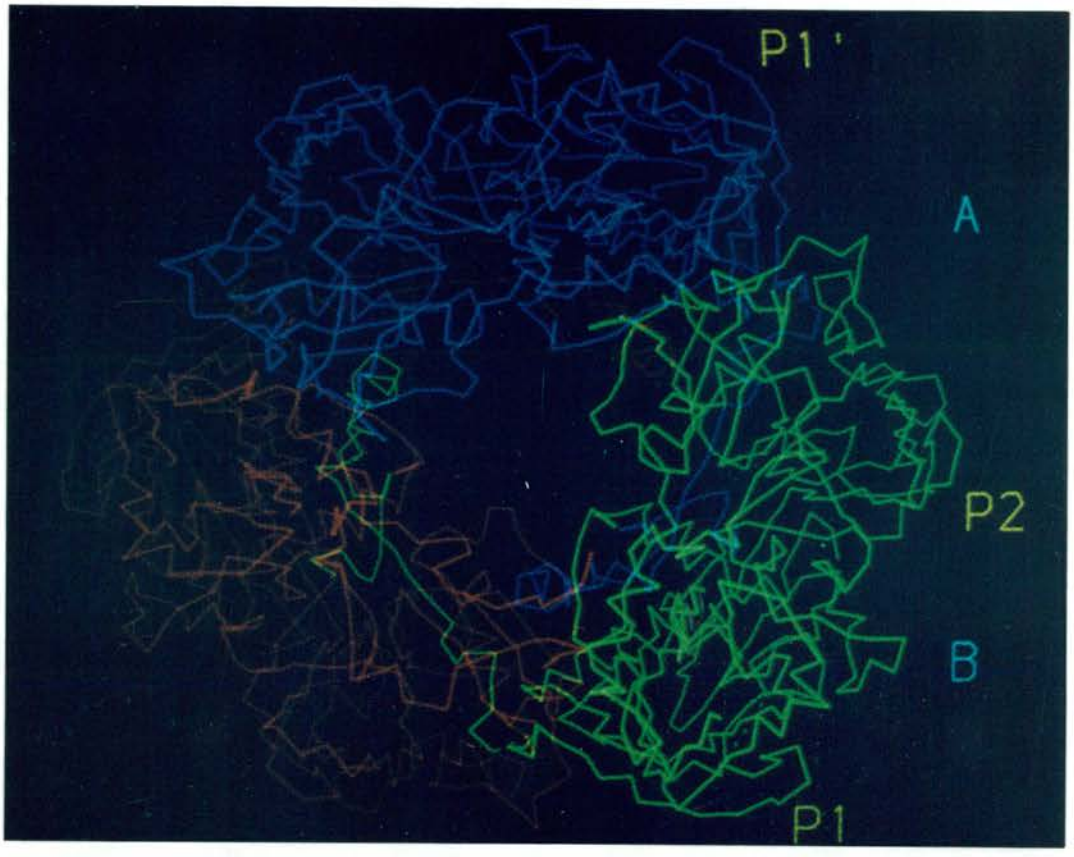


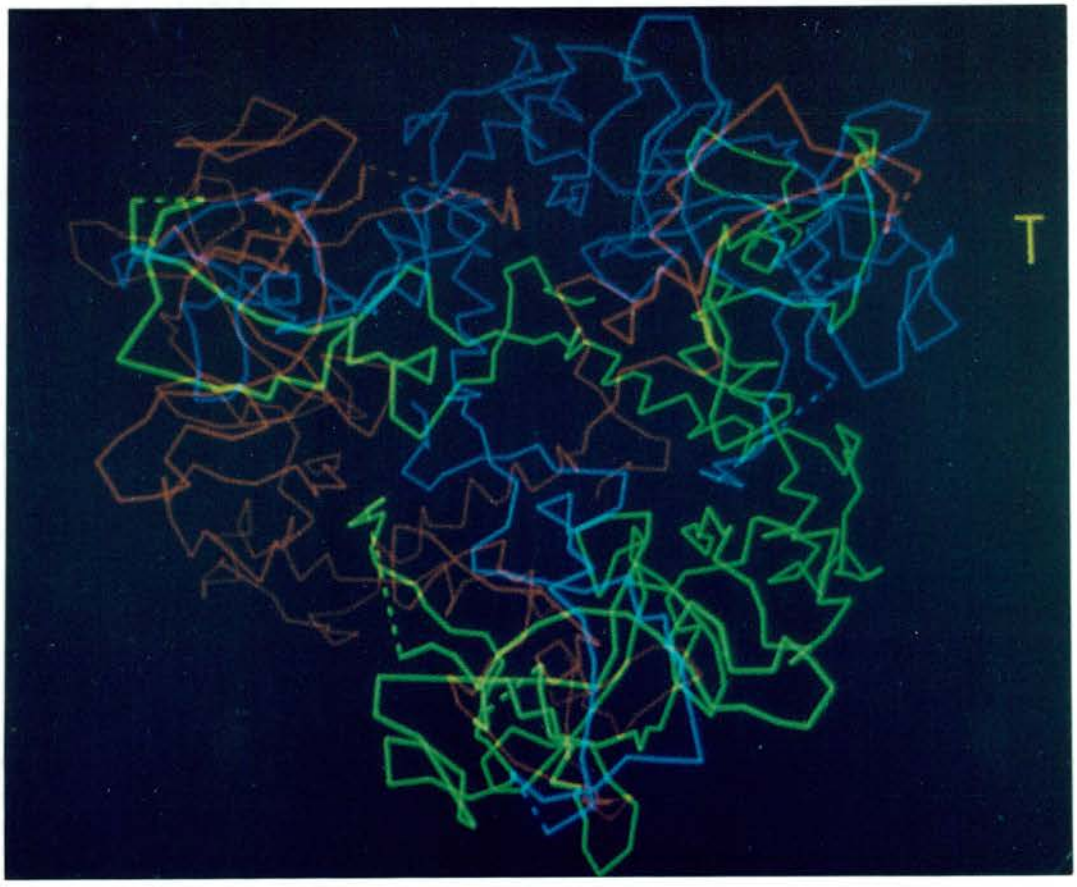
Figure ii. Alpha-carbon trace of the Ad2 hexon sub-unit normal to the base, orientated to be looking from within the central cavity of the trimeric molecule toward the B hexon:hexon contact face. The P1 and P2 domains are composed of two topologically identical eight-stranded flattened ß-barrels, and in the trimer the six barrels form the base of the hexon. The loops l_1, l_2 and l_4 which arise from the ß-barrels form three tower domains (T) which protrudes from the surface of the virion, while the l_3 loops are sandwiched between the P1 and P2 domains. The structures of three small regions remain unresolved and are shown as dotted lines. **(from Roberts *et al*,1986).**

Figure iii. Traces of the Ad2 hexon trimer (A) base and (B) tower domains in the same relative orientations, viewed form above. Each polypeptide of the trimer is given in a different colour. The top is composed of three tower domains (T). A and B denote the different contact faces of the base. **(From Roberts *et al*, 1986).**

(A)



(B)



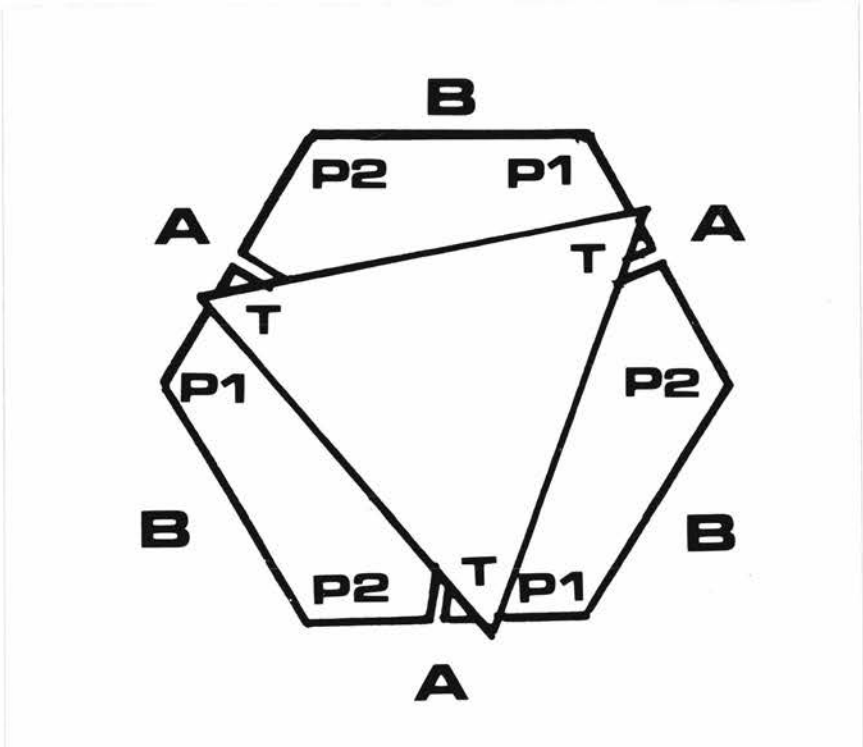


Figure iv. Schematic representation of the hexon trimer. The arrangement of P1 and P2 domains form two distinct faces, A and B. Each A face lies under an apex of the triangular tower.

Hexon possesses a central cavity or channel which varies from round at the base and 22Å across, to Y-shaped at the surface and approximately 7Å across. This channel narrows to 4Å at the base-tower junction. Thus it could be hypothesised that, as well as its obvious structural role, hexon may serve some function in the regulation of the transport of solutes and larger molecules across the capsid.

In addition to the structural differences within the hexon, there is a polarity with respect to charge, the top being predominantly negative and hydrophilic, whilst the base is hydrophobic. Of particular interest is a cluster of 16 contiguous glutamates and aspartates towards the top of the D-strand of the P1 domain of the Ad2 hexon, and evidence suggests that these may be responsible for pH-induced conformational changes at the surface of the virion. The *in vitro* formation of planar arrays and three-dimensional crystals of hexons can only be demonstrated below pH5 suggesting that hexon interactions are directly affected by pH. It has been observed that Ad2 hexon crystals formed at pH3.2 can be raised to pH5 but will crack if taken to pH5.5 (Burnett, 1985). This effect occurs near the isoelectric point and suggests an alteration in hexon packing which could be explained by a charge-induced conformational change. A similar acidic region is found in Ad5.

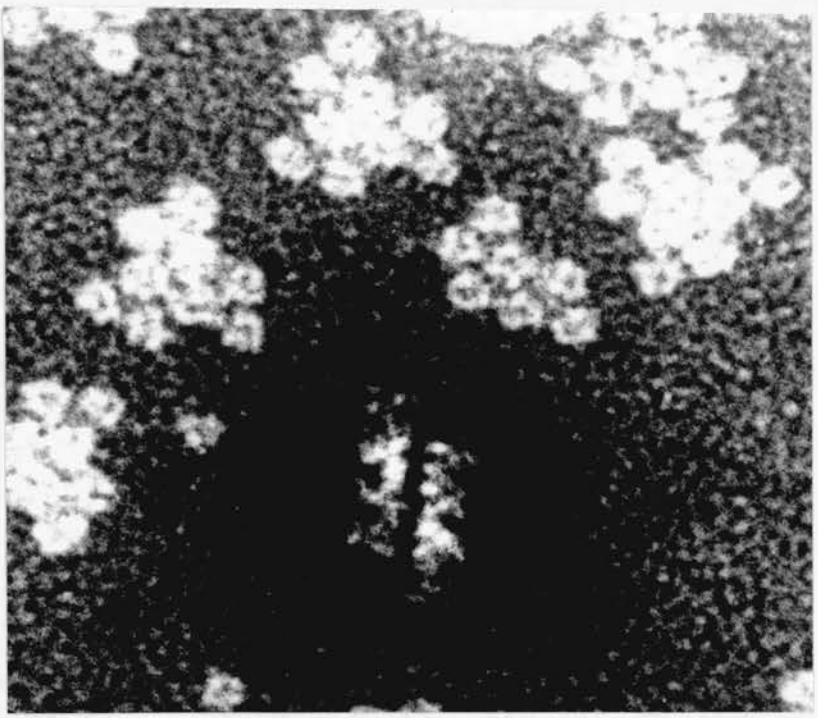
The capsomers lying at the vertices are surrounded by five neighbours, the peripentonal hexons (Ginsberg *et al*, 1966), and this led to the name penton. Penton consists of penton base in complex with a fibre which projects from the surface of the capsid, the interaction probably being hydrophobic. Fibres vary in length from 100Å to 370Å depending on serotype, and are

generally characteristic of subroup. The subgroup F adenoviruses possess fibres of a similar size to those of subgroup C (de Jong *et al*, 1983; Kidd and Erasmus, 1989; Pieniazek *et al*, 1989), which range from 230-310Å. Molecular weight studies predicted penton base to be trimeric and fibre dimeric (Green *et al*, 1983; Deveux *et al*, 1982), and the primary sequence of the Ad2 fibre polypeptide led to a structural model being proposed (Green *et al*, 1983). This comprises a short tail region at the N-terminus, a shaft of repeating 15 amino acid units and a knob at the C-terminus. Each repeating unit contains two ß-strands and two ß-bends, the strands forming a long narrow sheet. The sequences of other serotypes appeared to support this model, the fibre length being dictated by the number of repeating units in the shaft, which in Ad2, is 22. Correspondingly, the Ad40 (Kidd and Erasmus, 1989) and 41 (Pieniazek *et al*; 1989) proteins would predict the number of units to be 21 and 22 respectively. However more recent evidence is strongly in favour of a pentameric penton (van Oostrum and Burnett, 1985) and a trimeric fibre (van Oostrum and Burnett, 1985; Albiges-Rizo and Chroboczek, 1990), and this is further supported by electron microscope and X-ray diffraction analyses of fibre crystals (Devaux *et al*, 1990; Ruigrok *et al,* 1990). This indicates that a revision of the model, particularly in the shaft region is required. The N-terminus is well conserved across serotypes possibly indicating an anchorage function and reflecting structural constraints imposed by the penton base. Amino acids situated at the shaft : knob junction also display a high degree of conservation and may contain the site for interaction with cell receptors. Comparison between the type 40 and 41 sequence reveals high homology except in the shaft, with a

98% homology in the knob region. A more recent publication by Pieniazek *et al* (1990) reports the existence of a second Ad41 fibre gene upstream to that given previously (Pieniazek *et al*, 1989), with a clearly identifiable knob and tail domain, and a shaft which would be comprised of only 12 repeating units. Much of the sequence of this smaller gene is homologous to the first reported Ad41 fibre gene, and to that of Ad40. In the published type 40 fibre sequence paper (Kidd and Erasmus, 1989), an upstream open reading frame was noted encoding a 24kD polypeptide, for which no counterpart in Ad2 could be identified. This polypeptide exhibits striking similarity to the C-terminus of the putative second Ad41 fibre, and has regions of identity with respect to the Ad40 fibre. It would therefore appear that a second fibre gene exists in both enteric serotypes, and were these to be expressed, would make the subgroup F adenoviruses similar to avian serotypes in possessing fibres of two sizes.

Under mild conditions the virus dissociates, first releasing the pentons in conjunction with their peripentonal hexons, followed by 20 groups-of-nine hexons (GONs) (figure v). A model for hexon : hexon interactions has been proposed (Burnett, 1985), based on molecular close-packing and electron micrographs of GONs. This model predicts asymmetric contacts between A and B faces of neighbouring hexons within the GON, while the facets are joined by symmetric A:A or B:B contacts (figure v). Assembly studies with virion hexons suggest that at physiological pH, hexons form rings more readily than they form facets thereby implying that the vertex may be the nucleation site for the assembly of hexons into the virion (van Oostrum *et al*, 1987). The non-random dissociation pattern of the capsid into GONs can be explained by the location

(A)



(B)

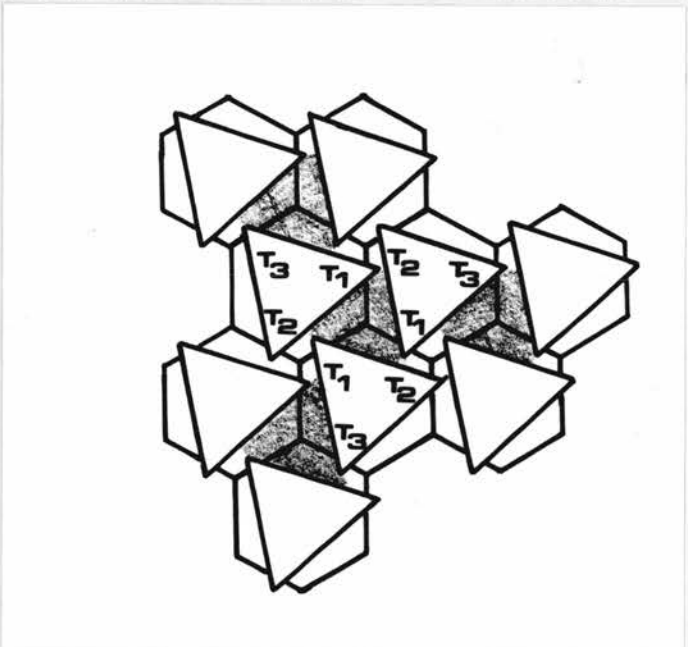


Figure v. (A) Electron micrograph of groups-of-nine (GON) hexons, released from the virion under mild dissociation. (B) Schematic diagram showing asymmetric A:B contacts within the GON. $T_{1,2}$ and $_3$ represent the three tower domains and show the different orientations of the hexons within the GON. The shaded regions show the proposed locations of polypeptide IX (from van Oostrum *et al*, 1987).

of the 14kD protein, polypeptide IX. Using scanning-transmission electron microscopy (Furcinitti *et al*, 1987), VP IX has been directly visualised as four trimers situated between the hexons which form the four large cavities on the upper surface of each GON (figure v). Its location on the surface of the capsid is supported by the observation that antibodies raised against VP IX will aggregate virions. Experiments with a deletion mutant of Ad5 have shown that infectious virus can be successfully assembled in the absence of VP IX (Colby and Shenk, 1981), however the mutant virus is more heat-labile than the wild-type, and releases individual hexons rather than GONs on dissociation. Both these observations suggest that polypeptide IX acts as a cement which stabilises the capsid.

At least three other polypeptides are capsid-associated, however their positions and roles are unclear. On treatment with pyridine the capsid dissociates and the hexons are released bound to a 24kD protein, polypeptide VI. Two copies of VP VI are present per hexon suggesting that it binds as a dimer. Since it is not detected by external enzymatic labelling of the virion it is assumed to be internal (Everitt *et al*, 1975). Cross-linking experiments reveal an interaction between VPVI and the core protein VP V (Everitt *et al*, 1975), and studies by Russell and Precious (1982) show it to possess DNA binding activity. These observations indicate that VP VI may provide a link between the capsid and the inner core. Protein IIIa is present in 60 copies per virion and is released along with penton under conditions which exclusively release components from the vertices (Everitt *et al*, 1973). Antibodies to VP IIIa cause

the virus to aggregate (Everitt *et al*, 1975), implying that it is at least partially exposed on the surface. Crosslinking experiments reveal close contact with penton, VP V and the peripentonal hexons, thus VP IIIa may extend through the capsid to the interior. Temperature-sensitive mutants lacking pIIIa fail to assemble virions properly. A third minor virion component about which little is known is VP VIII, found by sequential degradation to be weakly associated with hexon, although purified GONs appear to lack this protein. Because of its basic nature it is thought to be located on the inside of the capsid.

3.2. The virus core.

Virus particles mildly dissociated with sodium deoxycholate lose their capsid leaving only the virus core (Russell *et al*, 1971). The major components of the core are two basic proteins VP V and VP VII, the 35 kilobase-pair linear double-stranded DNA and two copies of a terminal protein covalently bound to each 5' end of the genome. The core is thought to be icosahedral with an edge of 380Å, and comprises a surface protein shell enclosing a chromatin-like DNA-protein complex. Electron microscopy of freeze-fractured negatively stained virions clearly reveals the presence of a shell enclosing the central core. The protein composing the shell has not been identified, however since VP VII forms such a close interaction with the DNA, it could be the other core protein, VP V. Crosslinking studies have shown that VP V is in contact with penton and the VP VII-protein complex, suggesting that it forms a bridge between the capsid and the core. Another possibility is VP VI, also shown to have core and capsid associations. There is no visible gap between the core

and the capsid implying that both structures possess the same icosahedral morphology. There is evidence that the pentons contribute structurally to the vertices of the core shell, since their removal renders the DNA sensitive to nuclease attack. Thus the shell, in addition to having a structural function in providing rigidity to the virion, could also play a physiological role in protecting the genome.

The organisation of the nucleoprotein complex in mature virus particles has been a controversial subject for many years. A eukaryotic-like nucleosome arrangement was the favoured model for some time, however more recently the following observations were made using Ad5 (Wong and Hsu, 1989), which suggests that the DNA is arranged in a series of supercoiled loop domains. Electron microscopy of virus particles exposed to chemical crosslinkers to fix the DNA conformation, allowed direct visualisation of supercoiled loops. The crosslinked sequences at the base of the loop were therefore in very close proximity, perhaps in the form of a tetrahelix. In addition, the DNA was found to be sensitive to Bal31 nuclease whose endonuclease activity is dependant on the DNA being in a supercoiled conformation and under torsional stress. Treatment with topoisomerase which induces relaxation of supercoiling rendered the DNA insensitive to Bal31. Limited digestion with Bal31 revealed that although each loop contained more than one site for cleavage they were mutually exclusive, presumably because once a break has been introduced the supercoiling would be eliminated, leaving the rest of the loop insensitive to further cleavage. By analysing the grouping of these mutually exclusive sites

and from the EM observations, Wong and Hsu predict there to be eight loop domains each being comprised of approximately 12% of the genome, and that the two ends of the genome containing the replication origins and the E1a and E4 promoters are outside these domains. Supercoiling of linear DNA requires a means of maintaining the rotation of the double helix, which would otherwise be released through rotation of the free ends. It is predicted that this function is performed by proteins binding at the boundaries of the supercoiled loops, and it is thought that these, or other proteins, somehow anchor the loops to the centre of the core. The arrangement of coding regions along the genome lends itself well to this model and it can be hypothesised that the domains are comprised as follows (the products of the coding regions are described later; briefly E1-E4 represent transcription units which code for early genes and L1-L5 code for the late genes).

loop1	2-14	map unit	E1
loop2	14-26	"	major late promoter, tripartite leader sequences, E2b
loop3	26-38	"	L1
loop4	38-50	"	L2
loop5	50-62	"	L3,
loop6	62-74	"	E2a
loop7	74-86	"	L4, E3
loop8	86-98	"	E4, L5

Viral DNA replication assays using subviral particles as templates (Leith *et al*,

1989), have shown that replication *in vitro* can proceed for at least 2000-3000 bases from either terminus. This suggests that the adenovirus genome is packaged within the virion in a conformation which will allow at least initial replication events to occur without the genome being uncoated. It could be that these regions of the genome are not included in the loop domains and therefore devoid of core proteins as suggested by Wong and Hsu (1989), or it may be that elongation can proceed in the presence of these basic proteins. If the termini are exposed, this conformation would lend itself well to early transcription, since the regions apparently available for replication covers almost the entire E1 and E4 transcription units, and it could be proposed that there is a topological mechanism for regulation of adenovirus gene expression. The structural and functional roles of the core proteins with respect to the loop domain model have not been elucidated, however, since protein VII has been shown to form a highly stable complex with the DNA it is suggested this protein may play an important role in conformation.

Studies using the viral protease deficient mutant H2ts1 which fails to process precursor polypeptides, reveals the presence of a core component with a molecular weight of 11kD. This has been identified as a product from the late transcription region 2 and a precursor to the virion component X (now thought to be two species of a similar size) and a small basic protein, mu. Mu binds very tightly to double-stranded DNA and can precipitate it from solution. It has been suggested that mu plays a role in condensing the genome to a size which can fit into the capsid (Anderson *et al*, 1989). Roles, if any, for the

majority of small molecular weight products resulting from precursor cleavage, have yet to be defined.

4. VIRION ASSEMBLY.

Late in adenovirus infection, host cell protein synthesis is shut down and viral message is preferentially translated. In Ad2 most virion components are vastly over-expressed, with perhaps only 10-20% being incorporated into mature viruses. The majority of polypeptides are transported rapidly to the nucleus during which time the capsid components begin to assemble into their respective structural units. Since the monomers of hexon, penton base and fibre are insoluble, it can be assumed that assembly into the freely-soluble multimeric forms occurs fairly instantaneously, although it can be up to 24hrs before the assembly of penton from penton base and fibre is complete. The formation of hexon capsomers has been shown to be dependent on a virally encoded 100kD phosphoprotein, which is present in large quantities in infected cell extracts but absent from mature virions. Data suggests that this polypeptide acts as a scaffolding protein, catalysing the folding of hexon into trimers. Experiments with a 100k-defective temperature-sensitive mutant reveal high levels of hexon polypeptide accumulating in the cytoplasm and a failure to assemble trimers. Antibodies prepared against the 100k polypeptide immunoprecipitate both 100k and hexon from infected cell extracts, implying a close interaction between the two (Cepko and Sharp, 1982). In addition, 100k antibodies inhibit hexon trimerisation *in vitro*.The production of monoclonal antibodies to the 100k polypeptide and to the hexon, which recognise different

antigen conformations, have allowed further insights into hexon morphogenesis. It has been demonstrated that 100k can become associated with the nascent hexon polypeptide on the ribosome and, since only pre-trimer hexon has been found in complex, it is assumed that dissociation occurs immediately following successful folding. The majority of 100k is found in oligomeric complexes to which monomeric hexon can bind. It is probable that three hexon polypeptides become associated with this complex, thus facilitating trimerisation. This scaffolding function is thought to be located in the C-terminus, although an involvement by the N-terminus has not been ruled out. A number of 100k temperature-sensitive mutants have been produced which permit trimer assembly, but not the production of capsids, thus indicating a further role for this protein in virion morphogenesis. The transportation of trimers to the nucleus is affected in the 100k mutant H2ts118, a defect shown by marker rescue to be located in the C-terminus, however a polypeptide VI mutant also fails to transport hexon. Rather than 100k being directly involved, it could be that subtle misassembly of trimers by H2ts118 affects the association between hexon and a transport protein (perhaps VP VI), thus inhibiting movement to the nucleus. The mutant H2ts107 has the ability to form hexon trimers but fails to assemble capsids at the non-permissive temperature. A subsequent reduction in temperature to 33°C results in the production of infectious virus, suggesting that 100k may play a scaffolding function in capsid formation. The H2ts107 mutation was mapped to the centre of the protein, between amino acids 300 and 400. In addition to 100k, large amounts of a 90k protein was found, and this was identified as a cleavage product of 100k.

Assembly intermediate particles of H2ts107 were obtained following temperature reduction, and the location of these two proteins examined by co-precipitation with antibodies to a variety of other capsid structures. 100k was associated with the vertex components III and IIIa, while 90k was found at the faces and edges of the capsid, and this could suggest that the non-vertex capsomers have been assembled for a longer time. Affinity chromatography has shown that 100k possesses a binding site for trimers, and this site is different from that which binds monomer (Morin and Boulanger, 1986).

Penton appears not to be produced in vast excess and therefore could be the rate limiting factor in adenovirus assembly. At 13hrs post-infection (p.i.). both virions and empty capsids can be recovered from cells, both being maximally produced at 24hrs p.i.. Labelling experiments indicate that hexon is very quickly assembled into immature capsid shells but production of mature virions takes much longer. Groups-of-nine hexons obtained by dissociating virus with deoxycholate can re-aggregate under acidic conditions to form dimers, rings of five and icosahedral shells (Pereira and Wrigley, 1974), however since it is not known whether the formation of GONs occurs prior to the production of empty capsids, it remains unclear as to whether this is the initiating step in virion assembly. Pentons and VP IIIa are also associated with immature particles but it is not known at what stage these or the other minor capsid proteins are incorporated. In addition there are three polypeptides, a 33kD, a 40kD and VP IVa_2, found only in immature capsids, and these, like 100K, may be scaffolding proteins which are lost following successful

assembly. Following capsid formation the DNA and core proteins are inserted. The DNA enters preferentially from the left-hand end, insertion being dependent upon a packaging signal located in the first 400 nucleotides of the left-hand end of the genome. The signal is cis-acting suggesting that it does not code for a protein, and this dependance on a packaging signal explains why cellular DNA does not get inserted into particles. The final maturation stage in virus assembly is the proteolytic processing of precursor polypeptides by a virally-encoded protease. A temperature-sensitive mutant of Ad2, H2ts1 (Weber, 1976) which was unable to process precursors, was found to contain a mutation in a gene coding for a 23kD polypeptide. It has therefore been proposed that this is the viral protease responsible for post-translational modification of a number of precursors including pIIIa, pVI, pVII, pVIII and the terminal protein. Viral particles are assembled in the absence of protease, however these are non-infectious although they do contain DNA which retains infectivity. Antigenic analysis of the vertex components of immature virions indicate that polypeptides III, IIIa and IV are not recognised by antibodies to the mature proteins until very late in the infectious cycle after the final proteolytic step has taken place. The protease has been characterised as a cysteine protease and a consensus sequence for its cleavage site has been proposed (Webster *et al*, 1989a,b). A number of different sequences can be cleaved by the enzyme, the three-dimensional structure of the recognition site being more important than the actual sequence. Following proteolysis the production of infectious virus has been accomplished and the cycle is complete.

5. THE STRUCTURE OF THE GENOME.

The nucleotide sequence of the entire genome of adenovirus type 2 has been determined, revealing a complex and compact arrangement of the genetic material. Data obtained for other serotypes suggests a general conservation of genetic organisation despite substantial sequence diversity. The locations of the major transcripts and the corresponding proteins are shown in figure vi. The adenovirus genes can be divided into the early genes expressed before the onset of viral DNA replication, and the late genes which are expressed after. In addition there is a group of intermediate genes which are transcribed prior to replication but are also present late in the infectious cycle. Early transcripts can be sub-divided into pre-early (which reach a maximum in under 2hrs post-infection (p.i.)) and early (2-6 hrs), while intermediate transcripts are maximal at 6-12 hrs p.i and late, at 12-36 hrs. The last three of these four stages are all regulated by virally-encoded products. A number of promoters are scattered throughout the genome, different ones being utilised at different stages of infection. There are six known promoters for the early genes, three for the intermediate and one, the major late promoter, for the late. Apart from polypeptide IX mRNA, all primary transcripts are processed into families whose members share common 5' and 3' ends, but differ in their internal splicing. The major transcription units of Ad2 and the more important gene products are discussed below. Limited sequence information is available for the corresponding subgroup F adenovirus gene products, and is included where relevant. By convention, the adenovirus genome is divided into map units (m.u.), 0 representing the extreme left-hand

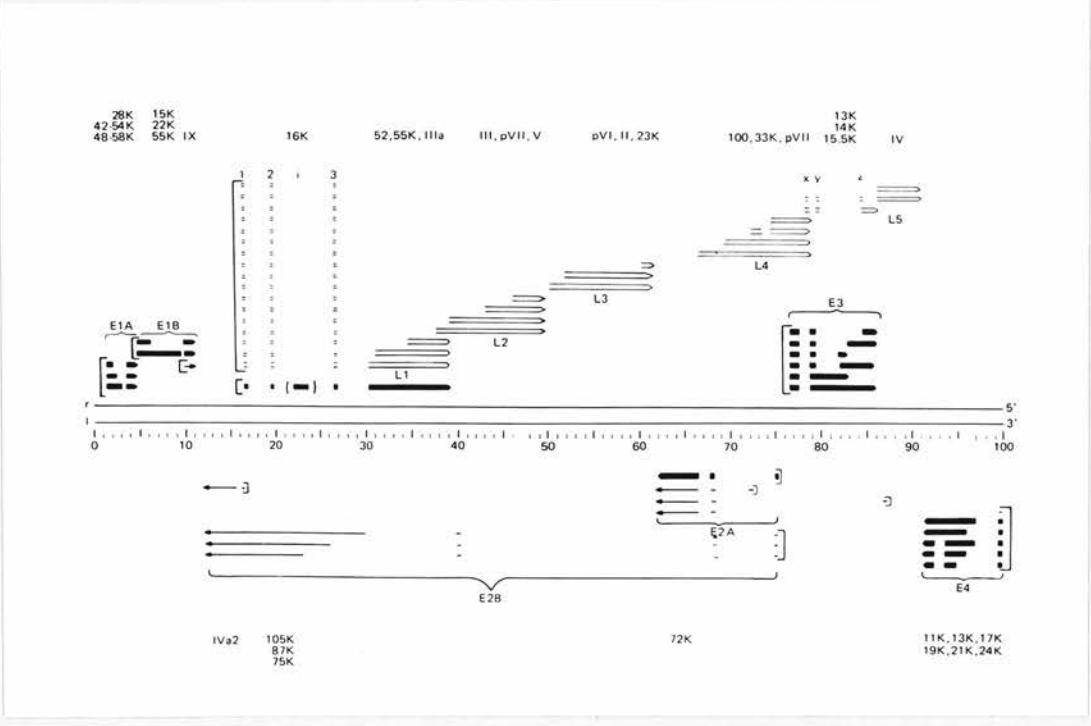


Figure vi. The genome (0-100 map units) of adenovirus type 2 showing the locations of the transcription units, the major transcripts including leader sequences, and the proteins. Early, intermediate and late transcripts are represented by solid, narrow and open arrows respectively. **(From Ginsberg H.S., 1984 p. 47).**

end and 100, the extreme right. The right strand is transcribed rightwards and the left strand leftwards.

Early region E1 (m.u. 1.3-11.2) is located at the left-hand end of the right strand and is sub-divided into E1a, E1b and region IX. E1a and E1b code for proteins which have important functions in the lytic cycle, cellular transformation and oncogenicity. The E1a region is transcribed from a promoter which functions in the absence of viral gene expression, producing two major mRNA species. E1a gene products directly or indirectly stimulate transcription of other early genes, and this function is essential in productive infection. E1a is also able to immortalise primary cells but requires E1b for complete transformation. In addition, E1a products can bind to certain cellular and viral enhancers thus regulating gene expression. During lytic infection the Ad5 E1b promoter gives rise to two differentially spliced mRNA species, a 22S encoding the 19K or 55K polypeptide depending on which initiation codon is used, and a 13S encoding only the 19K product. In Ad5 and Ad12 the 55K protein is required for a complete lytic cycle in HeLa cells, where it is involved in conjunction with an E4 protein, in the preferential transport of late viral mRNA from the nucleus to the cytoplasm (Bridge and Ketner, 1990), and is also responsible for switching off host cell functions (Babiss and Ginsberg, 1984). Late in infection viral mRNA is preferentially accumulated in the cytoplasm and is exclusively translated (Babiss *et al*, 1985; Williams *et al*, 1986). In contrast newly synthesised cellular mRNA fails to accumulate in the cytoplasm despite

continued nuclear synthesis, while translation of existing mRNA is inhibited. Ad12 E1b mutants indicate that there is a requirement for 55K in early mRNA expression and in DNA synthesis (Shiroki *et al*, 1986, Fukui *et al*,1984). Both polypeptides appear to be required for efficient, virion-mediated transformation although E1b has no transforming activity alone, and neither are required to maintain transformation once initiated. In addition to its role in transformation, the 19K protein plays a role in regulation of viral early gene expression, and in protection of newly synthesised viral DNA (White *et al*, 1988; Subramanian *et al*, 1984). Mutants lacking the 19K product display a large plaque phenotype, growing well in HeLa cells, but inducing rapid cytopathicity accompanied by degradation of cellular and viral DNA.

Subgroup F serotypes cannot be propagated in HeLa cells and, since viral DNA synthesis is undetectable, it is assumed that the block is prior to DNA replication. Both serotypes grow well in Graham 293 cells which constitutively express Ad5 E1a and E1b functions. The left terminal 12 and 11% of the Ad40 and 41 genomes respectively, are able to at least partially transform rodent cells *in vitro* (van Loon *et al*, 1985), and in this respect show certain similarities with other serotypes. The E1 regions of both genomes have been sequenced (van Loon *et al*, 1987a) and examined for the presence of regulatory elements. The first 50 nucleotides exhibit 82% homology to the corresponding region in Ad5, and most of the conserved binding sites for viral and cellular factors involved in DNA replication and found in other serotypes, are present in Ad40 and 41. This would suggest that the fastidious growth of these viruses is not due to peculiarities in this region. A number of regulatory elements required for

efficient expression of E1 genes are present in Ad5. Since some of these are absent from other serotypes it is difficult to predict whether all are likely to be important. However of the conserved elements, most can be identified in both Ad40 and 41 suggesting that the information is available for normal transcription from this region. The mRNA of rodent cells transformed by Ad40 and 41 was examined for the presence of E1 transcripts (van Loon *et al*, 1987a). In the Ad40 transformed cells three E1a transcripts were found similar to those present in other adenoviruses, however no E1b mRNA was detected. In contrast, only one E1a transcript was found in Ad41 transformed cells despite the existence of the necessary regulatory elements for transcription of the other two. One large E1b mRNA species was detectable. These results suggest that if a malfunction in the E1 region is responsible for the poor growth of the fastidious adenoviruses, the precise cause may differ between the two serotypes. Plasmids expressing type40 and 41 E1a have been shown to transactivate the Ad2 E4 promoter, although the level of stimulation is comparatively low (van Loon *et al*, 1987b). Using chloramphenicol acetyl-transferase expression vectors it was observed that the Ad40 and 41 E1a promoters were less active than that of Ad12 for *cis*-activation (Ishino *et al*, 1988). These observations led to the belief that the poor growth of the fastidious adenoviruses may be due to a malfunctioning E1a promoter, the subsequent low levels of E1a products resulting in reduced transactivation of the early genes. To address this question, the growth of Ad40 was studied in cell lines expressing various of Ad2 E1a or E1b functions (Mautner *et al*, 1989). This study revealed that efficient growth of Ad40 could be obtained by

supplying the E1b 55K protein *in trans*, either in cells which constitutively express it or by complementation with a helper virus. There was no such requirement for the 19K protein or for any E1a product. Despite the reported reduced E1a transactivating potential, Ad40 was able to act as a helper for an Ad5 mutant lacking functional E1a, and it can therefore be assumed that the cause of its poor growth is not due to an inability to transactivate early genes. Ad40 could not complement Ad5 55K mutants, however it was able to complement mutants lacking 19K. Analysis of E1b mRNA and proteins during infection by Ad40 in permissive cells (Mautner *et al*, 1990) showed that mRNA was made in small amounts and was only detected after the onset of DNA replication, however there was efficient transcription from the VP IX promoter. The 19K product was detectable only at late times while no trace of the 55K protein could be found. Presumably as a result of the lack of 55K, a failure to inhibit host cell functions was observed. In non-permissive cells, trace amounts of E1b mRNA were found in cytoplasmic extracts, but only after 42 hrs p.i. as opposed to 6 hrs for Ad5. This late appearance was not due to an inability to transport mRNA to the cytoplasm, as no accumulation in the nucleus was observed. It would seem likely that the gut cells which obviously support a high level of replication of Ad40, either possess a factor which enhances transciption from the E1b promoter, or which allows very efficient translation of E1b mRNA. Alternatively they may supply a cellular E1b-like function which eliminates the requirement for viral E1b.

Region IX contains the gene for the structural polypeptide VP IX, whose promoter lies within the 3' end of E1b. Transcription takes place after DNA

replication when it becomes the major product from the E1 region.

Late and intermediate genes (m.u. 11.2-31). This region contains numerous important sequences in both directions including the major late promoter (MLP) on the right strand. All mRNAs utilising the MLP commence with a tripartite leader sequence of approximately 200 nucleotides which significantly enhances the efficiency of translation of mRNA late in infection but not at early times. This leader is derived from the splicing together of three separate sequences from the right strand and does not contain an initiating AUG signal.

Two small RNAs known as virus-associated (VA) RNAs, for which no protein product has been found are transcribed from this region, and it is believed that they are involved in stabilising or splicing late mRNA. Both are transcribed by RNA polymerase III rather than RNA polymerase II. VA RNA_1 is involved in maintaining translation of viral message by preventing the activation of the double-stranded RNA-activated protein kinase, thus counteracting one of the cellular responses to adenovirus infection.

The left strand contains two regions of importance, the gene for the intermediate protein IVa_2 which plays a role in virion assembly (Persson *et al* 1979), and the early region E2b. E2b has been classified as an intermediate transcription unit and all mRNAs contain a tripartite leader sequence derived from spliced upsteam sequences. The major products of E2b are the adenovirus-specific DNA polymerase and the pre-terminal protein (pTP). pTP is covalently bound to the 5' termini of the adenovirus genome and plays an

important role in the initiation of viral DNA replication. Late in infection it is cleaved by the adenovirus protease to the 55kD terminal protein which remains bound to the DNA and may function in protecting the genome from exonuclease attack.

Late regions L1, L2 and L3 (m.u. 31-62). Early in infection the MLP is functional but transcription only proceeds as far as 39 m.u., which corresponds to the end of the L1 region. Later on in the infectious cycle there is a major switch bringing about the activation of the whole late transcription unit up to 99 m.u. The product from this unit is processed to yield five families of mRNA, L1-L5. All late mRNAs share a common 3' end within each family. The L1 mRNAs are transcribed from the MLP, both before the late switch and after, giving rise to proteins of 52k and 55k, both of unknown function, and to VP IIIa which is capsid associated. Region L2 yields penton, polypeptide V and the precursors to the core components, VP VII and VP X. Hexon, polypeptide VI and the 23k protease are the major products of the L3 family of mRNAs.

Early region E2a (m.u. 67-75). This region on the leftward transcribed strand codes for the multi-functional 72kD single-stranded DNA binding protein (DBP) which is involved in DNA replication, regulation of mRNA formation, host range specificity and cell transformation. Early in infection mRNA is transcribed from a promoter under the control of E1a, however at late times there is a promoter switch and E1a is no longer involved. The DBP gene and leader sequences for both subgroup F viruses have been determined (Vos et el,

1988). The promoter elements, splice sites and polyadenylation signals which flank the DBP gene in Ad2 can be identified in Ad40 and 41. However the non-coding regions are more compact and there is little sequence conservation even in the leader. The type 40 and 41 DBP genes share an overall homology of 72% and are of similar lengths, however they are considerably shorter than the DBP's of other serotypes. In Ad5 DBP has been divided into the N- and C- domains. The N-domain which is structurally and functionally important, shows a low level of conservation across serotypes and is shortened in both types 40 and 41. The C-domain which is involved in DNA replication and regulation of late gene expression, contains areas of homology between Ad40 and 41 and Ad2. Three areas believed to be important in DNA binding and an area which is potentially a metal-binding site, are all well conserved.

Late region L4 (m.u. 66-77). The major proteins from this region are the 100k polypeptide, the polypeptide VIII precursor and a 33kD protein. The 100k is a multifunctional protein. As well as its role in hexon folding and capsid morphogenesis discussed previously, it is thought to be required for efficient translation of viral late mRNA (Hayes *et al*, 1990). 100k binds to cytoplasmic RNA (Adam and Dreyfuss, 1987) and contains sequences conforming to the consensus RNA recognition motif found in RNA binding proteins. A link between the ability of 100k to bind to mRNA and its ability to stimulate translation is likely but yet to be established. A large portion of the 100k gene of type 40 (Vos *et al*, 1988), and all of the type 41 gene (Slemenda *et al*, 1990)

have been sequenced revealing a high degree of homology between the two. There are however, areas of marked divergence from the Ad2 and Ad5 sequences, particularly in those regions corresponding to the N-terminus of the protein. Since the functional domains of the 100K protein have not yet been fully mapped, the significance of these non-homologous regions remains unknown. The 33k protein which is also thought to be involved in virion morphogenesis, is coded for in the Ad41 genome (Slemenda *et al*, 1990).

Early region E3 (m.u.76-86). A large number of mRNAs are transcribed from this region, however few proteins have been identified. E3 is not essential for virus replication in tissue culture, although most of the proteins are products of conserved genes which suggests a functional importance *in vivo*. The E3 proteins characterised so far appear to play a role in avoidance of the host immune response to infection. The 19kD glycoprotein interacts with HLA class 1 antigens, preventing cell surface expression and thus protecting the cell against lysis by adeno-specific cytotoxic T-cells. The 14.7kD protein inhibits cytolysis by Tumour Necrosis Factor whose levels are stimulated by adenovirus infection.

Late region L5 (m.u. 86-91). L5 produces two mRNA species that code for a single protein, the fibre (VP IV), which in Ad2 has a molecular weight of 62kD. The fibre contains sites responsible for the interaction with cellular receptors and for haemagglutination. The sequence of the Ad40 (Pieniazek *et al*, 1989) and Ad41 (Kidd and Erasmus, 1989) fibres have been determined, predicting

proteins of just under 60kD. There is little overall conservation of amino acids between the serotypes sequenced to date, except in the tail of the shaft and the base of the knob suggesting that these are important functional domains. A second shorter fibre gene has been identified in Ad41 which predicts a protein of 41kD, and which also appears to be present in Ad40. However it is not known whether these are expressed during infection.

Early region E4 (m.u.91-94). At least six polypeptides are produced from this region located on the leftward transcribed strand. All share common 5' and 3' termini but differ in splicing. Translation of the E4 proteins reaches a maximum at 3 hrs post-infection, and thereafter declines. Roles have yet to be assigned to most of the E4 products, however they appear to be involved in host-cell shutdown and in the stabilty of late nuclear RNA. The 34kD protein interacts with the E1b 55kD product and together, act to inhibit cellular functions. The 17k product functions as a transactivator, enhancing transcription from the E2 region. E4 mutants exhibit a reduction in late protein synthesis.

From the limited sequence information available it seems that the subgroup F adenoviruses share the same overall genome organisation as the more conventional serotypes. There are areas of high amino acid conservation in the proteins whose sequences have been predicted, particularly in the functionally important domains. In general the non-coding regions appear to be more compact than those of Ad2. More information is required however, to

account for the unique properties of Ad40 and 41 with respect to their growth characteristics and tissue specificity.

6. ANTIGENICITY.

Two adenovirus proteins, the hexon and the fibre, are the major targets of neutralising antibody (Wilcox and Ginsberg, 1963; Kjellen and Pereira, 1968), and both carry a complex arrangement of antigenic sites.

Hexon. All human adenoviruses possess a common group-specific antigenic determinant (α), which is located within the hexon, but which is not immunoreactive in intact virus. In addition, hexon carries a type-specific component (ε), and inter- and intra-subgroup specific epitopes (Norrby, 1969; Norrby and Wadell, 1969). It is generally accepted that the type-specific moiety is responsible for virus neutralisation. Cross-absorption studies using anti-hexon antibodies have revealed relationships between hexons of different serotypes. For example Ad12 from subgroup A, shares few common hexon determinants with any of the other serotypes, whereas Ad4, the only member of subgroup E, displays close similarity with subgroup B serotypes, particularly Ad16. In addition, Ad4 shares epitopes with the pentons of subgroup B serotypes, and is closely related to subgroup C adenoviruses with respect to haemagglutination. Common intragroup antigenic sites have been demonstrated using sera raised against purified Ad5 hexon which is capable of neutralising infectivity of both type 5 virus and of the closely related type 1 (Kjellen and Pereira, 1968). Immunological cross-reactivity between the

surface antigens of the enteric adenoviruses and other serotypes has been examined using solid-phase immune electron microscopy (Svenssonn *et al*, 1983). An extensive two-way cross-reaction was observed between Ad40 and 41, however the only inter-group reaction was between Ad40 virions and antisera to Ad4. Interestingly, this was not reciprocal, and no similar observation was made between Ad41 and Ad4. Haemagglutination-inhibition tests demonstrated cross-reactivity between Ad40 and 41 suggesting common epitopes within the fibre, but no relationship to the other 39 serotypes was observed either by haemagglutination-inhibition tests or by neutralisation (de Jong *et al*, 1983). Recombinants between closely related serotypes have been found to occur naturally, and may possess hexon determinants characteristic of one serotype, and fibre determinants of another. Type-specific determinants have been demonstrated on the surface of the virion (Norrby, 1969; Willcox and Mautner, 1976). Antibodies to the type-specific epitope of the hexon of subgroup B adenoviruses, interfere with the agglutination activity, and this is assumed to be due to steric hindrance following binding of the antibodies to the peripentonal hexons, which thus blocks the site involved in erythrocyte attachment located on the fibre. This effect is only seen in the subgroups which possess short fibres. The distribution of antigenic sites has been studied by partial proteolysis (Pettersson, 1971). Digestion of native Ad2 hexons with trypsin induces three breaks in the polypeptide between residues 140 and 290, however this has no effect on antigenicity. Limited digestion with chymotrypsin or papain, removes more than 50% of the protein and destroys parts of the group-specific determinant (α). The type-specific site (ε) remains unaffected

and can only be destroyed by severe proteolysis, suggesting a very compact structure. Monoclonal antibodies (MAbs) have proved useful in analysing the different hexon epitopes. Native hexon trimers can be distinguished from the monomeric polypeptide by virtue of their different complement of antigenic sites as revealed by MAbs (Cepko *et al*, 1981). In addition, a number of sites which are exposed on trimers and groups-of-nine, appear to be masked in intact virions. Monoclonal antibodies against Ad5 hexons have revealed at least five different epitopes (Russell *et al*, 1981). Three MAbs have been developed by Herrmann *et al* (1987), one of which is specific for Ad40, one for Ad41 and one which recognises both viruses. Each MAb was shown to recognise hexon, and both the type-specific antibodies were neutralising. The MAb which reacted with both subgroup F viruses was non-neutralising, and in addition showed some reactivity with the subgroup C serotypes 2, 5, and 6, suggesting a common epitope.

Fibre and Penton Base. Fibre contains a varying number of epitopes, the antigenic complexity being related to the length. The type-specific determinant (γ) located in the knob, is a feature of all serotypes, and is the only epitope in the subgroup B viruses, which have the shortest fibres. Antibodies to γ block the attachment of virus to red blood cells. Subgroup A, C and D adenoviruses possess another type-specific epitope, ∂, located near the fibre base. Penton base is believed to carry antigenic determinants (Wadell and Norrby, 1969). However the high degree of homology between the types 2 and 5 pentons

(Neumann *et al*, 1988), almost 99%, would suggest that this protein is not subject to immunological pressure.

7. VIRUS PENETRATION.

Little is known about the mechanism by which adenoviruses enter the cytoplasm, however there is increasing evidence for internalisation by receptor-mediated endocytosis (Svensson, 1985). The penton base has been shown to be crucial for the release of virions from acidic endosomes into the cytoplasm. It is believed that the penton undergoes a change in conformation when subjected to the low pH within the vesicles, becoming increasingly hydrophobic and thereby interacting with, and penetrating, the endosomal membrane. It has been demonstrated that hexon sensitivity to cleavage by the endopeptidase dispase, is induced by reducing the pH from 7.5 to 5, thus implying that this protein also undergoes a change in conformation on exposure to the acidic compartment (Everitt *et al*, 1988).

Adenovirus infectivity can be neutralised by treatment with antibodies to fibre, penton and hexon. Anti-fibre antisera causes aggregation of virions which are still able to attach to cells, although only 15% become internalised, and these are not subsequently released into the cytoplasm (Wohlfart *et al*, 1985). This is presumably because the aggregates cannot undergo the necessary change in conformation required for escape from the vesicles. If viruses are allowed to attach to cells before treatment with fibre antibodies, they are not subsequently neutralised. Virions neutralised by anti-hexon and anti-penton antisera are internalised efficiently, however they become trapped within the endosomes

(Wohlhart *et al*, 1985). Virions already attached to cells are still susceptible to neutralisation, implying that the mechanism is not one of aggregation. Instead it would seem that anti-hexon antibodies inhibit the conformational changes in the hexon necessary for the correct exposure of the penton base, and this is supported by the observation that when bound to antibody, hexon remains insensitive to dispase cleavage at low pH. Similarly, the anti-penton antibodies directly cover the penton thus blocking the acid-induced exposure of the sites of interaction with the endosomal membrane. This conformation neutralisation pathway has been demonstrated for other viruses, including polio (Emini *et al*, 1983) and influenza (Possee *et al*, 1982).

Dispase digestion of hexon releases a 15kD product corresponding to amino acids 5 to 135-150. Under neutral conditions this region is expected to be buried on the interior thus not sensitive to dispase. However antisera raised against acid-fixed virions contain a sub-set of antibodies that are reactive against this fragment (Varga *et al*, 1990), and this sub-set is not seen in antisera to untreated or alkali-fixed virions. The exact position of the N-terminal 41 amino acids of the Ad2 hexon is not clear (Roberts *et al*, 1986). Varga *et al* (1990) suggest that by lowering the pH, the N-terminus becomes exposed and this could be mediated by the run of acidic amino acids between positions 146 and 161. Thus it appears that the conformation of the hexon and penton is crucial during adenovirus penetration.

It is clear that the immunological characteristics of adenoviruses are very complex and vary greatly between serotypes. The variation in the surface

properties of these viruses is likely to be a result of evolutionary changes in the genes encoding the surface proteins in order to escape immunological pressure. To gain an insight into the structure and antigenicity of the subgroup F adenoviruses, the nucleic acid sequences of the hexon genes were determined, and thus the amino acid sequences of the corresponding proteins. This information was used to examine the relationship between hexons of different serotypes, and to attempt a prediction of the likely structure of the enteric adenovirus hexons, based on the known Ad2 crystal structure.

EXPERIMENTAL PROCEDURES.

1. Cells and viruses.

Graham 293 cells (Graham et al,1977) were grown in monolayer in Glasgow-Modified Eagle's Medium (GMEM) containing 10% foetal calf serum. To obtain virus stocks, adenovirus types 40 (strain Dugan) and 41 (strain Tak) were adsorbed on to cells for one hour in serum-free medium and maintained in GMEM containing 2% serum. Infection was allowed to proceeded for 3-5 days until a good cytopathic effect (CPE) was observed. Crude virus stocks were prepared from infected cells frozen and thawed repeatedly, and extracted with trichlorofluoroethane (arcton). Pure virus stocks were prepared from arcton extracts by centrifugation through a caesium chloride step gradient. The gradient was assembled from 3ml 2M CsCl in TE (10mM Tris (2-amino-2-(hydroxymethyl)-1,3-propandiol) adjusted to pH8 with HCl, 1mM EDTA (ethylenediamine tetra-acetic acid)) overlaid onto 2ml 3M CsCl in TE. Arcton extracted virus (commonly 7ml per tube) was layered onto the gradient and spun at 30,000 rpm for 90 min. in a Beckman SW40Ti rotor. Banded virus was removed and stored at -20^{o}C in 50% glycerol.

2. Preparation of viral DNA for cloning.

Purified virus was diluted 1:5 in distilled water and treated with Proteinase K (100μg/ml) for one hour at 50^{o}C in the presence of 0.6% sodium dodecyl sulphate (SDS), 10mM Tris pH7.5 and 10mM EDTA. Following phenol and chloroform extraction the DNA was precipitated with 0.3M sodium acetate and 2.5 volumes of ethanol, taken up in 10mM Tris-HCl pH8, 1mM EDTA and stored at -20^{o}C.

DETERMINATION OF THE AD41 HEXON GENE SEQUENCE.

3. Production of probe for identification of the Ad41 hexon gene.

The plasmid pUB54 (obtained from Dr Vivien Mautner, MRC Virology Unit, Glasgow) contains the Ad2 Hind A fragment in the plasmid pAT153. pUB54 was digested with restriction enzymes *SmaI* and *XhoI* to yield a 420bp fragment corresponding to nucleotides 77-497 of the Ad2 hexon gene. This fragment was ligated into the plasmid pUC13 which had been cleaved with *SmaI* and *SalI* and dephosphorylated using calf intestinal alkaline phophatase (CIP) to stop self-religation. All ligations were carried out under standard conditions (Sambrook *et al*, 1989), incubating at room temperature for 16 hours in a reaction volume of 10µl, using T4 DNA ligase as instructed by the supplier. The ligation reaction was transformed into *Escherichia Coli* (*E. Coli*) strain JM83 made competent by the method of Hanahan (Hanahan, 1983). Cells were plated out on agar, and transformed colonies selected with ampicillin (100µg/ml) in the presence of bromo-chloro-indolyl ß-D-galactopyranoside (X-gal) (50µl of a 2% solution in dimethyl formamide). JM83 cells possess a defective lacZ gene which is complemented for by the plasmid pUC13 and enables the use of the chromogenic substrate X-gal to indicate the presence of plasmid, by production of a blue colony. The polylinker cloning site of pUC13 is located within the lacZ gene, hence insertion of a restriction fragment leads to the destruction of the complementation, yielding a white colony on transformation. Thus a colony containing the recombinant plasmid was identified and plasmid DNA prepared by the alkali lysis method followed by

caesium chloride/ethidium bromide centrifugation, as described (Hay *et al*, 1984). The plasmid was designated pCT1. To prepare the probe, pCT1 was digested with *EcoRI* and *HindIII* to release the hexon fragment which was resolved by agarose gel electrophoresis. The small fragment was cut out of the gel, electro-eluted in a small volume and ethanol precipitated. The fragment was radioactively labelled using the nick-translation method (Rigby *et al*, 1977), incorporating α-^{32}P dATP to a specific activity of approximately 50 μCi/μg DNA. Unincorporated radioactivity was removed by gel filtration through Sephadex G-50 (Pharmacia). Probe was never stored for more than two days, due to the inherent instability of internally labelled radioactive fragments.

4. Agarose gel electrophoresis and Southern blotting.

Viral DNA (typically 1-5μg) was digested with restriction enzymes under conditions recommended by the manufacturer. Fragments were resolved on a 1% agarose gel under standard conditions using the Tris-borate-EDTA (TBE) buffer system (0.1M Tris, 0.1M boric acid, 2mM EDTA, pH8.3) and visualised with ethidium bromide. Southern blotting was performed following the method of Southern (Southern, 1975) with minor modifications. Briefly, following electrophoresis the gel was placed in 1.5M NaCl, 0.5M NaOH for 1hr to denature the DNA, then neutralised in 1M Tris-HCl pH7.4, 1.5M NaCl. DNA was blotted onto a Hybond-N nylon membrane (Amersham) by capillary transfer in 10xSSC (1xSSC=0.15M NaCl, 0.015M tri-sodium citrate) and covalently bound with a 2 min. exposure to short-wave UV irradiation. The membrane was agitated gently in prehybridisation buffer (6xSSC, 5xDenhardts

(50xDenhardts=1% bovine serum albumen, 1% Ficoll, 1% polyvinyl pyrrolidone, 0.01% sodium azide), 0.1% SDS, 20μg/ml denatured salmon sperm DNA) for 30-120 min. at 65°C. Double-stranded radioactive probe was denatured with 0.5M NaOH for 10 min., neutralised with 0.5M acetic acid, added to the membrane in hybridisation buffer (25mM Tris-HCl pH 7.5, 7xSSC, 12xDenhardts,1mM EDTA, 0.5% SDS, 60μg/ml denatured DNA) and incubated for 16 hours at 65°C. The membrane was subjected to two 1hr washes in 2xSSC, 0.25% SDS, air-dried and exposed to X-ray film overnight.

5. Cloning the Ad41 hexon gene.

Ad41 DNA was digested with a series of restriction enzymes (1μg/digest), run on a 1% agarose gel and Southern blotted as described above. Hexon containing fragments were identified using the pCT1 fragment as a probe. *HindIII* digestion yielded a single hybridising species of approximately 1.9kb. 1μg Ad41 DNA was digested with *HindIII* and the fragments ligated into 100ng of pUC13 which had been cleaved with *HindIII* and dephosphorylated with CIP. This was transformed into competent JM83 cells as described previously. White colonies were streaked out onto fresh agar in duplicate and grown overnight. The colonies from one plate were lifted onto a nylon membrane, placed into denaturing solution (0.5M NaOH, 1.5M NaCl) for 7 min. followed by neutralising solution (1.5M NaCl, 0.2M Tris-HCl pH7.2, 1mM EDTA) for 3 min. After UV-irradiation the colony lift was probed with the hexon fragment from pCT1. A positively hybridising colony was used for large-scale preparation of plasmid, designated pCT3. Following DNA sequencing the insert present in

pCT3 was found to contain the 5' end of the hexon gene with a *BglII* site 120bp from the 3' end. pCT3 was digested with *BglII* and *HindIII* and the 3' terminal 120bp fragment isolated. This was nick-translated and used as a probe to identify an overlapping fragment large enough to contain the remainder of the hexon gene. 1μg Ad41 DNA was digested with *BglII* and a series of second enzymes. The procedure for identifying hexon containing fragments was as for the production of pCT3, using the pCT3-generated 120bp probe instead of the Ad2 420bp fragment. *BglII-EcoRI* digestion yielded a single hybridising species of approximately 2.7kb. This fragment was ligated into *BamHI-EcoRI* cleaved pUC13 following the procedure used to generate pCT3. By probing colony lifts the plasmid pCT6 was identified and isolated.

6. DNA Sequencing.

The plasmid pCT3 was cleaved with *HindIII* and the hexon fragment cloned into the *HindIII* site of the bacteriophage M13mp8 replicative form (RF) DNA. The reaction was transfected into competent *E.Coli* JM101 cells and recombinants selected as white plaques in a top agar supplemented with X-gal and isopropyl ß-D-thiogalactopyranoside (IPTG) (Messing *et al*, 1981). 20 white plaques were selected and single-stranded DNA prepared by polyethylene glycol precipitation. DNA sequencing was carried out by the dideoxy-chain terminating method (Sanger *et al*, 1977, 1980). α-^{32}P-dATP was incorporated into primed single-stranded DNA using Klenow polymerase, and the resulting sequencing ladders resolved on 0.5-2.5xTBE 6% polyacrylamide gradient gels (Biggin *et al*, 1983). Bands were visualised by overnight

autoradiography. Nucleotide sequences were read into a mainframe VAX-VMS computer using a Grafbar sonic digitiser in conjunction with a gel reading programme written and supplied by Dr A. Coulson of the Department of Molecular Biology, University of Edinburgh. Clones representing both orientations of the *HindIII* fragment were sequenced, each yielding approximately 150 bases of sequence information. Both sequences were compared to the Ad2 hexon gene which revealed a strong homology between one sequence and the region around 1kb of the Ad2 gene on the negative strand, confirming that the 5' end of the hexon gene had been cloned. Sequencing of the pCT3 HindIII insert was performed by the shotgun method. Briefly, 30μg pCT3 was sonicated for 30 sec using a microprobe. The fragmented DNA was treated with T4 DNA polymerase in the presence of dNTPs for 15 min at 37°C to produce blunt ends, and fractionated on a 1% agarose gel. Fragments between 300 and 800bp were excised, electro-eluted and precipitated. One-tenth of this was ligated into 100ng *SmaI*-cut dephosphorylated M13mp8 RF DNA. The reaction was transformed into JM101 cells and recombinant clones identified as white plaques in top agar supplemented with X-gal and IPTG. The resulting plaques were lifted onto nylon membranes, which were then treated to denaturation, neutralisation and fixing as for colony hybridization. Plaques were probed with the nick-translated 1.9kb *HindIII* insert isolated from pCT3. Single-stranded DNA was prepared from positively hybridising plaques (Sambrook *et al*, 1989) and sequenced. Information from gel readings was assembled into 1.2kb of contiguous sequence using the DB suite of Staden computer programmes (Staden,

1982a). Each base was sequenced an average of six times and included clones spanning both orientations. The 5' 700bp of the *HindIII* fragment was not covered by sufficient clones to provide accurate sequence and was therefore omitted from this work. In the same manner pCT6 was sequenced and assembled to within 100bp of the *EcoRI* cloning site.

DETERMINATION OF THE AD40 HEXON GENE SEQUENCE.

7. Identification and isolation of the Ad40 hexon gene.

Recombinant M13 clones containing fragments of the Ad41 hexon gene and generated in the sequencing described above, were used as probes to identify the Ad40 hexon gene. mA41.1 contained the region upstream from nucleotide 133 of the hexon gene on the negative strand and mA41.2, the region downstream from nucleotide 2602 on the positive strand. Single-stranded radioactive probes were generated from these clones by primer extension (Jeffreys *et al*, 1985) using the M13 universal primer, dCTP, dTTP, and dGTP in the presence of Klenow polymerase and α-^{32}P dATP. Following *BalI* digestion and alkali denaturation the DNA was fractionated on a 6% polyacrylamide gel containing 50% (w/v) urea. Single-stranded radioactive fragments were visualised by autoradiography, excised and electro-eluted. mA41.1 yielded a 129 nucleotide fragment spanning position 5 to 133 of the hexon gene, mA41.2 yielded a 185 nucleotide fragment extending from position 2602 to 8 nucleotides past the translational termination site. Ad40 genomic DNA was

digested (1µg per digest) with a series of restriction enzymes and fractionated on a 1% agarose gel. The resulting fragments were Southern blotted onto nylon in the manner described previously, and probed with mA41.1. After autoradiography, the probe was removed from the membrane by incubation in 0.4M NaOH for 30 min at 45°C, followed by a 15 min. wash in 0.1x SSC, 0.1% SDS, 0.2M Tris-HCl pH 7.5. The membrane was re-probed with A41.2. Both probes hybridised to a single 5.4kb fragment in *PstI* digested Ad40 DNA which was of sufficient size to contain the entire hexon gene. 1µg *PstI* digested Ad40 DNA was ligated into 100ng *PstI* cut, dephosphorylated pUC13 DNA and transformed into JM83. The hexon-containing recombinants were identified by colony hybridization, probing with mA41.1. Thus pCT40 was identified and isolated.

8. Sequencing the Ad40 hexon gene.

pCT40 was sequenced by the shotgun method described for Ad41. Following sonication, fragments were ligated into M13mp8 RF DNA, transfected into JM101 cells and plated out in top agar. Phage were transfered onto nylon membranes by the procedure used for colony lifts. The hexon inserts from pCT3 and pCT6 were isolated, nick-translated to provide a 'total hexon' probe and used to identify bacteriophage with Ad40 hexon inserts. 300 positive plaques were picked and sequenced providing most of the information needed for assembly of the entire gene. Extended gel electrophoresis (up to 8 hours) was performed on selected clones to provide further information. Only one area of approximately 220 bases was not covered by this method since extended

gel runs in excess of 8 hours proved difficult to read. This area corresponded to 1460-1680 of the sequence given in Results, figure 6. Primers were made (15 nucleotide) approximately 30 bases upstream and downstream of the unsequenced area and these were used in sequencing reactions in place of the M13 universal primer, to prime clones known to contain this region. By this method the sequence of the gene was completed.

DIAGNOSIS.

8. Synthesis and preparation of oligonucleotide probes.

Oligonucleotides were synthesised by the solid-phase phosphoramidite method on an Applied Biosystems automatic oligonucleotide synthesiser, according to the manufacturer's protocol. Oligonucleotides were cleaved from the solid-support column and deprotected with concentrated ammonia solution. Purification was performed following a standard procedure (Sambrook *et al*, 1989) using denaturing polyacrylamide gel electrophoresis, the oligonucleotide subsequently being visualised by u.v. shadowing on a fluorescent thin-layer chromatography plate. Full length oligonucleotide was passively eluted from excised gel slices, ethanol precipitated and redissolved in water, the concentration being determined from the absorbance at 260nm. The yield was generally 200-400μg per synthesis. To prepare a probe, 40ng oligonucleotide was 5' end-labelled using T4 polynucleotide kinase in the presence of 20μCi γ-^{32}P ATP. Unincorporated radioactivity was removed by

gel filtration through Biogel P-6 (Biorad). The specific activity of probes was in the region of 200-300 μCi/μg DNA.

9. Preparation of samples.

Clinical samples (10% faecal suspensions) were kindly supplied by Dr A.H. Kidd, (National Institute of Virology, Sandringham, South Africa) and Prof. U. Dessellberger (Regional Virus Laboratory, East Birmingham Hospital, Birmingham). Infected cell extracts were supplied by Barbara O'Donnell (MRC Virology Unit, Glasgow). All samples were treated with 100μg/ml proteinase K in the presence of 0.6% SDS for 1hr at 55°C, followed by phenol and ether extraction. 200μl sample was used per blot.

10. Dot blots.

Stool samples, infected cell extracts and plasmid DNA were denatured with 0.5M NaOH for 10 min., neutralised with 0.5M acetic acid, filtered onto nylon membranes using a vacuum filtration manifold and fixed with UV-irradiation. Membranes were treated with prehybridisation buffer for 1hr before overnight hybridisation with probe. Temperature for hybridisation was optimised at 45°C for oligonucleotides, followed by two 15min washes in 0.5xSSC, 0.25% SDS. For the Ad2 420bp probe, preparation, hybridisation and washing were as described for Southern blotting in sections 3 and 4. After washing under appropriate conditions for the probe used, membranes were subjected to overnight autoradiography.

RESULTS.

SECTION 1. Identification, cloning and sequencing of the type 40 and 41 hexon genes and predicted structrure of the proteins.

1.1 Isolation and sequence of the adenovirus type 41 hexon gene.

The Ad41 hexon gene was identified using a probe derived from the 5' end of the Ad2 gene, a region which has extensive homology with other human and animal adenoviruses (Kinloch *et al*,1984; Hu *et al*, 1984). Ad41 genomic DNA was cleaved with a series of restriction enzymes in order to generate a hexon-containing fragment of a suitable size for initial cloning and sequencing (1-2kb). The resulting fragments were resolved by agarose gel electrophoresis, Southern blotted and probed with the Ad2 hexon fragment (figure 1a and b). A single hybridising species of approximately 1.9kb was identified in *HindIII* digested Ad41 DNA which was ligated into *HindIII* cut pUC13 to yield the recombinant plasmid pCT3. Sequencing revealed that this 1.9kb insert contained the 5' third of the hexon gene. A 120bp *BglII-HindIII* fragment generated from the 3' end of the pCT3 insert was used as a probe to identify an overlapping fragment of sufficient size to contain the remainder of the Ad41hexon gene. Ad41 genomic DNA was double-digested with *BglII* and a series of second restriction enzymes, and the resulting fragments resolved on an agarose gel. This was Southern blotted and probed with the 120bp pCT3 *BglII-HindIII* fragment (figure 2). A single hybridising species of 2.7kb was identified in *BglII-EcoRI* digested Ad41 which was ligated into *BamHI-EcoRI* cut pUC13 to yield the plasmid pCT6. The location of the hexon gene with respect to these fragments and within the genome as a whole is shown in figure 3. Plasmids pCT3 and pCT6 were sonicated to generate random fragments which were ligated into the bacteriophage M13mp8. Hexon-containing clones

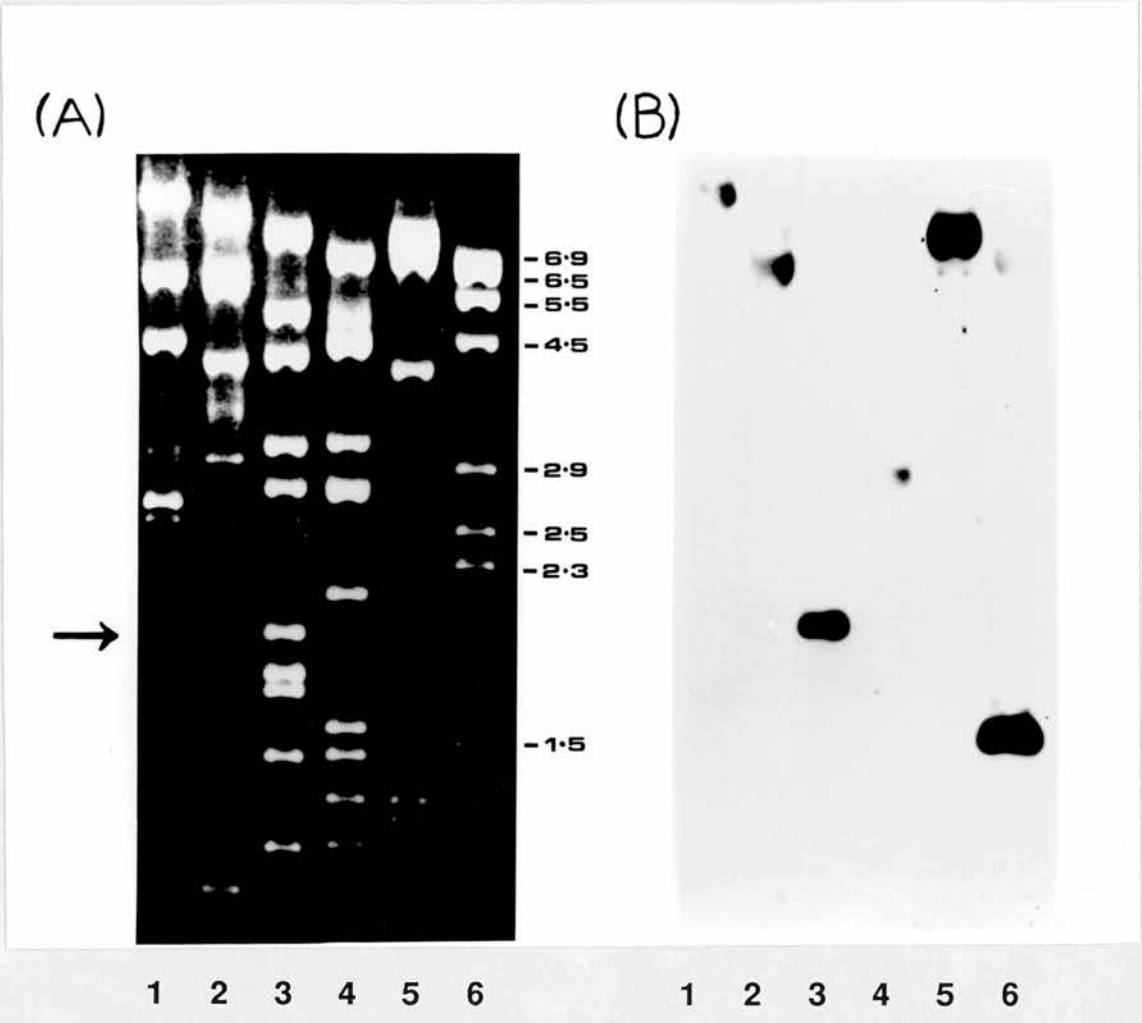


Figure 1. (A) agarose gel electrophoresis of adenovirus type 41 DNA cleaved with restriction enzymes: 1. *EcoRI*, 2. *HpaI*, 3, *HindIII*, 4. *SmaI* and 5. *BamHI*. Size markers, lane 6, were provided by adenovirus type 2 DNA cleaved with *SmaI* and are indicated in kb. (B) the gel was Southern blotted and probed with a nick-translated 420bp fragment isolated from *SmaI-XhoI* digested pUB54. The arrow indicates the 1.9kb *HindIII* fragment which was cloned into pUC13 to yield pCT3.

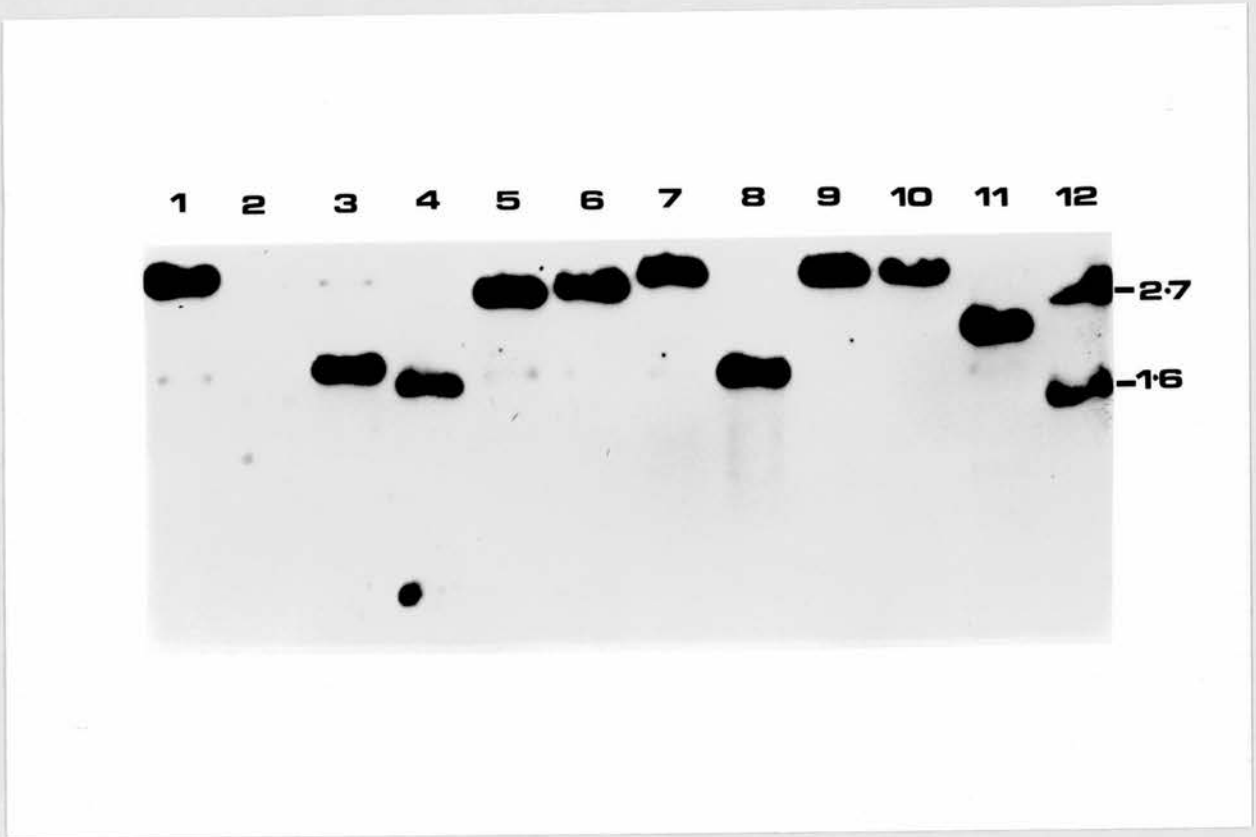


Figure 2. Southern blot of a 0.7% agarose on which a series of restriction enzyme digests of adenovirus type 41 DNA were run. DNA (1µg) was cleaved with *BglII* (1), or *BglII* and a second enzyme: *HindIII* (2), *SmaI* (3), *AccI* (4), *EcoRI* (5), *PvuII* (6), *DpnI* (7), *BamHI* (8), *XbaI* (9), *XhoII* (10) or *SphI* (11). Size-markers of 2.7 and 1.6kb were run in lane 12. The blot was probed with a nick-translated 120bp *BglII-HindIII* fragment from the 3' end of the hexon insert isolated from pCT3.

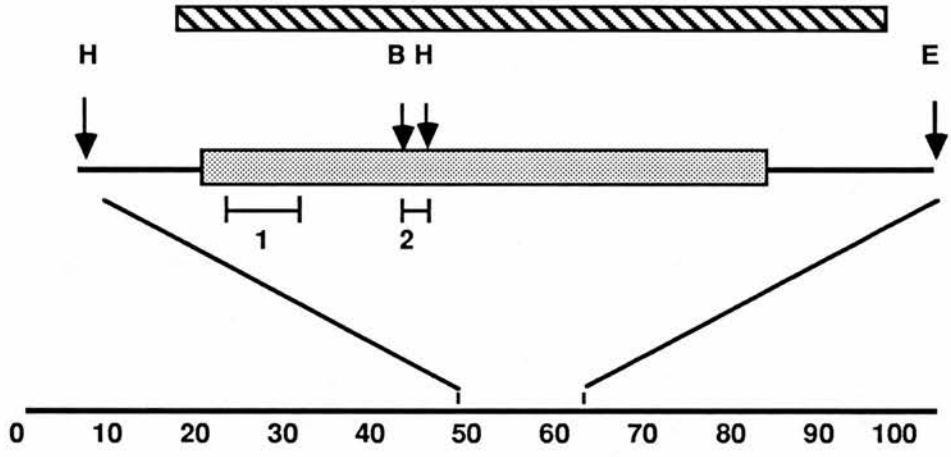


Figure 3. Genomic location of the adenovirus type 41 hexon gene. The physical map of type 41 (Takiff *et al*, 1984) was used to locate hexon containing restriction enzyme fragments on the viral genome (divided into 100 map units representing approximately 36kbp). The stippled area represents the hexon gene and the shaded bar indicates the region that was sequenced. Arrows show the sites of cleavage for enzymes *HindIII* (H), *BglII* (B) and *EcoR1* (E). Bracketed areas 1 and 2 indicate the regions corresponding to the probes used in the identification of plasmids pCT3 and pCT6 respectively.

were identified, and the nucleotide sequence of the gene determined by shotgun dideoxy-sequencing.

The assembled nucleotide sequence and translated major open reading frame are displayed in figure 4. The shotgun procedure did not yield sufficient M13 clones to enable accurate sequencing of the 5' end of the pCT3 insert or of the extreme 3' end of the pCT6 insert, and these regions are therefore omitted from figure 4. The hexon gene is represented by an open reading frame of 2778 nucleotides which would predict a polypeptide of 924 amino acids. This is smaller than the 967 amino acids of the Ad2 hexon (Akusjärvi *et al*, 1984), an observation which is supported by increased relative mobility on SDS-polyacrylamide gels (data not shown).

Figure 4. Sequence of the adenovirus type 41 hexon gene and flanking regions. Translation of the hexon gene in single letter amino acid code is shown above the middle base of each codon.

```
   1 GCGGCCCCAACCGTTGCCGTGGCCACTCCGGCTCGCAGAGTTCGCGGTGCCAATTGGCAG

  61 AGCACTTTGAACAGCATTGTGGGTTTGGGAGTGAAAAGCCTAAAACGCCGCCGGTGCTAT

 121 TAAAGTGGAACCAAGCTAAATACCCTTATTGTATGCGCCTCCTGTGTTACGCCAGACAGC
                                          M  A  T  P  S  M  M  P
 181 CGAGTGACGGGTCACCGCCAAGAGCGCCGCTTGCAAGATGGCCACCCCCTCGATGATGCC
       Q  W  S  Y  M  H  I  A  G  Q  D  A  S  E  Y  L  S  P  G  L
 241 GCAATGGTCTTACATGCACATCGCCGGGCAGGACGCCTCGGAGTATCTGAGCCCGGGCCT
       V  Q  F  A  R  A  T  D  T  Y  F  S  L  G  N  K  F  R  N  P
 301 GGTGCAATTTGCCCGCGCCACCGATACGTACTTCAGCCTGGGGAACAAGTTCAGAAATCC
       T  V  A  P  T  H  D  V  T  T  D  R  S  Q  R  L  T  L  R  F
 361 CACTGTGGCTCCGACCCACGATGTAACCACAGACAGGTCACAGCGACTGACGCTGCGATT
```

```
      V  P  V  D  R  E  D  T  A  Y  S  Y  K  V  R  F  T  L  A  V
 421 CGTGCCAGTCGACCGCGAGGACACCGCTTATTCTTACAAAGTGCGCTTTACGCTGGCCGT
      G  D  N  R  V  L  D  M  A  S  T  Y  F  D  I  R  G  V  L  D
 481 GGGCGACAACCGGGTGTTGGACATGGCCAGCACCTACTTTGACATCCGCGGCGTGCTGGA
      R  G  P  S  F  K  P  Y  S  G  T  A  Y  N  S  L  A  P  K  T
 541 TCGTGGCCCCAGCTTTAAACCCTACTCCGGAACCGCCTATAACTCCCTGGCTCCTAAAAC
      A  P  N  P  C  E  W  K  D  N  N  K  I  K  V  R  G  Q  A  P
 601 AGCACCCAATCCATGCGAATGGAAGGACAACAACAAAATAAAAGTTAGAGGGCAAGCTCC
      F  I  G  T  N  I  N  K  D  N  G  I  Q  I  G  T  D  T  T  N
 661 CTTTATAGGAACAAACATAAATAAAGACAATGGGATTCAGATAGGAACAGATACAACAAA
      Q  P  I  Y  A  D  K  T  Y  Q  P  E  P  Q  V  G  Q  T  Q  W
 721 TCAGCCTATTTATGCCGACAAAACCTATCAACCAGAGCCTCAAGTGGGTCAAACACAGTG
      N  S  E  V  G  A  A  Q  K  V  A  G  R  V  L  K  D  T  T  P
 781 GAACAGTGAAGTTGGCGCCGCTCAAAAGGTAGCGGGTAGGGTGCTGAAAGATACCACGCC
      M  L  P  C  Y  G  S  Y  A  K  P  T  N  E  K  G  G  Q  A  S
 841 AATGTTACCATGCTACGGGTCTTACGCCAAACCAACAAATGAAAAGGGCGGTCAGGCTAG
      L  I  T  N  G  T  D  Q  T  L  T  S  D  V  N  L  Q  F  F  A
 901 TCTCATAACGAACGGAACCGATCAAACTCTAACAAGCGACGTTAACTTACAGTTCTTTGC
      L  P  S  T  P  N  E  P  K  A  V  L  Y  A  E  N  V  S  I  E
 961 CTTGCCAAGCACTCCTAATGAGCCAAAAGCCGTTCTTTACGCTGAAAACGTGTCCATTGA
      A  P  D  T  H  L  V  Y  K  P  D  V  A  Q  G  T  I  S  S  A
1021 AGCGCCCGACACCCATTTGGTGTACAAGCCAGATGTAGCTCAGGGAACCATAAGTTCGGC
      D  L  L  T  Q  Q  A  A  P  N  R  P  N  Y  I  G  F  R  D  N
1081 AGATCTTTTAACGCAGCAGGCAGCGCCCAACAGACCGAACTACATTGGCTTTAGGGATAA
      F  I  G  L  M  Y  Y  N  S  T  G  N  M  G  V  L  A  G  Q  A
1141 CTTTATCGGCCTGATGTACTACAACTCCACAGGCAATATGGGTGTATTGGCTGGGCAAGC
      S  Q  L  N  A  V  V  D  L  Q  D  R  N  T  E  L  S  Y  Q  L
1201 TTCACAGCTAAATGCTGTAGTGGACTTGCAAGACAGGAACACTGAGTTATCATACCAACT
      M  L  D  A  L  G  D  R  S  R  Y  F  S  M  W  N  Q  A  V  D
1261 TATGCTGGACGCACTTGGCGATCGGAGCAGATATTTTTCTATGTGGAATCAGGCTGTTGA
      S  Y  D  P  D  V  R  I  I  E  N  H  G  V  E  D  E  L  P  N
1321 CAGTTACGACCCCGACGTAAGGATCATTGAGAACCACGGAGTGGAGGACGAACTGCCAAA
      Y  C  F  P  L  G  G  S  A  A  T  D  T  Y  S  G  I  K  A  N
1381 TTACTGCTTTCCGCTGGGAGGGTCTGCAGCTACAGACACGTACTCTGGCATAAAGGCCAA
      G  Q  T  W  T  A  D  D  N  Y  A  D  R  G  A  E  I  E  S  G
1441 TGGCCAAACCTGGACTGCAGACGACAATTATGCCGACAGAGGGGCAGAAATTGAATCTGG
      N  I  F  A  M  E  I  N  L  A  A  N  L  W  R  S  F  L  Y  S
1501 GAACATTTTTGCCATGGAAATCAATTTGGCGGCCAATCTCTGGCGCAGCTTCTTATACTC
      N  V  A  L  Y  L  P  D  S  Y  K  I  T  P  D  N  I  T  L  P
1561 CAATGTAGCTTTGTACTTGCCTGACTCATACAAGATTACGCCAGACAACATTACACTGCC
      E  N  K  N  T  Y  A  Y  M  N  G  R  V  A  V  P  S  A  L  D
1621 CGAAAACAAGAACACCTATGCCTACATGAACGGTCGGGTGGCGGTTCCTAGCGCCCTCGA
      T  Y  V  N  I  G  A  R  W  S  P  D  P  M  D  N  V  N  P  F
1681 TACCTACGTAAACATCGGGGCACGGTGGTCTCCAGATCCCATGGACAATGTTAACCCCTT
      N  H  H  R  N  A  G  L  R  Y  R  S  M  L  L  G  N  G  R  Y
1741 CAATCACCACCGTAACGCCGGTCTGCGCTATCGATCCATGCTCTTGGGCAACGGGCGTTA
      V  P  F  H  I  Q  V  P  Q  K  F  F  A  I  K  N  L  L  L  L
1801 CGTACCCTTCCACATTCAAGTCCCCCAGAAGTTTTTTGCCATTAAAAATCTCCTCCTCTT
      P  G  S  Y  T  Y  E  W  N  F  R  K  D  V  N  M  I  L  Q  S
1861 ACCGGGTTCCTACACCTACGAGTGGAACTTCAGGAAGGACGTTAACATGATCCTCCAGAG
      S  L  G  N  D  L  R  V  D  G  A  S  V  R  F  D  S  I  N  L
1921 CAGTCTGGGTAACGACCTGCGGGTCGACGGAGCCAGCGTCAGGTTCGATAGCATTAATCT
      Y  A  N  F  F  P  M  A  H  N  T  A  S  T  L  E  A  M  L  R
1981 GTACGCCAACTTTTTTCCCATGGCTCACAACACCGCCTCCACCTTGGAAGCTATGCTGCG
      N  D  T  N  D  Q  S  F  N  D  Y  L  C  A  A  N  M  L  Y  P
2041 CAATGACACCAACGACCAATCGTTTAACGACTACCTCTGCGCTGCAAACATGCTTTATCC
      I  P  S  N  A  T  S  V  P  I  S  I  P  S  R  N  W  A  A  F
2101 CATTCCTTCTAATGCTACTAGCGTTCCAATTTCTATTCCTTCGCGGAACTGGGCCGCTTT
      R  G  W  S  F  T  R  L  K  T  K  E  T  P  S  L  G  S  G  F
2161 TCGGGGCTGGAGTTTTACTCGACTGAAAACCAAAGAAACCCCCTCTTTGGGTTCGGGATT
      D  P  Y  F  T  Y  S  G  S  V  P  Y  L  D  G  T  F  Y  L  N
2221 TGATCCCTACTTCACCTACTCTGGCTCCGTTCCATACTTGGATGGTACCTTTTACCTCAA
```

```
      H  T  F  K  K  V  S  I  M  F  D  S  S  V  S  W  P  G  N  D
2281 CCACACTTTTAAAAAGGTGTCCATCATGTTTGACTCCTCTGTGAGTTGGCCTGGCAATGA
      R  L  L  T  P  N  E  F  E  I  K  R  T  V  D  G  E  G  Y  N
2341 CCGACTGCTTACCCCCAATGAGTTTGAAATCAAACGCACCGTGGATGGGGAAGGGTACAA
      V  A  Q  C  N  M  T  K  D  W  F  L  I  Q  M  L  S  H  Y  N
2401 CGTGGCTCAATGTAACATGACCAAAGACTGGTTTCTGATTCAGATGCTTAGTCATTACAA
      I  G  Y  Q  G  F  Y  V  P  E  S  Y  K  D  R  M  Y  S  F  F
2461 CATTGGATATCAGGGTTTTTATGTACCCGAAAGCTACAAAGATAGAATGTATTCCTTTTT
      R  N  F  Q  P  M  S  R  Q  V  V  N  T  T  T  Y  K  E  Y  Q
2521 CCGAAACTTCCAGCCAATGAGTCGTCAGGTGGTAAACACTACCACCTACAAGGAATATCA
      N  V  T  L  P  F  Q  H  N  N  S  G  F  V  G  Y  M  G  P  T
2581 GAATGTTACCCTTCCTTTCCAGCATAATAACTCAGGCTTTGTGGGATACATGGGACCCAC
      M  R  E  G  Q  A  Y  P  A  N  Y  P  Y  P  L  I  G  Q  T  A
2641 CATGCGGGAGGGACAAGCTTACCCCGCTAACTACCCTTACCCCCTTATTGGTCAAACGGC
      V  P  S  L  T  Q  K  K  F  L  C  D  R  T  M  W  R  I  P  F
2701 CGTGCCAAGCCTGACACAGAAAAAATTCCTGTGCGATCGCACCATGTGGCGCATTCCGTT
      S  S  N  F  M  S  M  G  A  L  T  D  L  G  Q  N  M  L  Y  A
2761 TTCCAGCAACTTCATGTCTATGGGGGCACTGACCGACCTGGGGCAGAACATGCTGTACGC
      N  S  A  H  A  L  D  M  T  F  E  V  D  P  M  D  E  P  T  L
2821 CAATTCCGCCCACGCCCTCGACATGACTTTTGAGGTCGATCCCATGGATGAGCCCACACT
      L  Y  V  L  F  E  V  F  D  V  V  R  I  H  Q  P  H  R  G  V
2881 TCTTTATGTTTTGTTTGAAGTTTTCGACGTCGTGCGCATCCACCAGCCCCACCGCGGCGT
      I  E  A  V  Y  L  R  T  P  F  S  A  G  N  A  T  T  *
2941 CATCGAGGCCGTCTACCTCCGTACGCCGTTCTCGGCCGGTAACGCCACCACATAAGAAGC

3001 TGGCCAATGGGCTCCAGCGAGCAGGAGCTGGTCGCCATCGCGCGTGACCTGGGCTGCGGA

3061 TCCTACTTTCTAGGCACCTTTGACAAGCGTTTCCCGGGCTTTATGGCTCCGAACAAGCTG

3121 GCCTGCGCTATTGTGAACACGGCCGGGCGTGAAACAGGGGGCGTTCACTGGTTGGCCCTG

3181 GCCTGGAACCCCAAGAGCCATACCTGTTACCTCTTTGACCCGTTTGGCTTTTCAGACGAG

3241 CGGCTCAAACAAATTTACCAGTTTGAATATGAGGGCCTACTAAAGCGAAGCGCTTTGGCC

3301 TCCACGCCTGACCACTGCATTACCCTTGTGAAATCCACCCAGACCGTTCAGGGGCCTTTT

3361 TCTGCCGCCTGTGGTCTTTTCTGCTGCATGTTTTTGCATGCTTTTATACACTGGCCCAGT

3421 AATCCAATGGAGCAAAATCCCACAATGGATCTTCTCACGGGCGTACCGAACAGCATGCTT

3481 CAAAGCCCCCAGGTTGAACCCACCCTGCGTCGGAATCAGGAACGATTGTATCGTTTTCTG

3541 ACCCAACACTCTCCCTACTTTCGTCGTCACCGCGAGCGAATTGAAAAAGCTACAGCTTTT
```

1.2. Isolation and sequence of the adenovirus type 40 hexon gene.

The Ad40 hexon gene was identified by hybridisation to homologous regions of the Ad41 gene, the sequence of which had been determined. Two recombinant M13 phage generated during the Ad41 shotgun procedure

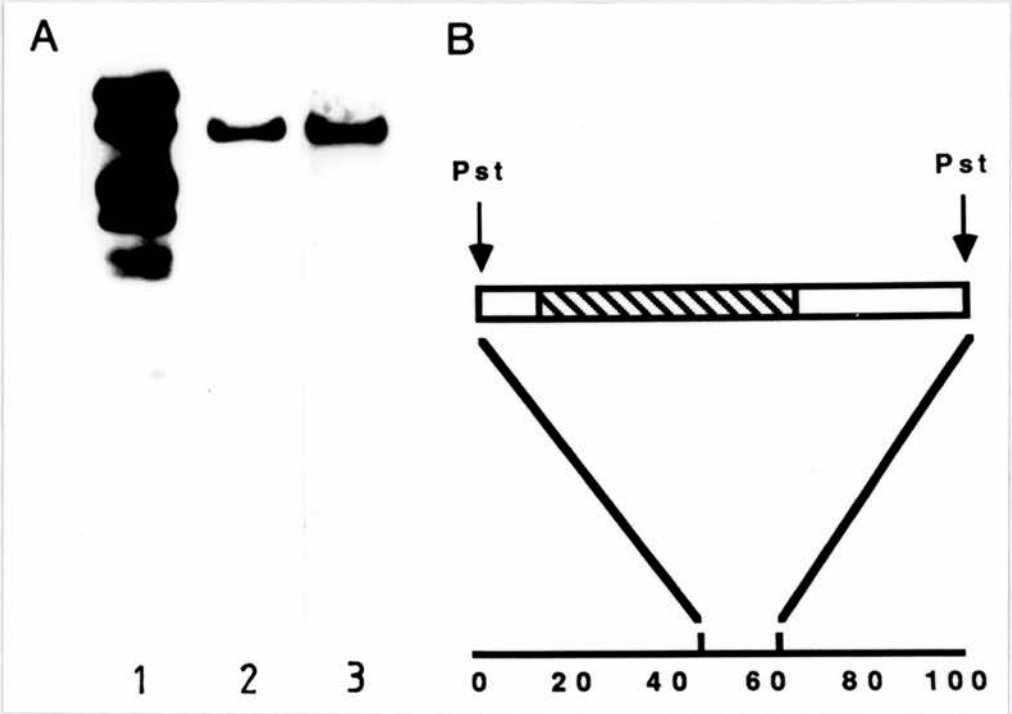


Figure 5. Identification of the Ad40 hexon gene by hybridisation to homologous regions of Ad41 DNA. Genomic DNA was digested with restriction enzymes, fractionated by 0.7% agarose gel electrophoresis, transferred to a nylon membrane and probed. Only the blot of the *Pst1* digest is shown in this figure, the other digests producing multiple hybridising species or fragments which were too large. (A) Lane 1, Ad2 DNA cleaved with *HindIII* and probed with nick-translated Ad2 DNA to provide size-markers; lane2, Ad40 DNA cleaved with *PstI* and probed with mA41.1 representing the 5' end of the hexon gene; lane 3, hybridised probe was removed from the membrane and re-probed with mA41.2 representing the 3' end of the hexon gene. (B) Position of the Ad40 hexon gene (shaded region) within the cloned *PstI* fragment and its location within the genome as a whole (represented by 100 map units, equivalent to approximately 36kbp). The location within the genome was determined by alignment of restriction sites in the hexon gene with the published restriction maps (Takiff *et al*, 1984).

contained the 5' and the 3' ends of the hexon gene, (mAd41.1 and mAd41.2 respectively). These were used to generate radioactive, single-stranded probes for identifying hexon fragments in a southern blot of Ad40 genomic DNA that had been cleaved with a series of restriction enzymes. The blot was probed with mAd41.1, stripped and reprobed with mAd41.2. *PstI* digestion yielded a single hybridising species of sufficient size to contain the entire hexon gene in one fragment (figure 5a). This *PstI* fragment was ligated into pUC13 to generate the plasmid pCT40 which was thus used as a source of DNA for sonication followed by shotgun sequencing to determine the nucleotide sequence of the hexon gene. The location of the hexon gene within the *PstI* fragment and within the genome as a whole is given in figure 5b.

The complete nucleotide sequence and corresponding translation of the Ad40 hexon gene is shown in figure 6. The 2772 nucleotide open reading frame predicts a polypeptide of 922 amino acids which is two amino acids shorter than the corresponding Ad41 polypeptide and forty-five amino acids shorter than that of Ad2.

Figure 6. Nucleic acid sequence of the Ad40 hexon gene and predicted amino acid sequence of the corresponding polypeptide. Translation in single letter code is shown above the middle base of each codon.

```
    M  A  T  P  S  M  M  P  Q  W  S  Y  M  H  I  A  G  Q  D  A
1 ATGGCCACCCCCTCGATGATGCCGCAATGGTCTTACATGCACATCGCCGGGCAGGACGCC
```

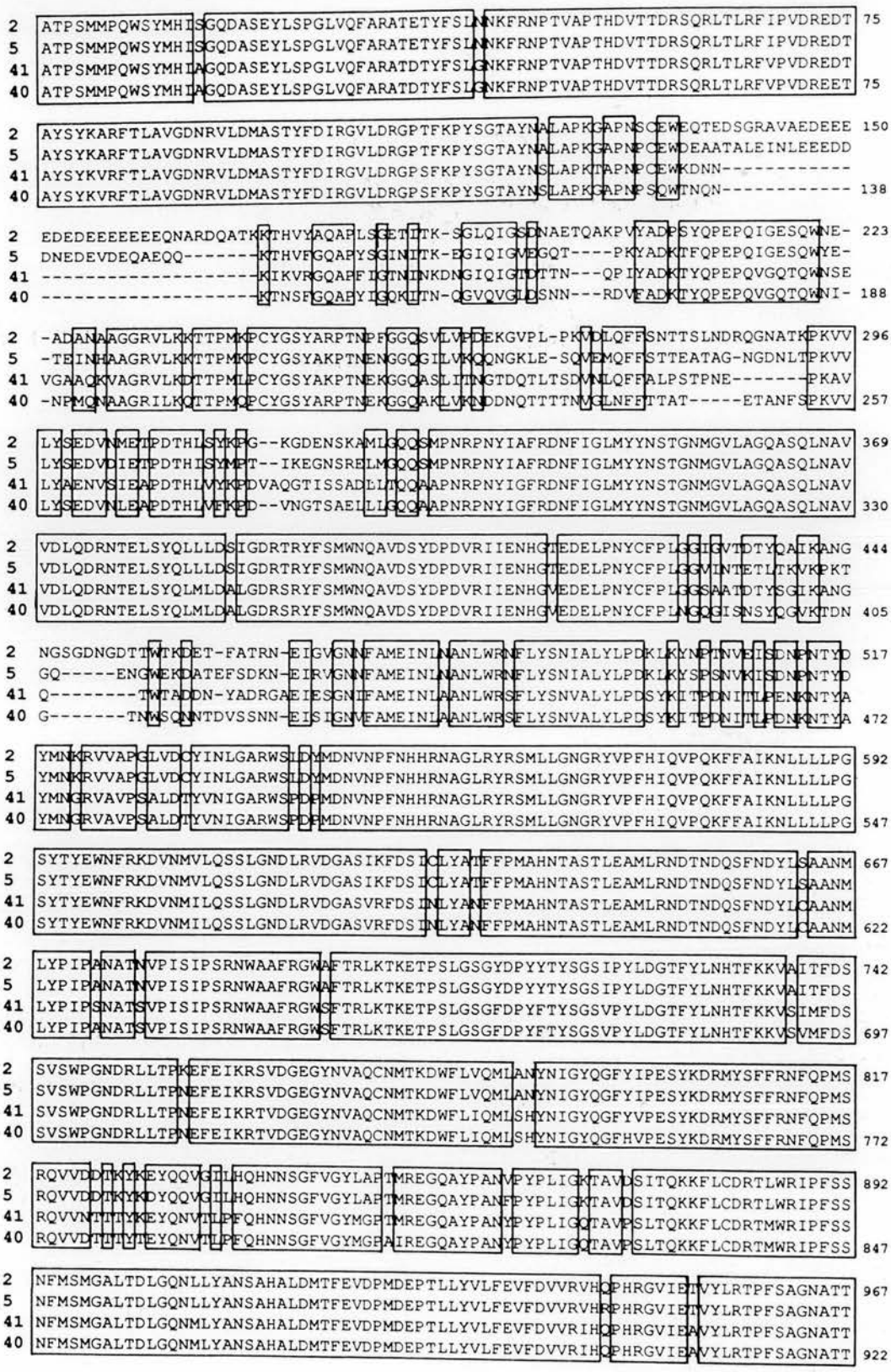
```
     S  E  Y  L  S  P  G  L  V  Q  F  A  R  A  T  D  T  Y  F  S
  61 TCGGAGTACCTGAGCCCGGGCCTGGTGCAGTTCGCCCGTGCCACCGATACCTACTTCAGC
     L  G  N  K  F  R  N  P  T  V  A  P  T  H  D  V  T  T  D  R
 121 CTGGGGAACAAGTTCAGAAACCCCACCGTGGCTCCCACCCACGATGTAACCACAGACAGG
     S  Q  R  L  T  L  R  F  V  P  V  D  R  E  E  T  A  Y  S  Y
 181 TCGCAGCGACTGACGCTGCGCTTCGTGCCCGTCGACCGCGAGGAAACCGCCTACTCTTAC
     K  V  R  F  T  L  A  V  G  D  N  R  V  L  D  M  A  S  T  Y
 241 AAAGTGCGCTTTACGCTGGCCGTGGGCGACAACCGGGTTTTGGACATGGCCAGCACCTAC
     F  D  I  R  G  V  L  D  R  G  P  S  F  K  P  Y  S  G  T  A
 301 TTTGACATCCGCGGCGTGCTGGATCGTGGTCCCAGCTTTAAACCCTATTCGGGCACTGCA
     Y  N  S  L  A  P  K  G  A  P  N  P  S  Q  W  T  N  Q  N  K
 361 TACAACTCCCTGGCCCCCAAAGGTGCTCCCAATCCTAGCCAGTGGACAAACCAAAACAAA
     T  N  S  F  G  Q  A  P  Y  I  G  Q  K  I  T  N  Q  G  V  Q
 421 ACAAACTCCTTTGGACAAGCTCCCTATATAGGACAAAAAATCACCAATCAGGGCGTGCAA
     V  G  L  D  S  N  N  R  D  V  F  A  D  K  T  Y  Q  P  E  P
 481 GTGGGCTTAGACTCCAACAATCGCGATGTGTTTGCCGATAAAACGTACCAACCGGAGCCT
     Q  V  G  Q  T  Q  W  N  I  N  P  M  Q  N  A  A  G  R  I  L
 541 CAAGTGGGGCAGACGCAATGGAACATTAATCCAATGCAAAACGCTGCGGGAAGAATACTA
     K  Q  T  T  P  M  Q  P  C  Y  G  S  Y  A  R  P  T  N  E  K
 601 AAACAAACCACGCCCATGCAGCCATGTTATGGGTCATACGCTAGACCAACAAACGAAAAA
     G  G  Q  A  K  L  V  K  N  D  D  N  Q  T  T  T  T  N  V  G
 661 GGAGGTCAAGCCAAGCTGGTAAAAAATGACGACAATCAGACCACAACAACAAACGTAGGT
     L  N  F  F  T  T  T  A  T  E  T  A  N  F  S  P  K  V  V  L  Y
 721 TTAAACTTTTTTACCACTGCCACTGAGACCGCTAATTTTTCACCAAAGGTGGTTCTGTAC
     S  E  D  V  N  L  E  A  P  D  T  H  L  V  F  K  P  D  V  N
 781 AGCGAGGATGTTAACTTAGAAGCGCCCGATACCCACCTTGTGTTTAAGCCAGATGTCAAC
     G  T  S  A  E  L  L  L  G  Q  Q  A  A  P  N  R  P  N  Y  I
 841 GGCACAAGTGCCGAGCTTTTACTGGGTCAGCAGGCCGCTCCCAATCGACCTAATTACATT
     G  F  R  D  N  F  I  G  L  M  Y  Y  N  S  T  G  N  M  G  V
 901 GGTTTTAGGGACAACTTCATTGGTTTGATGTACTACAATTCCACTGGCAACATGGGAGTG
     L  A  G  Q  A  S  Q  L  N  A  V  V  D  L  Q  D  R  N  T  E
 961 CTGGCCGGGCAAGCTTCTCAGCTCAACGCAGTGGTGGACTTACAAGATAGAAACACGGAG
     L  S  Y  Q  L  M  L  D  A  L  G  D  R  S  R  Y  F  S  M  W
1021 CTGTCTTACCAGTTAATGCTTGACGCTTTAGGGGATCGGAGTCGATACTTCTCCATGTGG
     N  Q  A  V  D  S  Y  D  P  D  V  R  I  I  E  N  H  G  V  E
1081 AACCAGGCAGTGGACAGCTATGACCCAGACGTGAGAATTATTGAAAATCATGGCGTGGAA
     D  E  L  P  N  Y  C  F  P  L  N  G  Q  G  I  S  N  S  Y  Q
1141 GACGAGCTCCCCAACTATTGCTTTCCTCTTAATGGGCAAGGAATATCTAACAGTTACCAA
     G  V  K  T  D  N  G  T  N  W  S  Q  N  N  T  D  V  S  S  N
1201 GGCGTGAAGACTGACAATGGAACTAACTGGTCTCAGAATAATACAGACGTCTCAAGCAAC
     N  E  I  S  I  G  N  V  F  A  M  E  I  N  L  A  A  N  L  W
1261 AATGAAATTTCCATTGGCAATGTGTTTGCCATGGAGATTAATCTGGCGGCTAACTTGTGG
     R  S  F  L  Y  S  N  V  A  L  Y  L  P  D  S  Y  K  I  T  P
1321 AGAAGCTTCTTGTACTCAAATGTAGCCTTGTACTTGCCTGACTCTTACAAAATAACCCCC
     D  N  I  T  L  P  D  N  K  N  T  Y  A  Y  M  N  G  R  V  A
1381 GATAACATTACTTTACCCGACAACAAAAATACATATGCCTACATGAACGGTCGGGTTGCC
     V  P  S  A  L  D  T  Y  V  N  I  G  A  R  W  S  P  D  P  M
1441 GTCCCCAGCGCCCTGGATACATACGTAAACATCGGGGCGCGGTGGTCTCCAGACCCCATG
     D  N  V  N  P  F  N  H  H  R  N  A  L  G  R  Y  R  S  M  L
1501 GACAACGTTAATCCCTTTAACCACCACCGCAATGCTGGTCTGCGCTACCGTTCTATGCTC
     L  G  N  G  R  Y  V  P  F  H  I  Q  V  P  Q  K  F  F  A  I
1561 CTGGGTAACGGCCGCTACGTGCCTTTTCACATCCAAGTGCCCCAGAAATTTTTCGCCATT
     K  N  L  L  L  L  P  G  S  Y  T  Y  E  W  N  F  R  K  D  V
1621 AAAAATCTCCTGCTCCTGCCCGGGTCCTACACCTATGAGTGGAACTTCCGGAAGGATGTT
     N  M  I  L  Q  S  S  L  G  N  D  L  R  V  D  G  A  S  V  R
1681 AACATGATTCTCCAAAGCAGTCTCGGTAACGACCTCAGGGTCGATGGAGCCAGCGTCAGG
     F  D  S  I  N  L  Y  A  N  F  F  P  M  A  H  N  T  A  S  T
1741 TTTGACAGCATTAACCTGTATGCCAACTTTTTCCCCATGGCTCACAACACCGCTTCCACC
     L  E  A  M  L  R  N  D  T  N  D  Q  S  F  N  D  Y  L  C  A
1801 TTGGAAGCAATGCTTCGTAATGATACCAACGATCAGTCTTTCAACGACTACCTCTGCGCC
     A  N  M  L  Y  P  I  P  A  N  A  T  S  V  P  I  S  I  P  S
1861 GCAAACATGCTTTACCCCATACCCGCCAACGCTACTAGCGTGCCCATTTCTATTCCTTCG
```

```
      R  N  W  A  A  F  R  G  W  S  F  T  R  L  K  T  K  E  T  P
1921 CGAAATTGGGCTGCTTTTCGGGGGTGGAGTTTTACTAGACTAAAAACTAAAGAAACCCCC
      S  L  G  S  G  F  D  P  Y  F  T  Y  S  G  S  V  P  Y  L  D
1981 TCTTTGGGGTCCGGGTTTGATCCATATTTCACCTACTCTGGCTCCGTCCCATACTTGGAT
      G  T  F  Y  L  N  H  T  F  K  K  V  S  V  M  F  D  S  S  V
2041 GGCACCTTTTACCTGAACCATACTTTTAAAAAGGTGTCCGTTATGTTCGACTCCTCTGTG
      S  W  P  G  N  D  R  L  L  T  P  N  E  F  E  I  K  R  T  V
2101 AGCTGGCCTGGTAACGACCGACTACTTACTCCCAACGAGTTTGAAATCAAACGCACCGTG
      D  G  E  G  Y  N  V  A  Q  C  N  M  T  K  D  W  F  L  I  Q
2161 GATGGGGAAGGATACAACGTGGCTCAATGTAACATGACCAAGGACTGGTTCCTCATACAA
      M  L  S  H  Y  N  I  G  Y  Q  G  F  H  V  P  E  S  Y  K  D
2221 ATGCTCAGTCACTACAATATTGGCTACCAGGGTTTCCACGTACCAGAAAGCTACAAGGAC
      R  M  Y  S  F  F  R  N  F  Q  P  M  S  R  Q  V  V  D  T  T
2281 AGGATGTACTCCTTTTTCCGAAACTTCCAACCCATGAGCCGCCAGGTGGTAGACACTACC
      T  Y  T  E  Y  Q  N  V  T  L  P  F  Q  H  N  N  S  G  F  V
2341 ACCTACACGGAGTATCAGAATGTAACTCTCCCTTTCCAGCATAATAACTCTGGCTTTGTA
      G  Y  M  G  P  A  I  R  E  G  Q  A  Y  P  A  N  Y  P  Y  P
2401 GGATACATGGGACCTGCCATACGGGAGGGACAAGCTTACCCCGCCAACTATCCATACCCC
      L  I  G  Q  T  A  V  P  S  L  T  Q  K  K  F  L  C  D  R  T
2461 CTTATTGGTCAGACGGCCGTACCAAGCCTGACTCAGAAAAAATTTCTTTGCGATCGTACC
      M  W  R  I  P  F  S  S  N  F  M  S  M  G  A  L  T  D  L  G
2521 ATGTGGCGCATTCCCTTTTCCAGCAACTTTATGTCTATGGGGGCCCTGACTGACCTGGGG
      Q  N  M  L  Y  A  N  S  A  H  A  L  D  M  T  F  E  V  D  P
2581 CAAAACATGCTGTACGCCAACTCCGCCCACGCGCTCGACATGACTTTTGAGGTGGACCCC
      M  D  E  P  T  L  L  Y  V  L  F  E  V  F  D  V  V  R  I  H
2641 ATGGATGAGCCCACACTTCTCTATGTTCTGTTCGAAGTTTTCGACGTTGTGCGCATCCAC
      Q  P  H  R  G  V  I  E  A  V  Y  L  R  T  P  F  S  A  G  N
2701 CAGCCGCACCGCGGCGTCATCGAGGCCGTCTACCTGCGTACGCCGTTCTCGGCCGGTAAC
      A  T  T  *
2761 GCCACCACATAA
```

1.3. Comparison of the enteric adenovirus hexons with other known hexon sequences.

To examine the similarities between hexons, the four human adenovirus hexons whose sequences have now been determined were aligned (figure 7). Padding characters have been included to facilitate maximal alignment. In the following comparisons each padding character is taken to be equivalent to one mismatch. When the Ad41 hexon polypeptide is compared to its Ad2 counterpart an overall identity of 77% is observed, with more than one-third of

Figure 7. Sequence alignment of the hexons from adenovirus types 2, 5, 40 and 41. Residues conserved in all four serotypes are boxed. Amino acids are considered conserved on the basis of the following criteria: exchanges between hydrophobic residues (Ala, Gly, Leu, Ile, Val, Cys and Met); exchanges between aromatic residues (Phe, Tyr, Trp and His); or exchanges between hydrophilic residues of a similar size (Glu, Asp, Gln, and Asn; Arg for Lys; Thr for Ser). All changes involving Pro are treated as significant. Amino acid sequence numbering is shown for both Ad2 and Ad40.



the non-identical amino acids representing conservative substitutions. The changes are not evenly distributed throughout the polypeptides but appear to be localised in discrete areas. This creates regions of comparatively low homology, which are interspersed with regions of much higher than average homology. For example the amino acids comprising regions 1-134 and 491-924 show greater than 90% identity, while in regions 135-296 and 409-490 this drops to 49% and 58% respectively, (the numbering corresponds to the Ad41 amino acid sequence). The regional pattern of distribution is very similar to that observed in comparison made between the hexons of Ad5 and Ad2 (Kinloch *et al*, 1984), which are both members of the subgroup C adenoviruses. It can be seen that even in the highly variable regions certain positions are occupied by amino acids which are conserved across all four serotypes. These amino acids may have a strategic importance in the chain folding of the polypeptide, and may therefore be responsible for conserving the overall structure of the hexon. The most striking difference between Ad41 and Ad2 is the absence in the former of 32 amino acids present in Ad2. This region in Ad2 (139-170) is extremely acidic, containing 16 contiguous glutamic and aspartic acid residues, and is thought to be responsible for conformational changes at the surface of the virion on exposure to different pHs.

Similarly, comparison of the Ad40 and Ad2 hexons reveals an overall identity of 76%. If the sequence is divided into regions in a similar manner to that described above, the homology again rises to more than 90% at the N- and

C-termini while in the two central regions it falls to 40% and 52% respectively. The 32 amino acid highly acidic region present in Ad2 but not in Ad41, is similarly absent from the Ad40 hexon.

When Ad40 and Ad41 are aligned an overall identity of 88% is observed. The major changes are concentrated in the regions 131-287 and 390-425 where the homology drops to 56% and 27% respectively. The remaining 79% of the polypeptide displays 98% identity. This is in support of the general observation that there is greater sequence conservation between members of the same subgroup. Thus the subgroup F serotypes (Ad40 and Ad41) show higher homology to each other than to the subgroup C serotypes (Ad2 and Ad5), and *vice versa*.

1.4. Graphic representation of pairwise comparisons between hexon amino acid sequences.

To examine further the homology patterns detailed in the previous section, a series of pairwise comparisons were performed. In this analysis padding characters were inserted to allow maximal alignment of each pair of sequences. The aligned hexons were then divided in to runs of 25 amino acids and the percentage identity plotted (figure 8). The approximate location of the surface loops l_1,l_2,and l_4 are shown on each plot, these being identified by comparison to the three-dimensional structure of the Ad2 hexon. The comparisons Ad41 versus Ad2 and Ad40 versus Ad2 exhibit homology profiles

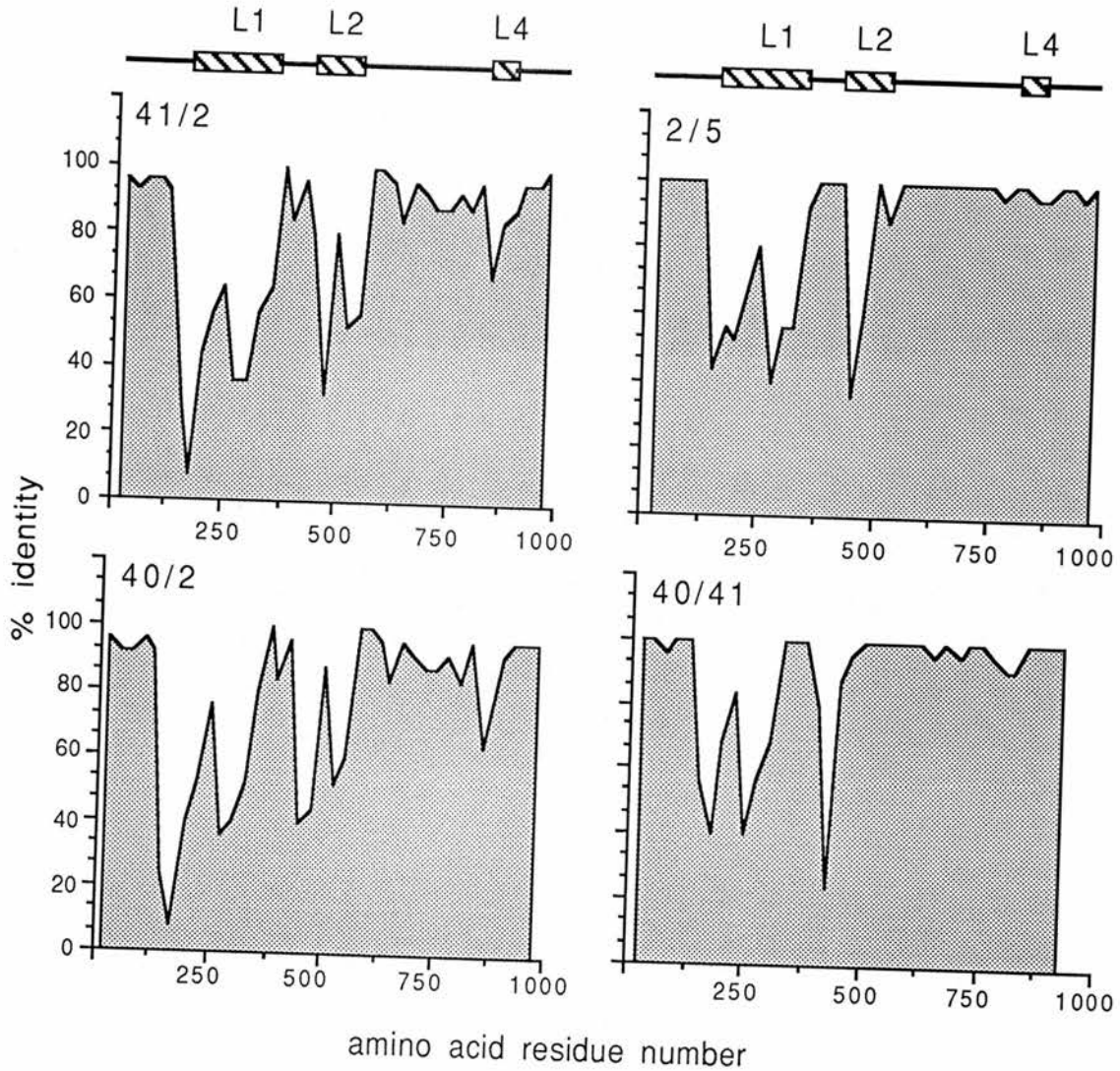


Figure 8. Histograms representing paired homology comparisons between known human adenovirus hexon sequences. Sequences were aligned and the percentage identity calculated for each consecutive group of 25 amino acids. Deletions and insertions were interpreted as mismatches. I_1, I_2 and I_4 indicate the positions of the surface loops, based on the Ad2 sequence.

that are almost superimposable. Loops 1 and 2 can each be divided into two domains of low homology, while loop 4 is more highly conserved. Intrasubgroup comparisons (Ad40 versus Ad41 and Ad2 versus Ad5) also yield virtually identical profiles, however these differ from the intergroup pairings with respect to the loop 2 region. While loop1 appears to possess two areas of low sequence conservation between subgroup members, loop2 only significantly differs in the first half, the second half being reasonably well conserved. If the amino acids which align with the second half of the Ad2 loop2 are examined (ie. 498-543 of the Ad2 hexon, the sequences of the four serotypes are given in figure 7), 23 out of the 46 are common to all four serotypes. Of the 23 which differ, 22 are conserved between Ad40 and Ad41 and 20 between Ad2 and Ad5. Furthermore, at each of these 23 positions the subgroup C sequence is different to that of the subgroup F.

1.5. Are the P1 domains of the Ad40 and 41 hexons structurally similar to that of Ad2?

The P1 and P2 domains of Ad2 are topologically identical, consisting of eight-stranded, flattened ß-barrels arranged in a "jelly-roll" topology (Richardson,1981). This "jelly-roll" formation is found in the coat proteins of small, spherical plant viruses (Olson *et al*, 1983; Rossmann *et al*, 1983), rhinoviruses (Rossmann *et al*, 1985) and poliovirus (Hogle *et al*, 1985). Although not unique to viral proteins this topology is a frequently observed feature of viral architecture (Liljas, 1986). The question to be asked is whether

the ß-barrel formation of the P1 domain is conserved in the Ad40 and 41 hexons, despite the absence of 32 consecutive amino acids present in Ad2, some of which contribute towards this domain? From the three-dimensional model of the Ad2 hexon it would appear that the D-strand is integral to the overall structure of the P1 domain and therefore, that significant changes to the amino acid sequence could not be tolerated. It is difficult to envisage how the enteric adenoviruses can accommodate such a major alteration in sequence and still maintain the stability of the hexon. One would have to assume that residues present in Ad40 and 41 can form a rather different D-strand and this would be accompanied by a shorter I_1 loop which would fold differently to its Ad2 counterpart. Structural data (Roberts *et al*,1986) indicates that, of the 16 amino acids which comprise the D-strand of Ad2, 5 are located on the upper external surface of the base and are therefore not involved in hexon-hexon contacts. Thus, only eleven residues (130-140) are important for the integrity of the P1 domain. Listed below are the eleven amino acids which comprise the D-strand of Ad2 and the corresponding residues of types 40, 41 and 5.

	130									
Ad2	Asn	Ser	Cys	Glu	Trp	Glu	Gln	Thr	Glu	Asp	Ser
Ad40	Asn	Pro	Ser	Gln	Trp	Thr	Asn	Gln	Asn	Lys	Thr
Ad41	Asn	Pro	Cys	Glu	Trp	Lys	Asp	Asn	Asn	Lys	Ile
Ad5	Asn	Pro	Cys	Glu	Trp	Asp	Glu	Ala	Ala	Thr	Ala

Of the amino acids predicted to occupy the equivalent position in Ad41, four are identical to Ad2 and three represent conservative changes. To determine whether the altered amino acids have the potential to form a ß-sheet we utilised the secondary structure prediction programme of the University of Wisconsin Genetics Computer Group Software Analysis package. Applying the method of Robson-Garnier (Garnier *et al*, 1978) it was predicted that these amino acids could form a ß-sheet. However it was noted that the programme was very inaccurate in predicting many of the known structural features of the hexon determined by crystallography. This is hardly surprising for a molecule of the size and complexity of the hexon, however it serves to demonstrate that little weight can be attached to predictions obtained from such programmes.

Recently, the structure of the Ad2 hexon has been refined (R factor of 21%) enabling an investigation of this problem on a stereochemical basis. The following modelling work and that in Section 1.6. was performed by our collaborators Dr Roger Burnett and Dr Ramachandran Murali of the Wistar Institute, Philiadelphia. Using the molecular modelling programme FRODO an attempt was made to build a D-strand based on a hybrid of Ad40 and 41. It was considered that residues 130-134 were sufficiently well conserved to be omitted from the model building. Keeping the Ad2 backbone constant, the altered D-strand residues were substituted, and displayed in a standard conformation. When a side-chain suffered steric contact with neighbouring atoms, it was rotated by varying the torsion angles so that it could be fitted without violating van der Waal's contacts. Torsion angles were always well

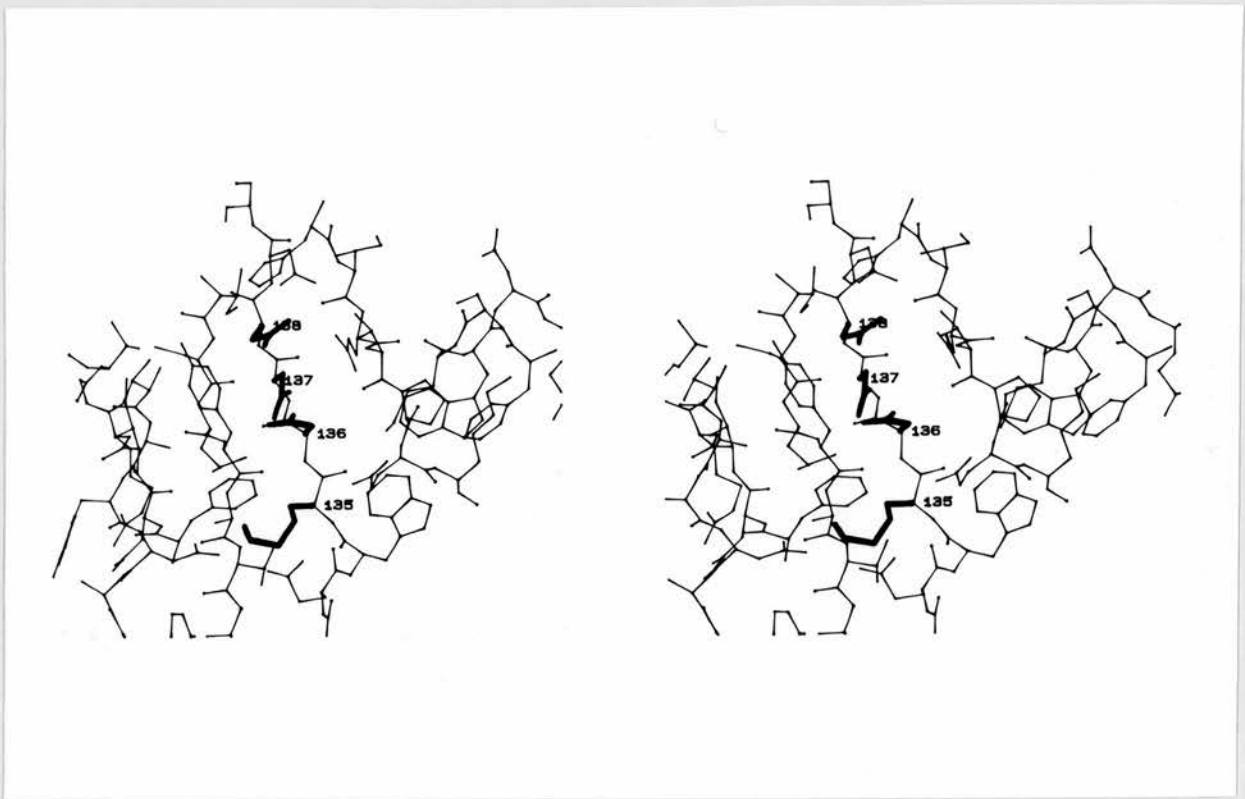


Figure 9. A stereo-view of the D-strand of the P1 ß-barrel, with the modelled residues from Ad40 and 41 shown in bold. Amino acids in the D-strand of the Ad2 hexon were changed for the corresponding Ad40 or 41 residues, and the side chains fitted as described in the text. Note that the ß-strand cannot continue above residue 138 due to the bulky histidine side-chain, seen just above residue 139.

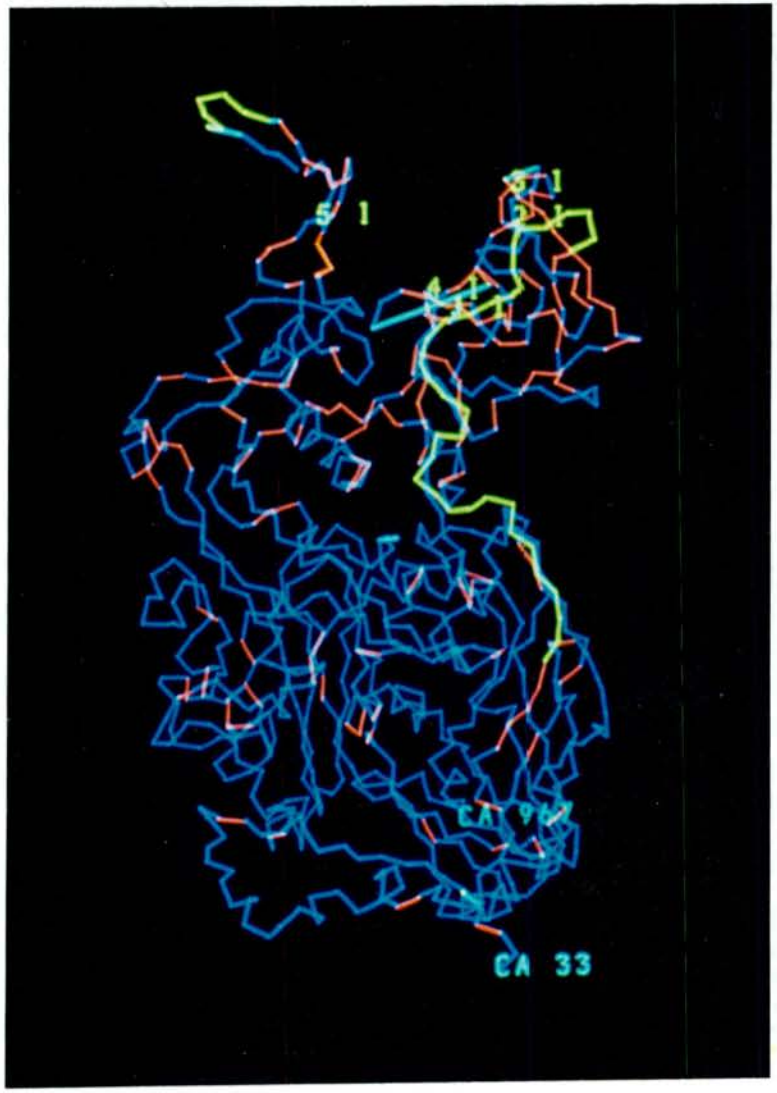
within statistically permitted values. Of the amino acids occupying position 135, the Ad41 lysine was chosen for modelling as this was the more bulky. Since this could be successfully substituted, it was therefore assumed that the smaller threonine of Ad40 could also be accomodated. The following three residues of Ad40 also fitted well, however the D-strand could not be continued beyond position 138 due to the bulky histidine side chain seen just above residue 139 (figure 9). Although there is sufficient space for a small amino acid, the lysine present in both Ad40 and 41 is too large. Upon formation of the trimer the available space is reduced even further as the top of the D-strand becomes buried internally. These results suggest that the amino acid changes can be accommodated while maintaining the eight-stranded ß-barrel, however there would be a shortening of the D-strand to 9 residues, and loop1 of both types 40 and 41 would commence earlier than in the subgroup C viruses.

1.6. Structural predictions using the three-dimensional model of the adenovirus type 2 hexon.

The high degree of homology between hexons has allowed those of the subgroup F adenoviruses to be fitted to the three-dimensional X-ray crystallographic structure of Ad2 (Roberts *et al*, 1986). The positions of amino acid substitutions, 'insertions'and 'deletions' relative to the Ad2 sequence are displayed in figures 10, 11, and 12. Figure 10 represents Cα traces of the Ad2 hexon showing the location of the Ad41 mutations. Clearly visible is the region

Figure 10. Alpha-carbon trace of the Ad2 hexon monomer showing the positions of the altered amino acids found in Ad41, (A) view of the molecule normal to the base orientated to be looking from within the central cavity of thesub-unit towards the B hexon:hexon contact face; (B) viewed from the side. Colour coding is as follows: amino acid substitutions - red; residues absent in Ad41 - yellow; regions of Ad2 for which the structure has not yet been determined - light blue; insertions - "I".

(A)



(B)

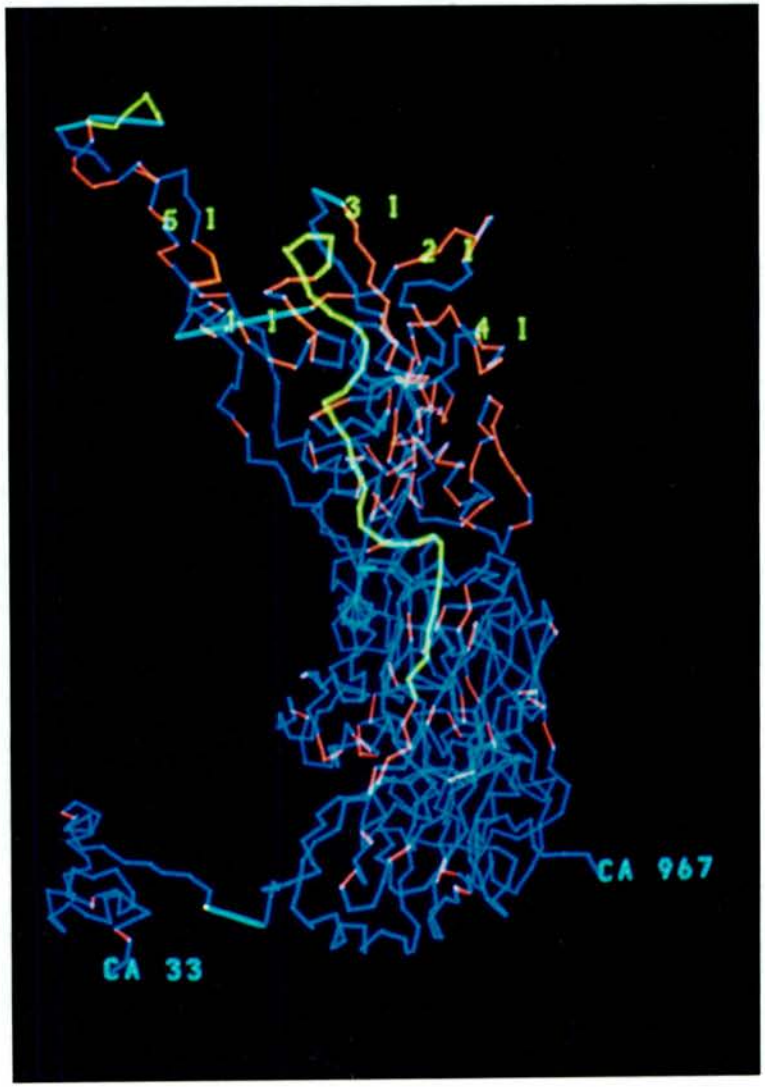
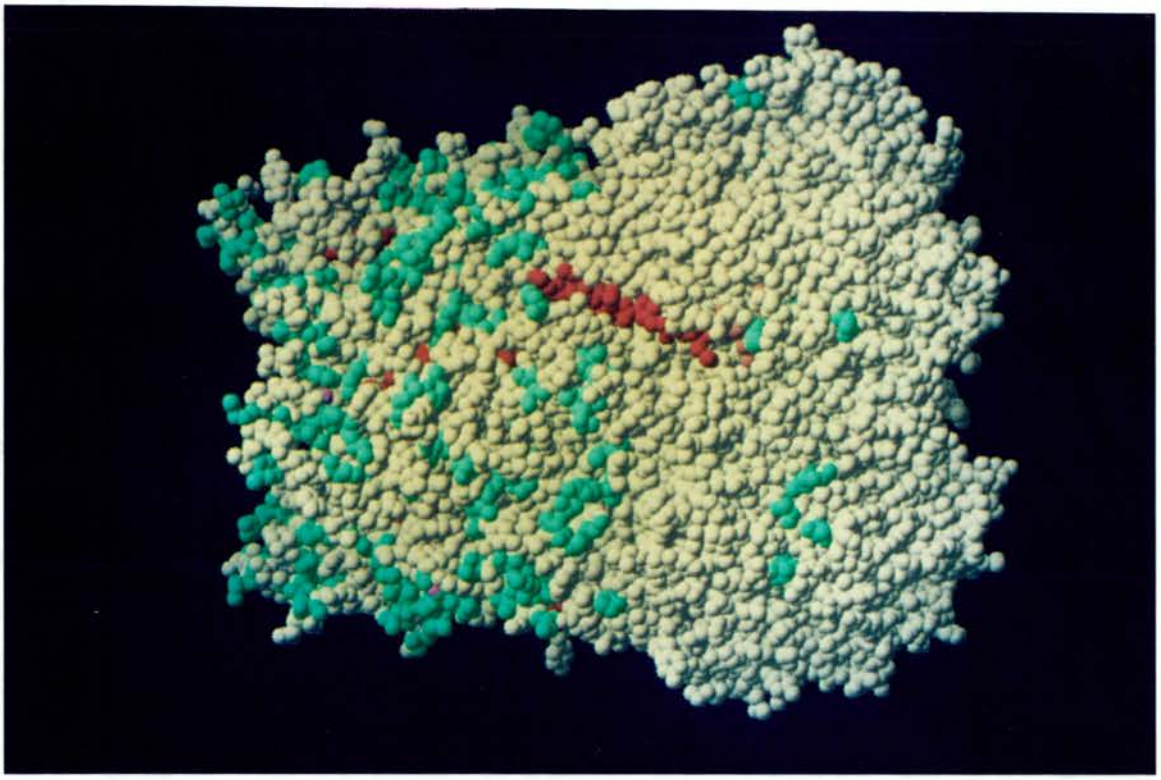
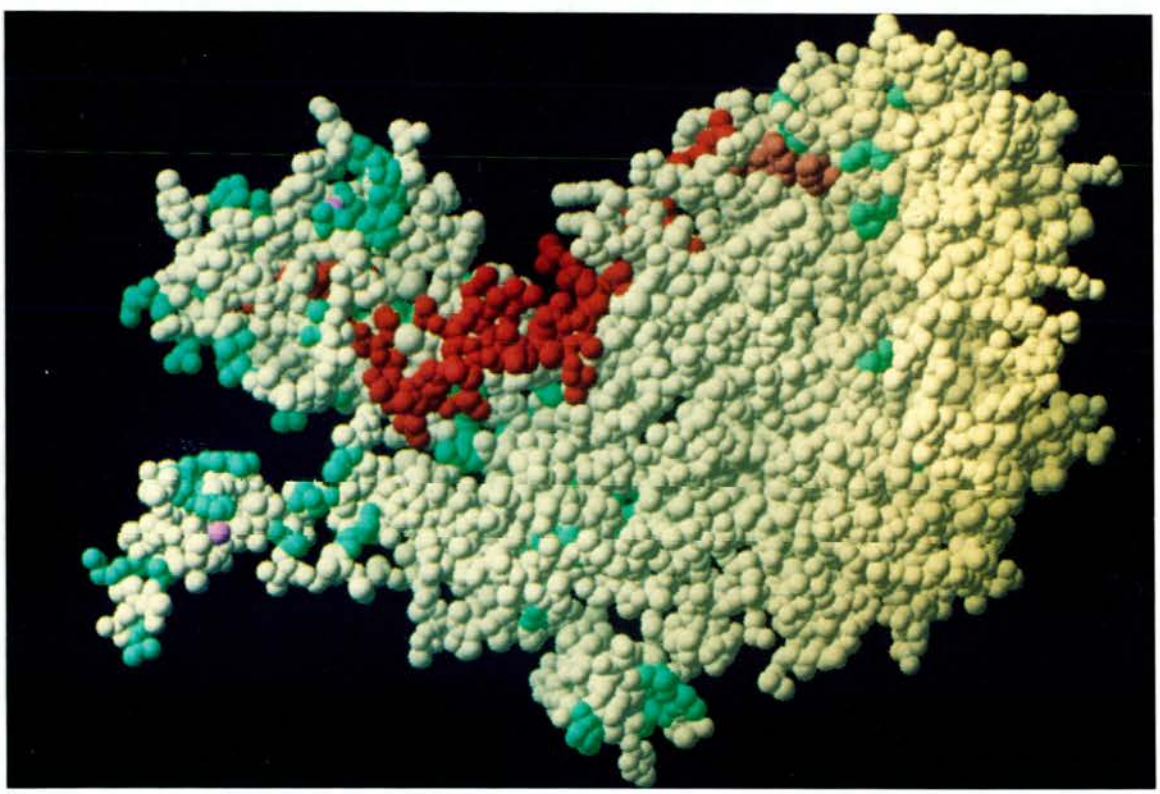


Figure 11. (A) space-filling CPK model viewed from within the central cavity to show changes in the surface residues. Red atoms represent position of deletions in Ad40 relative to Ad2, pink are insertions and green are substitutions. The brown atoms indicate the modelled part of the D-strand shown in figure 9. One of the insertions in loop 1 is invisible in this view as it is buried inside the molecule. (B) The Ad40 changes are represented on the trimer. Colour coding as for (A).



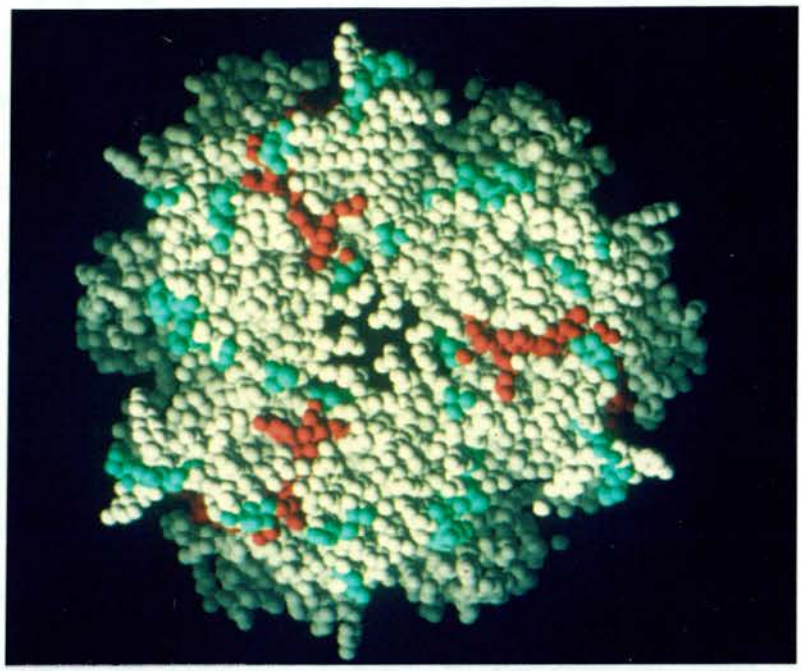
(B)



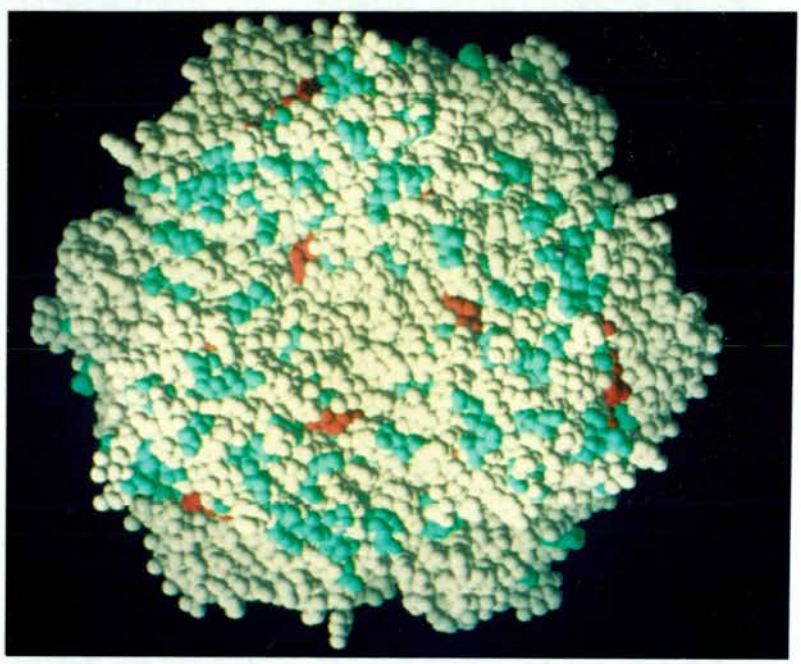
(A)

Figure 12. Space-filling models of the Ad40 hexon trimer viewed from above looking downward towards the interior of the virion, with colour coding as in figure 11. (a) intact hexon; (b) sectioned at the base:tower junction with the top part removed to show internal changes; (c) sectioned through the ß-barrels with the upper part removed.

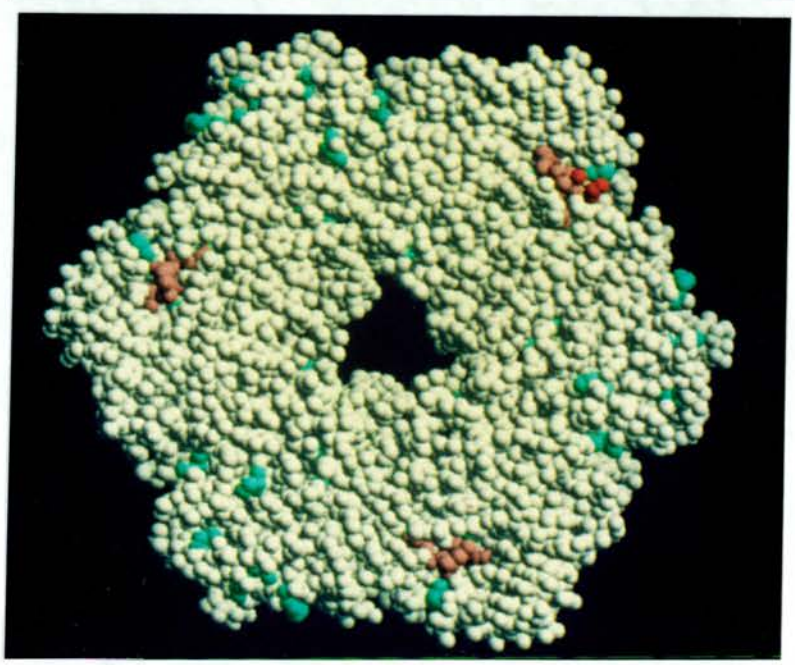
(a)



(b)



(c)



absent in Ad41, extending from the top of loop1 down into the P1 domain. The majority of the amino acid substitutions are concentrated in loop1 and loop2, with the base being well conserved. Space-filling models allow an even clearer visual representation of sequence changes, and these have been produced to show the positions of the mutations in the Ad40 sequence, in the context of the monomer and of the trimer as a whole (figure 11 A and B respectively). Figure 12 a,b and c show the positions of internal changes viewed from the top and from within the hexon. Figures 12a and 11B show clearly that the tower contains the majority of the mutations, while the majority of residues which are absent or inserted in the Ad40 hexon, are buried internally. The D-strand which suffers a number of alterations in Ad40, is partially exposed. A section through the junction between top and base of the trimer (figure 12b) reveals that the majority of changes are on the periphery with the bulk of the core residues conserved. A further section through the ß-barrels (figure 12c) shows the high conservation in the P1 and P2 domains. The brown atoms show the position of the Ad2 D-strand residues substituted during molecular-modelling of the 'enteric' D-strand. The pseudo-hexagonal form of the base is particularly clear in this view.

1.7. Features of the sequence upstream and downstream of the adenovirus type 41 hexon gene.

Shotgun sequencing of pCT3 and pCT6 yielded DNA sequence information

```
     A  A  P  T  V  A  V  A  T  P  A  R  R  V  R  G  A  N  W  Q
   1 GCGGCCCCAACCGTTGCCGTGGCCACTCCGGCTCGCAGAGTTCGCGGTGCCAATTGGCAG
     S  T  L  N  S  I  V  G  L  G  V  K  S  L  K  R  R  R  C  Y
  61 AGCACTTTGAACAGCATTGTGGGTTTGGGAGTGAAAAGCCTAAAACGCCGCCGGTGCTAT
     *
 121 TAAAGTGGAACCAAGCTAAATACCCTTATTGTATGCGCCTCCTGTGTTACGCCAGACAGC
                                          M  A  T  P  S  M  M  P
 181 CGAGTGACGGGTCACCGCCAAGAGCGCCGCTTGCAAGATGGCCACCCCCTCGATGATGCC
      Q  W              R  T  P  F  S  A  G  N  A  T  T  *
 241 GCAATGG............CGTACGCCGTTCTCGGCCGGTAACGCCACCACATAAGAAGC
           M  G  S  S  E  Q  E  L  V  A  I  A  R  D  L  G  C  G
3001 TGGCCAATGGGCTCCAGCGAGCAGGAGCTGGTCGCCATCGCGCGTGACCTGGGCTGCGGA
     S  Y  F  L  G  T  F  D  K  R  F  P  G  F  M  A  P  N  K  L
3061 TCCTACTTTCTAGGCACCTTTGACAAGCGTTTCCCGGGCTTTATGGCTCCGAACAAGCTG
     A  C  A  I  V  N  T  A  G  R  E  T  G  G  V  H  W  L  A  L
3121 GCCTGCGCTATTGTGAACACGGCCGGGCGTGAAACAGGGGGCGTTCACTGGTTGGCCCTG
     A  W  N  P  K  S  H  T  C  Y  L  F  D  P  F  G  F  S  D  E
3181 GCCTGGAACCCCAAGAGCCATACCTGTTACCTCTTTGACCCGTTTGGCTTTTCAGACGAG
     R  L  K  Q  I  Y  Q  F  E  Y  E  G  L  L  K  R  S  A  L  A
3241 CGGCTCAAACAAATTTACCAGTTTGAATATGAGGGCCTACTAAAGCGAAGCGCTTTGGCC
     S  T  P  D  H  C  I  T  L  V  K  S  T  Q  T  V  Q  G  P  F
3301 TCCACGCCTGACCACTGCATTACCCTTGTGAAATCCACCCAGACCGTTCAGGGGCCTTTT
     S  A  A  C  G  L  F  C  C  M  F  L  H  A  F  I  H  W  P  S
3361 TCTGCCGCCTGTGGTCTTTTCTGCTGCATGTTTTTGCATGCTTTTATACACTGGCCCAGT
     N  P  M  E  Q  N  P  T  M  D  L  L  T  G  V  P  N  S  M  L
3421 AATCCAATGGAGCAAAATCCCACAATGGATCTTCTCACGGGCGTACCGAACAGCATGCTT
     Q  S  P  Q  V  E  P  T  L  R  R  N  Q  E  R  L  Y  R  F  L
3481 CAAAGCCCCCAGGTTGAACCCACCCTGCGTCGGAATCAGGAACGATTGTATCGTTTTCTG
     T  Q  H  S  P  Y  F  R  R  H  R  E  R  I  E  K  A  T  A  F
3541 ACCCAACACTCTCCCTACTTTCGTCGTCACCGCGAGCGAATTGAAAAAGCTACAGCTTTT
```

Figure 13. Nucleotide sequence upstream and downstream of the Ad41 hexon gene. Translation of the 3' end of the pVI gene, and the majority of the protease gene is given in single letter amino acid code above the middle base of each codon. The middle of the hexon gene has been omitted (and replaced by dots), leaving only the sequence of the extreme 5' and 3' ends.

flanking the Ad41 hexon gene (figure 13). Comparison of the upstream region with the Ad2 genome revealed that pCT3 contains the 3' end of the gene encoding pVI, the precursor to the structural protein VP VI. Figure 14a represents an alignment of the C-termini of the Ad2 and Ad41 pVI polypeptides. This shows that there is high homology at the extreme C-terminus which decreases further into the polypeptide.

Examination of the region downstream of the hexon gene (figure 13), reveals an uninterrupted open reading frame which extends to the end of the sequence determined. In Ad2 the corresponding region codes for the proposed 23kD viral protease. Comparison of the predicted Ad41 amino acid sequence with the Ad2 protease (figure 14b) shows that out of 198 residues, 155 are identical and many of the changes are conservative.

The Diagon programme (Staden, 1982b) which utilises a score-matrix to identify homology between protein or nucleic acid sequences, was applied to the sequence upstream of the type 41 hexon gene. Nucleotides 1-400 (figure 4) were compared to nucleotides 18628-19027 of the Ad2 genome and the resulting homology plot is displayed in figure 15. This clearly demonstrates the similarity between the 3' ends of the pVI genes and between the 5' ends of the hexon genes. It is also shows that there is little homology between the intergenic sequences of Ad41 and Ad2. Both non-coding regions which flank the Ad41 hexon gene are shorter than their Ad2 counterparts.

a) pVI

```
Ad41   AAPTVAVATP ARRVRGANWQ STLNSIVGLG VKSLKRRRCY
                    * *   *** ********** * *******
Ad2    LRRAASGPRS MRPVASGNWQ STLNSIVGLG VQSLKRRRCF
```

b) 23K

```
                                                                    50
Ad41   MGSSEQELVA IARDLGCGSY FLGTFDKRFP GFMAPNKLAC AIVNTAGRET
       ******** * *  ***** * **** ***** **  * **** **********
Ad2    MGSSEQELKA IVKDLGCGPY FLGTYDKRFP GFVSPHKLAC AIVNTAGRET

                                                                   100
Ad41   GGVHWLALAW NPKSHTCYLF DPFGFSDERL KQIYQFEYEG LLKRSALAST
       ***** * ** ** * *****  ****** ** ** ******  ** *** **
Ad2    GGVHWMAFAW NPRSKTCYLF EPFGFSDQRL KQVYQFEYES LLRRSAIASS

                                                                   150
Ad41   PDHCITLVKS TQTVQGPFSA ACGLFCCMFL HAFIHWPSNP MEQNPTMDLL
       ** **** ** ** **** ** ********** ***  **  * *  **** *
Ad2    PDRCITLEKS TQSVQGPNSA ACGLFCCMFL HAFANWPQTP MDHNPTMNLI

                                                                   198
Ad41   TGVPNSMLQS PQVEPTLRRN QERLYRFLTQ HSPYFRRHRE RIEKATAF
       ******** * *** ****** ** ** **   ****** *    *  ** *
Ad2    TGVPNSMLNS PQVQPTLRRN QEQLYSFLER HSPYFRSHSA QIRSATSF
```

Figure 14 a) comparison of the predicted sequence of the C-terminus of the Ad41 pVI polypeptide with that of Ad2. b) comparison of the predicted Ad41 protease sequence with that of Ad2. Identical amino acids are indicated by asterisks.

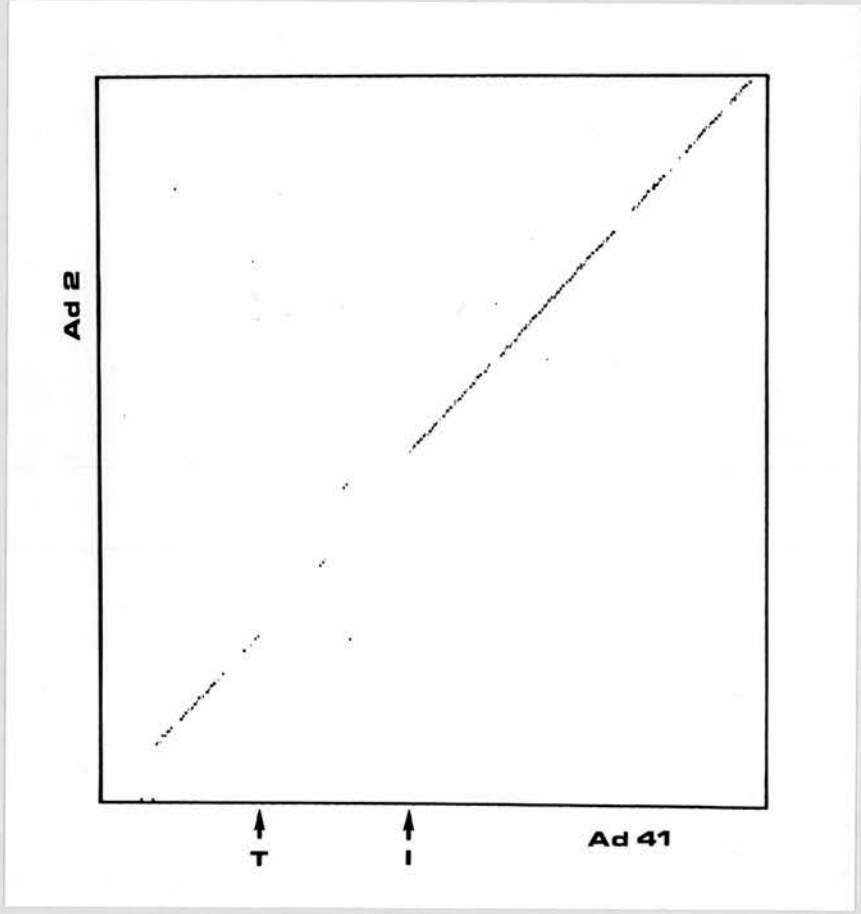


Figure 15. Comparison by Diagon plot analysis of the upstream regions of the Ad2 (nucleotides 18628-19027) and Ad41 (fig. 4, 1-400) hexon genes. The parameters applied were a span of 13 and a score of 9. A dot indicates that at least 9 out of 13 consecutive nucleotides are identical, thus a line of dots indicates significant homology between two sequences. T denotes the termination codon of the polypeptide VI precursor gene, and I denotes the initiation codon of the hexon gene.

SECTION 2. **The use of hexon-derived oligonucleotide probes for diagnosis of Ad40 and 41 infection.**

2.1. Design and specificity of probes.

An attempt was made to develop a method for diagnosis of enteric adenovirus infection in clinical samples using oligonucleotide probes in a dot-blot assay system. Examination of the Ad40 and Ad41 hexon gene sequences, allowed identification of oligonucleotides which were likely to be unique to each virus. Oligonucleotide Ad41/1 was derived from a sequence coding for residues predicted to be in loop1 and which are relatively non-conserved across the four serotypes. Ad41/2 and Ad40/1 were derived from a region coding for the highly conserved C-terminus of the hexon, but which contained a number of third-base changes at the codon level. The nucleotide sequences of all four serotypes in the regions yielding the information for the design of these probes are given in figure 16.

Hybridisation and washing conditions were established for each probe using dot-blots onto which Ad40, Ad41 and Ad2 hexon-containing plasmids had been spotted. Overnight hybridisation at 45°C followed by two 15 min washes in 0.5xSSC, 0.25% SDS was found to be suitable for all three probes, and under these conditions no heterotypic cross-reactivity was observed (figure 17). Each plasmid was serially diluted ten-fold across the dot-blot, thus allowing the sensitivity of each probe to be determined. Oligonucleotide probes displayed sensitivities down to 100pg plasmid DNA following overnight autoradiography. When the blots were probed with the nick-translated 420bp Ad2 hexon fragment (described in Results Section 1. 1), the sensitivity

a) loop 1 region.

```
Ad2    GTC TAT GCC CAG GCT CCT TTG TCT GGA
               *   **  **  **  *       **
Ad5    GTA TTT GGG CAG GCG CCT TAT TCT GGT
               *   **  **  **  *       **
Ad40   TCC TTT GGA CAA GCT CCC TAT ATA GGA
               *   **  **  **  *       **
Ad41   GTT AGA GGG CAA GCT CCC TTT ATA GGA    (Ad41/1)
```

b) C-terminal region.

```
Ad2    C CTG GTG CAG ATG TTG GCC AAC
         **   *  **  ***  *       *
Ad5    C CTG GTA CAA ATG CTA GCT AAC
         **   *  **  ***  *       *
Ad40   C CTC ATA CAA ATG CTC AGT CAC   (Ad40/1)
         **   *  **  ***  *       *
Ad41   T CTG ATT CAG ATG CTT AGT CAT   (Ad41/2)
```

Figure 16. Nucleic acid sequence of the two regions of the hexon gene which provided the information for designing specific oligonucleotide probes. Nucleotides which are conserved across the four genes are indicated by asterisks.

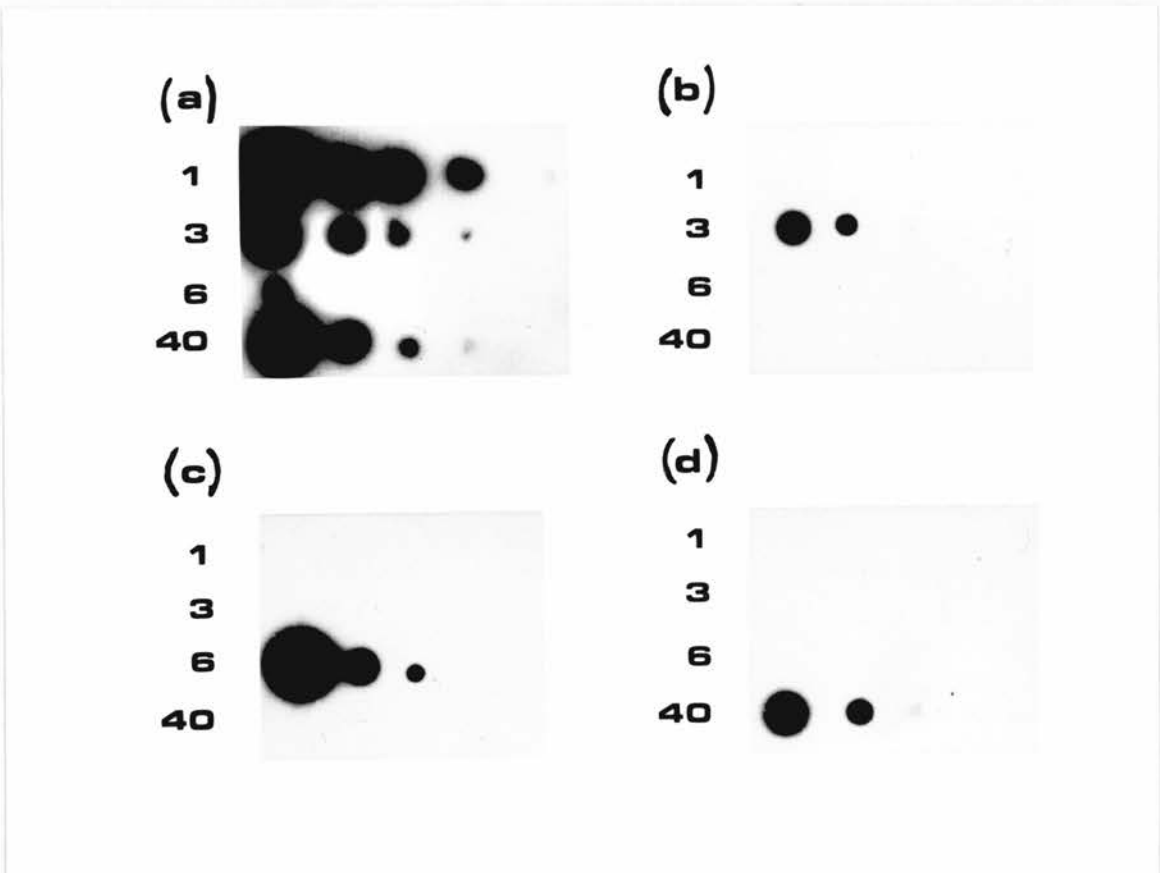


Figure 17. Dot-blot analysis of plasmid DNA to demonstrate the specificity of the oligonucleotide probes. Four identical blots were assembled consisting of 10 fold dilutions of pCT1, pCT3, pCT6 and pCT40, starting from 10ng (left-hand side) to 1pg (right-hand side) DNA. Blots were probed with a) nick-translated Ad2 420bp hexon fragment, b) Ad41/1, c) Ad41/2, d) Ad40/1. Following overnight hybridisation, oligonucleotide probe filters were washed in 0.5xSSC / 0.25% SDS at 45°C, and the Ad2 probe filter in 2xSSC / 0.25% SDS at 65°C.

increased to 1pg for homologous DNA. The Ad2 probe did not hybridise to pCT6 DNA, since this plasmid lacks the 5' end of the hexon gene. The signal seen in the lowest dilution of pCT6 using the Ad2 probe results from homology between the pUC13 polylinker sequences present both in the probe and in the plasmid.

Ad40/1 and Ad41/1 were analysed for type-specificity by hybridising to dot-blots of genomic DNA from a number of adenovirus serotypes. While attempts were made to ensure that the amount of DNA was equivalent for each type, only a small amount of DNA was available for simian Ad7 and Ad12. A control blot was therefore included using the Ad2 hexon probe, to give an indication of the quantity of DNA present. Both oligonucleotide probes hybridised only to homotypic DNA, showing no visible cross-reactivity with the other serotypes (figure 18).

2.2. Analysis of stool samples and infected cell extracts.

Having established their specificity and sensitivity, the oligonucleotides were analysed for their usefulness as diagnostic probes in a series of dot-blots of clinical samples and infected cell extracts. Virus isolates originating from stool samples and passaged twice in 293 cells, were supplied by Barbara O'Donnell. These had been identified as positive for adenovirus by EM, and 10 out of the 11 further characterised as Ad41 by restriction enzyme analysis of

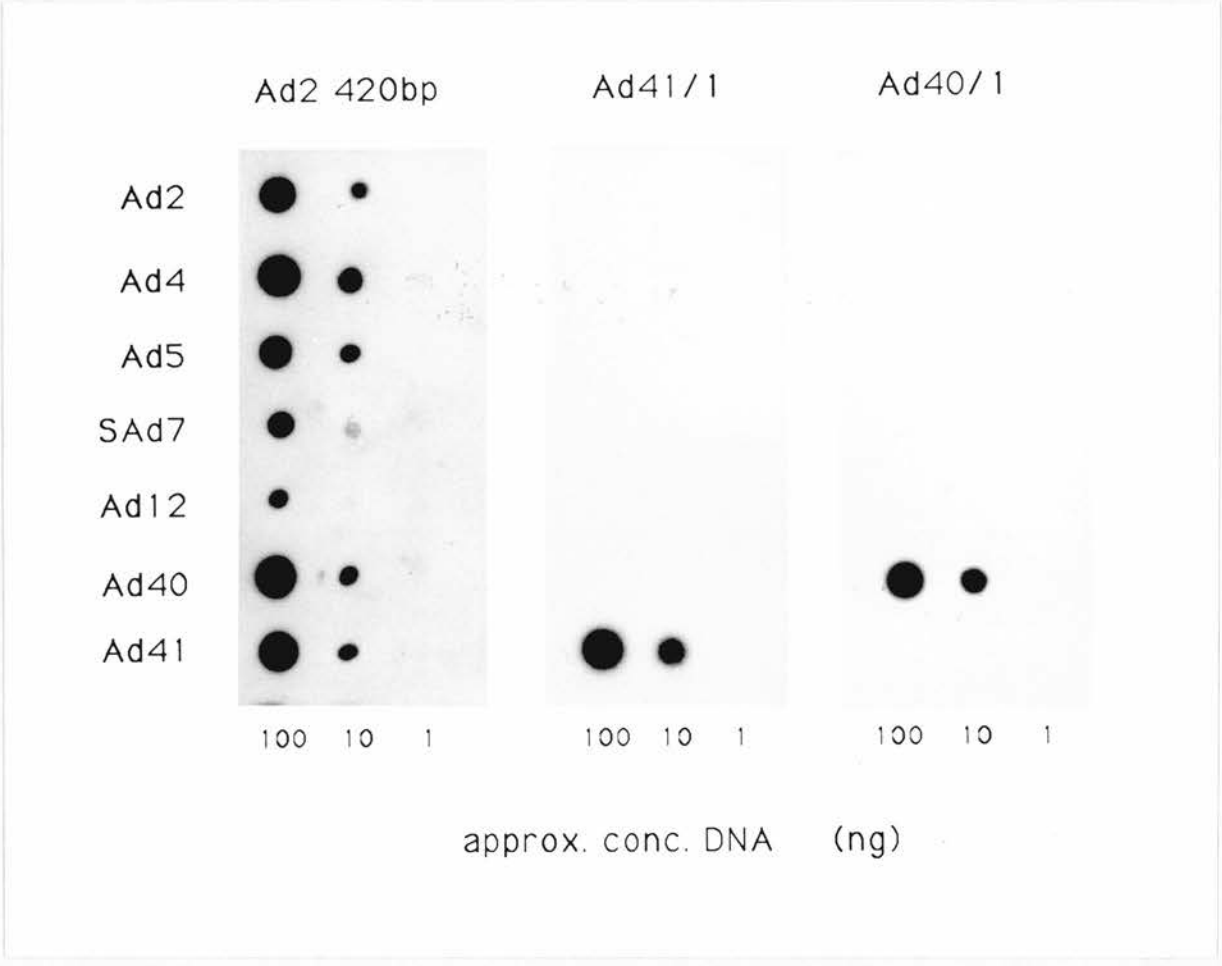


Figure 18. Dot-blot analysis of genomic DNA from different adenovirus serotypes (Ad2, Ad4, Ad5 simian Ad7, Ad12, Ad40 and Ad41) to test the type-specificity of the oligonucleotide probes. The probes were, a) nick-translated Ad2 420bp hexon fragment; b) Ad41/1 and c) Ad40/1. Washing conditions were as stated in figure 17.

labelled DNA. One sample had shown no cpe in 293 or Chang conjunctival cells and was supplied untyped as a first passage extract. Samples were blotted onto nylon and probed with Ad41/2 (figure 19). The stringency of washing of the oligonucleotide probed filters was reduced to maximise the chances of detecting positives, although under these conditions a small amount of cross-reactivity with the Ad40 plasmid pCT40 was observed. Six strong positives (1, 3, 4, 5, 6, 9) and two possible positives (10, 11) were identified. A second blot was assembled to test for the presence of adenovirus using the Ad2 probe. Using this probe 8 definite positives were identified: the six strong positives revealed by Ad41/2, one of the possible positives also revealed by Ad41/2 (sample 11) and sample 8 which had not previously been detected. Sample 10 was not positive with the Ad2 probe suggesting that the weak signal obtained with Ad41/2, and which was only around the edge of the spot, was non-specific. Virus was undetectable in the sample which failed to give a cpe in tissue culture, and in sample 2.

Eleven stool samples extracts were supplied by Prof. U. Desselberger. These had been found positive for adenovirus by EM, and typed as Ad41 using an Ad41-specific antibody in an ELISA test. Using the Ad41/2 probe, 3 out of 10 were identified as positive (figure 20). These were samples 2, 8 and 11. Using the Ad2 hexon probe, two additional samples (9 and 10) along with the three identified with Ad41/2 were positive. Samples 1, 3, 4, 5, 6, and 7 were negative with both probes.

Twelve stool extracts, typed by restriction enzyme analysis following growth in 293 cells, were supplied by Dr A.H. Kidd. Of these, 7 had been found to contain Ad40 and 5, Ad41. These samples were probed with each of the three oligonucleotides. The nick-translated Ad2 hexon probe was used as a means of identifying adenovirus DNA which may not hybridise to the oligonucleotide probes. The results are shown in figure 21. Of the Ad40 stool extracts, 6 out of 7 were positive for adenovirus using the Ad2 probe. Of these only 4 were positive with Ad40/1, N5234 and FB363 not being detected with this probe. Since FB273 was negative with both probes, it was assumed that the amount of virus was below the level of detection.. Of the Ad41 stool extracts, 5 out of 5 were positive with Ad41/1, Ad41/2 and the Ad2 probe.

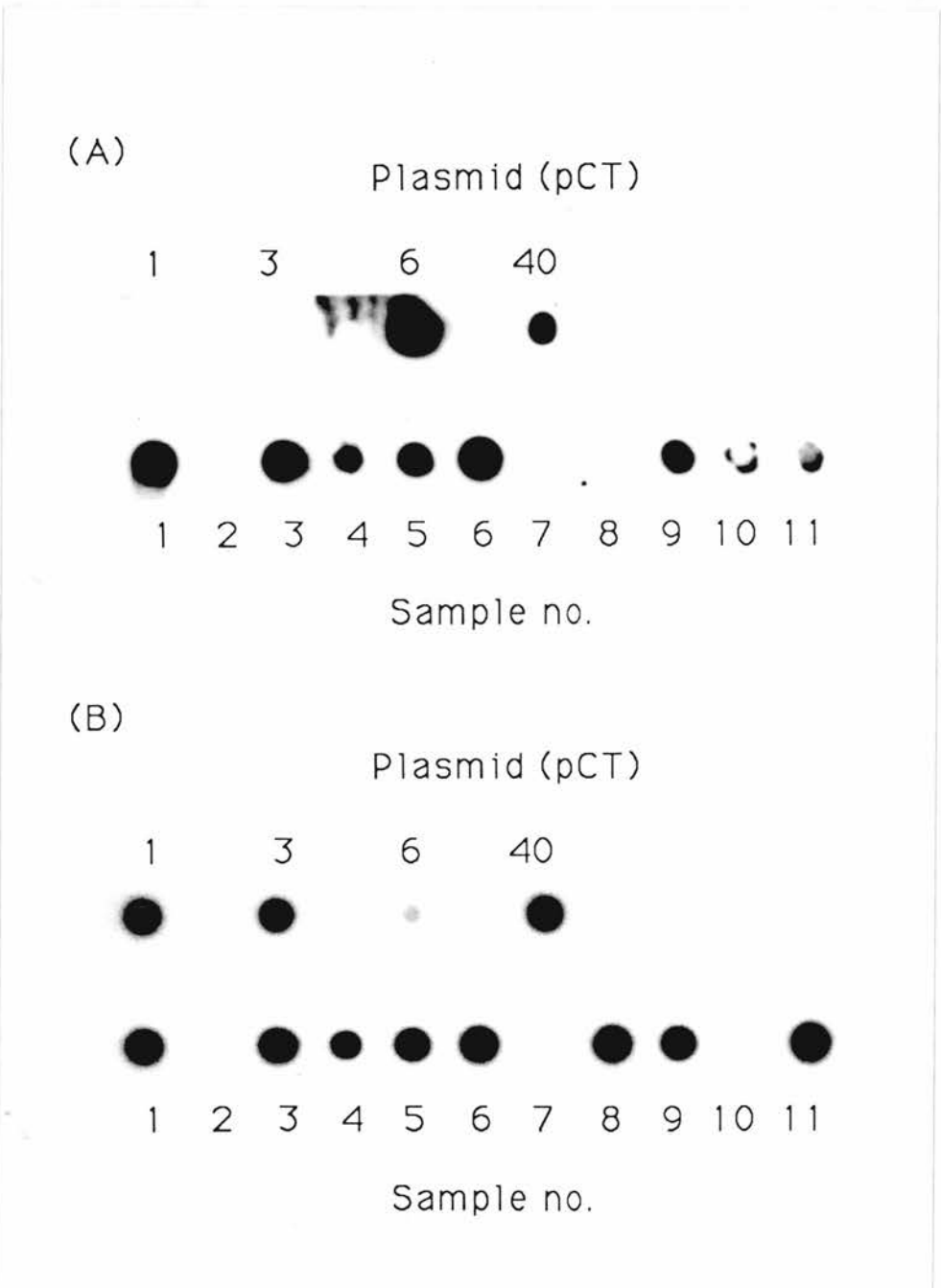


Figure 19. Dot-blot analysis of clinical samples passaged in 293 cells, supplied by Barbara O'Donnell. Filters were probed with a) Ad41/2; b) Ad2 hexon fragment. The oligonucleotide blot was given two 15min washes in 1xSSC, 0.25% SDS. The blot probed with the Ad2 fragment was washed under the standard conditions.

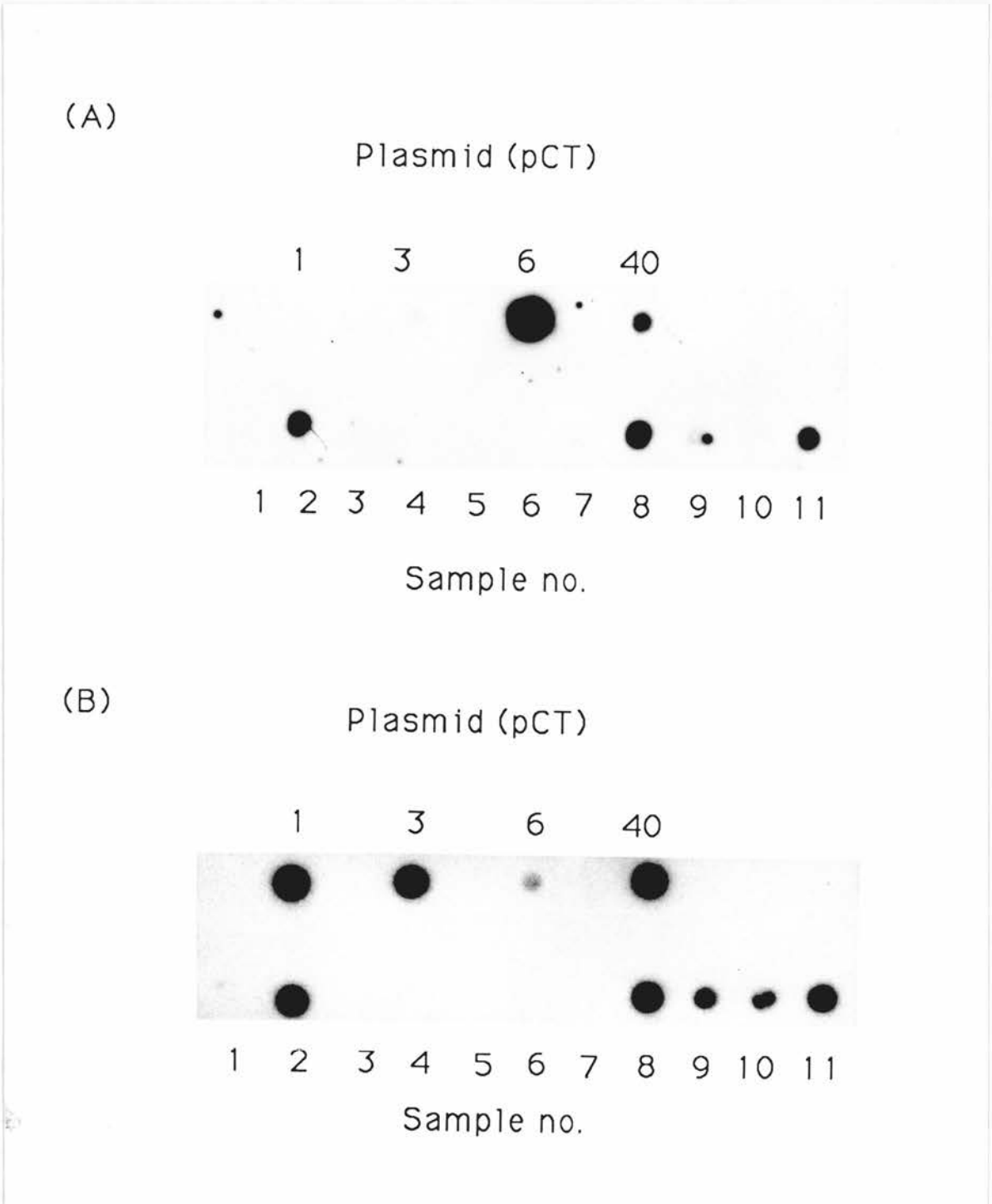


Figure 20. Dot-blot analysis of stool extracts supplied by Prof. Desselberger. Eleven samples, typed by conventional methods as Ad41 positive, were probed with a) Ad41/2 and b) with the Ad2 hexon fragment. The filters were washed in the manner describe in fig.19.

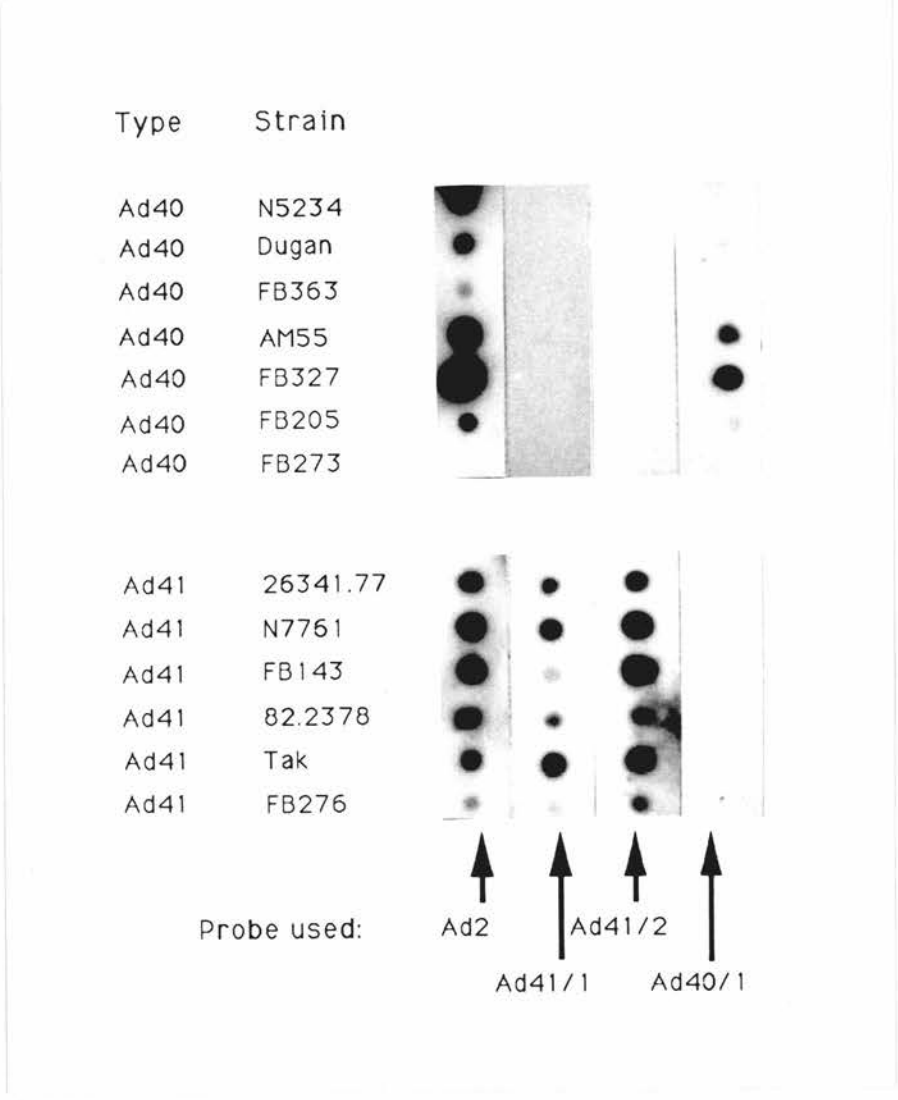


Figure 21. Dot-blot analysis of Ad40 and 41 stool extracts supplied by Dr A. Kidd. Samples were probed with each of the three oligonucleotide probes and with the Ad2 hexon fragment. Filters were washed in the manner described in fig. 17, to give maximum specificity.

DISCUSSION.

SECTION 1: The structure of the adenovirus type 40 and 41 hexon.

The high degree of sequence homology between hexon genes of different serotypes was invaluable during the initial stages of this work, firstly in the identification of the Ad41 gene using a homologous fragment of Ad2, then in the identification of the Ad40 gene with corresponding regions of Ad41. The nucleic acid sequence predicts proteins of 922 and 924 amino acids for the hexons of Ad40 and 41 respectively, which is significantly shorter than the 967 amino acids of Ad2, and the 951 amino acids of Ad5. This brings the number of hexons sequenced to date to five, four human (Ad2, 5, 40 and 41) and one bovine (BAV-3). In the absence of detailed structural information, little can be done to assess the significance of amino acid sequence variation, and such was the case when the Ad2 (Akusjärvi *et al*, 1984), Ad5 (Kinloch *et al*, 1984) and BAV-3 (Hu *et al*, 1984) hexon sequences were determined. However, comparison of primary sequences can give a good indication of regions of structural or functional importance, and may give a clue to the evolutionary paths of different serotypes. In the study by Kinloch *et al* (1984), the Ad5 and Ad2 hexons were aligned revealing eight separate, small 'deletions' in the Ad5 sequence relative to Ad2. Closer examination of the aligned sequences reveal that the non-conservative amino acid changes are concentrated in certain regions of the polypeptide. There is considerable overall homology between the two sequences with almost complete identity over the N-terminal 131 residues. However, this is followed by an abrupt drop in homology until residue 327 where the sequences return to near identity. A second drop is noted at position 431 and this continues until 479, before returning to high homology which is maintained to the C-terminus. Native Ad2 hexon possesses three sites

which are sensitive to trypsin, and it was therefore proposed that these are in regions exposed on the virion surface (von Bahr-Lindström *et al*, 1982), Interestingly, these sites are found to be located within the regions of low homology between Ad2 and Ad5, which thus correlates low amino acid conservation with exposure on the upper surface of the trimer. These cleavage sites are, however, absent from the Ad5 hexon and it was suggested by Kinloch *et al*, (1984) that in evolutionary terms Ad5 may have arisen as a protease resistant variant of Ad2. Since hexon contains a number of immunoreactive sites, immunological selection by type-specific antibodies is also likely to play a role in generating discrete serotypes within an adenovirus subgroup. An unusual feature which is apparent in the Ad2 hexon, is an extremely acidic region between residues 134 and 168. This feature is found to be conserved in Ad5, although the number of acidic amino acids is reduced from 21 to 16.

A comparable study performed by Hu *et al* (1984), reported similar observations to those detailed by Kinloch *et al*, (1984). Over 70% of the bovine hexon shares greater than 80% homology with Ad2, with the areas of low homology being confined to residues 130-331, 431-474 and 820-877 (numbering with respect to the Ad2 hexon). The bovine adenovirus 3 hexon, like the hexons from human serotypes, contains a number of subgroup- and type-specific antigenic determinants, however none of these epitopes are conserved between the two species. Thus it can be hypothesised that these immunogenic sites fall within the regions of least homology between bovine and human hexons.

When the Ad40 and 41 primary sequences are aligned with the other human serotype counterparts, this regional distribution of conserved domains is maintained. The areas of lowest homology lie between residues 135 and 285 (135-324 in Ad2), and 390 and 425 (430-470 in Ad2). To maximally align the hexons a number of gaps have to be inserted, five in Ad41 and seven in Ad40, relative to the Ad2 sequence, and these are referred to as "deletions" in the enteric hexon sequence. A number of gaps also have to be inserted in the Ad2 hexon, two to align with Ad40 and five to align with Ad41, however these are small, representing only one or two amino acids. These are referred to as "insertions" in the enteric hexons relative to Ad2. The terms insertion and deletion are not intended to imply that one sequence evolved from another, only to denote whether a particular sequence in the Ad2 hexon is correspondingly present or absent from the enteric virus hexons. The majority of the deletions in the Ad40 and 41 hexon are small apart from one region of 32 amino acids absent from both enteric sequences, which corresponds to residues 139 to 170 of the Ad2 hexon. This removes all of the highly acidic domain noted in both subgroup C hexons, and the physiological significance of this is a matter for speculation. In Ad2 this region is thought to be responsible for pH-induced conformational changes at the surface of the virion. During penetration, adenovirus is taken into acidic endosomes by receptor-mediated endocytosis. Following conformational changes at the virion surface, sites on the penton are exposed which interact with the endosomal membrane causing release of the virus into the cytoplasm. Whether the acidic region present in the subgroup C adenoviruses has an involvement in this process is not known,

however it seems reasonable to suggest that it may, since hexon antibodies have been shown to block this process. That the enteric adenoviruses then lack this region, could imply their uptake be less efficient than, or slightly different to, that of Ad2. This may partially explain the observation that a cytopathic effect in Ad40 and 41 infected cells takes longer to appear than with more conventional serotypes, which could suggest a slower proliferation. However, the absence of this acidic region could be beneficial in allowing the enteric adenoviruses to inhabit the highly acidic environment of the gut without undergoing pH-induced changes in conformation which, one assumes, would affect infectivity. Interestingly BAV-3 also lacks this acidic domain, however in contrast to Ad40 and 41 it is a pathogen of the respiratory system.

The elucidation of the three-dimensional structure of the Ad2 hexon (Roberts *et al*, 1986) has confirmed many of the tentative proposals concerning the domains of structural importance that were deduced from the primary sequence. Due to the high degree of amino acid sequence conservation between the Ad2 and the Ad40 and 41 proteins, it is possible to predict the likely structure of the enteric adenovirus hexons. The trimeric Ad2 hexon can be divided into two structural domains, the base and the tower (this can be seen clearly in Results figure 11b, and in Introduction figure iii). The base is comprised of three copies of two, eight-stranded ß-barrels which form the core of the capsid, and which are involved in contacts with neighbouring hexons (Introduction: figure. iiia). Arising from the basal domains are loops which form the surface of the virion, and although not contributing directly to the structure

of the capsid, confer great stability by virtue of their considerable interweaving (Introduction: figure. iiib). Examination of the aligned primary sequences allowed a prediction of the domains of structural importance, these being the regions of high homology between serotypes, and comparison with the Ad2 hexon model reveal these to correspond to the P1 and P2 ß-barrels. In contrast, the variable regions are confined to the l_1 and l_2 loops, and to a lesser extent l_4. A fourth loop, l_3, is well conserved, however it does not contribute to the virion surface, being sandwiched between the P1 and P2 domains. Within the base it is clear that spatially there is little room for accommodating amino acid substitutions, and this would appear to be the underlying reason for the high level of conservation. In contrast, the loop domains seem to be able to accommodate a multitude of amino acid substitutions, deletions and insertions without affecting the integrity of the hexon. It is noticeable however, that within the highly variable domains certain positions are occupied by residues which are conserved across the four serotypes, suggesting that these may be of strategic importance, perhaps in chain folding and in maintaining the overall framework of the hexon.

A major question to be addressed given the structural constraints placed upon the P1 and P2 domains, was : "How can the hexons of Ad40 and 41 tolerate the 32 amino acid deletion relative to Ad2, which extends from the top of loop1 down into the D-strand of the P1 domain and still conserve the integrity of the ß-barrel?" Since all virus coat proteins for which a structure has been determined contain domains with this same 'jelly-roll' topology, it would

seem unlikely for the enteric adenoviruses to be an exception. This problem was investigated using molecular modelling, however since this required considerable expertise and very powerful computer facilities, neither of which were available at St.Andrews, this aspect of the structural analysis was performed by our collaborators, R. Burnett and R. Murali at the Wistar Institute, Philadelphia. The Ad2 D-strand is formed by residues 130-145, however the latter six are located on the upper surface of the base and are therefore not important to the structure of the barrel. Comparison of the residues in Ad2 with those in Ad40 and 41 show that 130-134 are not significantly altered, while 135-138 are, and this is followed by a "deletion" from 139-170. Can the residues which follow 134 in types 40 and 41 be accommodated into the P1 domain and thus conserving the D-strand? To answer this question, a D-strand was built by substituting the Ad2 residue with the corresponding "enteric" residue, the bulkiest option being selected from Ad40 and 41 at each position. Thus Glu, Gln, Thr and Glu were replaced by Lys (from Ad41), and Asn, Gln and Asn (from Ad40). Following position 138 a constriction due to a histidine residue is encountered, which is overcome in Ad2 by the insertion of a small amino acid, aspartate. However, the corresponding residue in both Ad40 and 41 is lysine which is too bulky to continue the D-strand and is therefore assumed to signal the start of loop1. The new model suggests that the enteric adenovirus hexons are likely to possess D-strands which conserve the integrity of the P1 ß-barrel but which are seven amino acids shorter than their Ad2 counterpart. It is of interest that adenovirus type 5 also contains a deletion relative to Ad2 in this region, albeit much smaller, the residues which occupy

positions 135-138 being Asp, Glu, Ala, Ala. These are either structurally similar, or smaller than the Glu, Gln, Thr, Glu of Ad2, and hence would also be expected to be accommodated in the D-strand. Where the histidine constriction is encountered Ad5 has a threonine residue which, like the aspartate of Ad2, is small and therefore likely be accommodated in the space available. Thus it is predicted that the Ad5 hexon also possesses a D-strand, although whether it is as long as that of Ad2, has not been determined.

That the enteric adenovirus hexons are likely to possess P1 domains which are structurally similar to the Ad2 hexon is reassuring and, since the amino acid sequences corresponding to the P2 domains are extremely highly conserved, strongly implies that with respect to the hexon base, Ad40 and 41 closely resemble Ad2. In contrast, the surface regions are not conserved which implies that the architecture of the towers is significantly different, and this would allow each serotype to present a unique surface conformation to the environment.

Computer graphic programmes for modelling amino acid sequences have become extremely sophisticated over the last few years, and now present a very powerful tool for visualising complex three-dimensional protein structures. In addition, they allow one to pinpoint the locations of amino acid subsitutions, deletions and insertions found in other closely related proteins, and therefore highlight regions which are likely to be structurally different. The following modelling work was also performed by R. Burnett and R. Murali. The changes in the Ad40 and 41 hexons with respect to that of Ad2 have been

represented on a series of α-carbon traces and space-filling CPK models of the Ad2 protein. Two α-carbon traces of the hexon show the positions of the Ad41 changes, while spacing-filling models show the Ad40 changes with respect to the monomer, trimer, and through sections of the hexon interior. These serve to highlight the observations already made regarding the conservation of the base and the variability of the towers. In all the views, aspects of the deletion in loop1 and the P1 D-strand are clearly visible, and reveal that the major part of it is located on the exterior of the trimer. The region of the D-strand used for modelling the "enteric" D-strand, is shown on the CPK models. Despite the sophistication of the molecular modelling programmes it is impossible to predict structures of complex proteins such as the hexon, purely from the amino acid sequence. So, although the bases of the Ad40 and 41 hexons are probably sufficiently similar to that of Ad2 for the altered amino acids to be substituted and new models produced, the loops are far too complex and the number of changes too great, to even attempt any structural predictions.

To address the question "are some hexons more homologous than others?", a series of pairwise comparisons similar to that previously applied to Ad5 and Ad2 by Kinloch *et al* (1984), were performed and homology profiles produced. This analysis was intended to highlight only the relative sizes of the non-homologous domains rather than the actual amino acid changes *per se*. Using this approach inter- and intragroup comparisons were made, Ad40 and 41 being members of subgroup F and Ad2 and Ad5, subgroup C. The

intergroup homology profiles for Ad40 versus Ad2 and Ad41 versus Ad2 are superimposable, thus indicating that they are similar in the way they differ from Ad2. Similarly Ad40 and Ad41 differ from each other in virtually the same manner that Ad5 differs from Ad2, resulting in intragroup homology profiles which again are superimposable. However when intra- and intergroup profiles are compared, there appears one major difference between them and this lies within the loop2 domain. Examination of the loop2 sequences (approximately 50 amino acids) from members of different subgroups reveals a low level of conservation, however residues around the centre show greater homology which thus divides the loop into two domains of low homology. Comparisons between members of a subgroup also reveals high variability within loop2, however this is confined to the first half of the sequence which in turn corresponds to the first domain of the two observed in the intergroup profiles. Thus it appears that while the second domain is well conserved within a subgroup, it shows variation between subgroups. Hexon carries a complex array of antigenic determinants, none of which have been accurately mapped or characterised. A common group-specific determinant has been demonstrated and is located in the base, while a number of type-specific determinants, some of which are neutralising, are thought to be predominantly located at the surface. In addition to these, inter- and intra- subgroup sites have been demonstrated. The high degree of variability observed within the tower region of hexons from different serotypes is a strong indication that it is subject to immunological pressure. Since the greatest diversity is found in loops I_1 and I_2, it seems reasonable to suggest that these contain the type-specific

immunoreactive sites. However it would be naive to consider these sites to be linear, it being far more likely that, due to the complex interweaving of the loops, each epitope would be formed by a number of widely spaced regions of the polypeptide. This rules out a prediction of the precise location of epitopes based solely on sequence variation. However, it could be suggested from the observations made following the pairwise comparisons, that the second half of loop2 which displays high homology within but not between subgroups, contains a component of the intra-group specific epitope.

In conclusion, it can be said that while certain predictions are possible regarding the structure of the hexons of these enteric adenoviruses, accurate models can only be obtained by X-ray crystallography. This, of course, depends on the production of sufficient protein for crystallisation and this task is now being undertaken. Both Ad2 and Ad5 proteins are vastly over-expressed during infection, and are therefore reasonably easy to purify from infected cell extracts. Unfortunately this is not the case with Ad40 and 41, and since growth in tissue culture is not particularly prolific, the use of expression systems is currently under investigation. The elucidation of the enteric adenoviruses' hexon structures may shed light on the location of the specific immunoreactive sites, however these can only be properly characterised by mapping the binding sites of monoclonal antibodies (MAbs). Dr H. van der Avoort at the Rijksinstituute voor Voksgezondheiden, Bilthoven, Netherlands has raised three Ad41-specific neutralising MAbs which exhibit strain specificity (personal communication). MAb 5-8 neutralises prototype strain 79-18025, and variant

strains 81-2378 and 83-17952, MAb 1-23 neutralises only the first two and MAb 3-10 neutralises just the prototype strain. The non-reactivity of the Ad41 variants with MAb 1-23 runs parallel with the loss of a *HindIII* site at 56% of the genome, while the non-reactivity of MAb 3-10 coincides with the appearance of an additional *EcoRI* site at 54%. These restriction sites are predicted to be in the hexon gene, the position for the *HindIII* mutation suggested by Dr van der Avoort at 56%, according to figure 3, should be fairly central. The sequence of the Ad41 hexon gene reveals two *HindIII* sites, one at approximately 1kb and the other at 2.4kb, neither of which could be said to be central. This anomaly is probably due to the fact that, since the entire sequence of the Ad41 genome has not been determined, the locating of genes in terms of percentage is a fairly arbitary business. Although it is not possible here to determine which of the two *HindIII* sites coincide with the reactivity of MAb 1-23, it is interesting to note that both fall in regions which correspond to protein sequences which are identical across all four serotypes and which are outwith the loops. This could imply that the type-specific epitopes are not exclusively located on the virus surface. However, it is also possible that the variants possess a number of mutations in the hexon gene, and the appearance or disappearance of these restriction sites bears no relation to the location of the epitope. It would be of considerable interest to determine the sequence of the hexon genes of these variants, so that the recognition sites of the MAbs may be further defined. The generation of mutant viruses which escape neutralisation by type-specific hexon monoclonal antibodies presents another approach for studying epitopes. Sequencing of the hexon genes of the ensuing mutants would thus

reveal the amino acids which contribute towards the formation of the epitope. Since the majority of sites are expected to be located in the variable regions, it would be possible to sequence only selected portions of the hexon gene using specific primers, hence increasing the number of mutants which one could sensibly study by reducing the tedious task of sequencing the entire gene. Antigenicity remains an area of adenovirus research yet to be extensively explored and could be particularly valuable towards determining the evolutionary history of these viruses.

Examination of the regions flanking the hexon gene reveal the location of sequences encoding portions of two other proteins. Upstream is the extreme 3' end of the gene which encodes the precursor to polypeptide VI, while downstream is the majority of the gene which codes for the proposed adenovirus protease. In Ad2, polypeptide VI is produced following proteolytic cleavage of a 249 amino acid precursor which loses 32 amino acids from the N-terminus during virus maturation. Although pVI is believed to be hexon-associated and possibly to provide a link between the capsid and the core, the functional domains have not yet been determined. Thus little can be said about the small amount of sequence obtained for Ad41, except that pVI appears to be coded for and that exhibits some homology to the Ad2 protein.

Remarkably high homology is observed between the Ad2 23k protease and the predicted Ad41 sequence, and this is maintained throughout the length of the protein. The protease is responsible for the maturation of a number of internal adenovirus proteins during assembly and although it has been

characterised as most like a cysteine protease, the location of the active site has not been identified. Since virus lacking functional protease activity is non-infectious, it must be assumed that both enteric serotypes possess a functional protein and therefore that the active site is conserved. The complete amino acid sequences of the Ad40 and 41 proteases have subsequently been determined and published (Vos *et al*, 1988), and are given as 205 and 214 residues respectively, longer than the 204 residues of Ad2. Thus the C-terminal 16 residues are absent from the sequence given in figure 14b. Vos *et al* (1989) also report a high level of amino acid conservation between the Ad41 and Ad2 proteases, with the exception of the extreme C-termini, the Ad41 protein being 10 amino acids longer. The published Ad41 sequence is identical to that reported here.

This information adds to the list of proteins which appear to be coded for by the enteric adenoviruses, and which together imply that the overall genome organisation and viral architecture is similar to that of Ad2 and Ad5. Hexon, fibre, pVI, pVIII, the 33k and 100k polypeptides, the protease and the single-stranded DNA binding protein have all been identified, and at least partially sequenced. From this information it is difficult to get any clue towards the cause of the unusual growth characteristics of the subgroup F viruses, the only obvious structural differences being the absence of the acidic domain in the hexon, and the possibility that they may possess heteromeric fibres (Pieniazek *et al*, 1990). It would seem much more likely that their unusual characteristics are, at least in part, due to peculiarities in the E1 coding region,

particularly in E1b. Ad40 has been shown to be defective in this region during infection *in vitro* (Mautner *et al*, 1989 and 1990), and in rodent cells transformed by Ad40 no E1b transcripts can be found (van Loon *et al*, 1987a). In contrast, Ad41 transformed rodent cells produce at least the large transcript from E1b but appear to lack two of the three E1a mRNA's (van Loon *et al*, 1987a). Whether these observations reflect the *in vivo* situation has not yet been established, since the cell type which supports replication of the enteric adenoviruses within the gastrointestinal tract has not been precisely defined. However, from the high titre of enteric adenoviruses present in the gut during infection it can be concluded that they replicate extremely efficiently, thus it has been suggested (Mautner *et al*, 1990) that the natural target cells either possess specific factors reponsible for E1 activation or, like the Ad5 transformed 293 cell line, constitutively express functions which compensate for defects in E1. More information is required for both serotypes regarding their proteins, mode of infection, host-cell specificity and the involvement of cellular factors, before the reason for their fastidious growth characteristics may be ascertained.

SECTION 2: **The use of oligonucleotide probes for the detection of adenovirus type 40 and 41.**

Identifying the presence of DNA of interest using radioactive oligonucleotide probes has become a successful and widely used technique, not least in the field of diagnosis. It does however, require at least some knowledge of the nucleotide sequence of the DNA to be detected, before specific probes can be made. The determination of the entire coding sequence of the Ad40 and 41 hexon genes has presented a plethora of information from which to design probes. The use of the hexon gene as a source of information has major advantages over other regions of the genome. Due to the number of different serotypes for which the nucleotide sequence of the hexon gene is known, it is easy to predict regions of the gene likely to be unique to each virus. Certain sequences corresponding to the surface loops are extremely unlikely to be duplicated in any other adenovirus serotype, and thus provide an ample source of information for designing type-specific probes. However, due to the high variability in these regions it is also possible that certain strains within a serotype may contain mutations which would render them unable to hybridise to a particular probe. Taking this into account it may not be advantageous to select probes from the regions subject to the greatest variation, unless strain variation is to be studied. Hexon also contains regions which are not thought to be susceptible to immunological pressure and are therefore less prone to strain-specific mutations, thus unique sequences identified within these conserved areas are more likely to detect all strains within a serotype. Three probes were thus derived, Ad41/1 from a variable region corresponding to loop 1, and Ad41/2 and Ad40/1 corresponding to the conserved C-terminal domains. In dot-blots of plasmids containing hexon fragments from different

serotypes each probe was found to be specific for the virus from which it was derived. End-labelling of these oligonucleotides with ^{32}P, generated probes which could detect hexon-containing plasmid DNA down to a level of 100pg. In contrast, a nick-translated Ad2 hexon fragment could detect plasmid DNA down to a level of 1pg, and although it displayed greatest sensitivity against the Ad2 hexon-containing plasmid, it also hybridised strongly with pCT3 and pCT40. Since pCT6 lacks the region of the hexon gene which is homologous to this probe, it did not bind to this fragment. The dot-blots shown in figure 18 contain DNA from representatives of subgroup A, C, E, F and of a non-human serotype (types 12, 2 and 5, 4, 40 and 41, and simian adenovirus 7, respectively). Unfortunately no DNA was available from subgroup B or D members. Oligonucleotides Ad41/1 and Ad40/1 were specific for their respective serotypes and showed no cross-reactivity with any of the other viruses. It would obviously be necessary to test all 39 non-subgroup F serotypes before being able to establish complete specificity, however given that none of the above types are recognised it is highly likely that this reflects the overall situation. The likelihood of completely unrelated gut pathogens containing identical sequences to those of the probes, and therefore interfering with diagnosis in clinical samples, was assumed to be negligible. The results obtained from dot-blots of plasmid and purified genomic DNA demonstrate the increased specificity of oligonucleotides probes compared to large DNA fragments, however it is also clear that this is at the expense of sensitivity.

To ascertain whether the sensitivity of these short probes would be sufficiently high to reveal Ad40 or 41 DNA in clinical samples, a series of stool

extracts and infected cell extracts were analysed. The first set of samples to be analysed were infected cell extracts, the virus originating from stool samples and passaged twice. All but one had been typed as Ad41, the untyped sample showed no cpe and did not produce sufficient DNA for restriction enzyme analysis. Ad41/2 was used in preference to Ad41/1 due to the greater possibility of certain strains avoiding detection with the latter. With this probe six definite (1, 3, 4, 5, 6, 9) and two possible (10, 11) positives were revealed. Using the Ad2 probe the same six samples were strongly positive, as was one of the possibles (11) indicated by Ad41/2 and sample 8 which had not been previously detected. Sample 10 did not hybridise, and indeed the distribution of radioactivity on the spot on the Ad41/2 autoradiograph was more suggestive of non-specific background binding than a specific interaction. A negative result was also obtained for sample 2. The presence of adenovirus in samples 2 and 10 had been revealed by EM, and, although they grew well enough in tissue culture to yield DNA for restriction enzyme analysis, they did not appear to hybridise to either of these probes. It is certainly possible that Ad41/1 may fail to hybridise due to strain variability, however it is totally inconceivable that the region which hybridises to the Ad2 probe is deleted. Thus the likely explanation for being unable to demonstrate the presence of any adenovirus DNA these samples, is either that the DNA had been degraded by nucleases, or that the level was below that detectable with this test. Sample 8 hybridised only to the Ad2 probe suggesting either that it was a variant not able to hybridise to Ad41/1, or that it had been wrongly typed by restriction enzyme analysis. Sample 7 was supplied untyped due to a failure of any virus

to grow and, since it gave no signal even with the more sensitive Ad2 probe, remains uncharacterised. Thus Ad41/2 has proved a perfectly adequate diagnostic tool when viruses from stool samples are amplified in tissue culture, however this is no more than should be expected.

For the routine screening of the number of samples which commonly pass through a diagnostic laboratory, it would obviously be advantageous to avoid unnecessary culturing. Thus the performance of Ad41/2 was assessed directly on stool extracts. Of the eleven Ad41 clinical samples supplied by Prof. Desselberger, three were positive (2, 8 and 11) using Ad41/2, however a further two (9 and 10) were found positive with the Ad2 probe. It should be noted that samples 9 and 10 appear to contain less DNA than the other three, and since the signal obtained with Ad41/2 was fairly weak, it can be assumed that they were not detected by the oligonucleotide probe due to a low level of DNA rather than a mutation in sequence.

A series of stool extracts supplied by Dr Kidd were hybridised to all three oligonucleotide probes and to the Ad2 probe. Six of the seven Ad40 samples were positive with the Ad2 probe and it was assumed that FB273 contained insufficient DNA for detection. Four of the six samples positive with the Ad2 probe were diagnosed as Ad40 with Ad40/1, and of the two that didn't hybridise, one (FB363) was indicated to contain very little DNA with the Ad2 probe and can therefore be assumed to be beyond detection with an oligonucleotide. N5234 did not hybridise to Ad40/1 despite the presence of a reasonable amount of DNA and although the reason for this is not certain, it

could be that in spite of this probe having been designed from what was hoped to be a conserved region, the relevant sequence contains a mutation in this strain. All of the six Ad41 stool extracts were shown to contain adenovirus DNA with the Ad2 probe, and were confirmed as Ad41 with both Ad41/1 and Ad41/2. It should be noted that FB143 appeared to hybridise more strongly to Ad41/2 than to Ad41/1 and this again could indicate some strain variation in nucleotide sequence. In contrast to the Ad40/1 probe, Ad41/1 is derived from the variable loop region where mutations are likely to be much more common, thus it may be expected not to recognise every strain equally.

It is regrettable that since a limited quantity of each clinical sample was available it was only possible to perform the dot-blot analyses once. Obviously it is difficult to make any firm conclusions from the results here, as there may be a degree of variability when experiments like these are repeated. This is particularly important to bear in mind when trying to make claims regarding the detection of strain variation with oligonucleotide probes. However, the results as they stand here would suggest that differences in target sequence have been detected in Ad41 strain FB143 and in Ad40 strain FB5234 using this set of probes. Another observation is that the oligonucleotides were more successful at diagnosing the presence of virus in the samples supplied by Dr Kidd than in those supplied by Prof. Desselberger. This may be a reflection of the stage of infection at which the samples were obtained, however no information is available regarding this. With this number of samples it is not easy to assess the suitability of these radioactive oligonucleotide probes to the diagnosis of enteric adenovirus infection. The studies certainly show that the

probes can diagnose infection in crude stool extracts, however it is likely that if this were the only method to be used, some positives would be missed. The samples which were typed by restriction enzyme analysis following tissue culture have not been typed by any other technique so it is impossible to tell how they would stand up to other methods of diagnosis, and therefore to assess the relative merits of this approach. Similarly, the samples which were typed by ELISA have, to my knowledge, only previously been tested by this method. However it should be said that, since the ELISA test requires no prior tissue culture step, it is more applicable to use in a diagnostic laboratory than one involving radioactive probes. The major drawback to this is the lack of availability of specific antisera. In addition, the study of the spatial and temporal distribution of variants of any one serotype would require the production of a number of strain-specific monoclonal antibodies which is not an easy task. The use of specific oligonucleotides for the detection of enteric adenoviruses is not readily applicable to routine use in a hospital diagnostic laboratory for two main reasons. Firstly, the level of sensitivity obtained here would suggest that without culturing, a small percentage of positives would fail to be detected, and secondly, the use of radioactivity makes the technique relatively expensive and certainly hazardous. However it should be noted that techniques generally available at present such as latex agglutination, are also prone to insensitivity. The likelihood of the diagnosis of enteric adenovirus infection becoming a routine hospital procedure is small, since the seriousness of the malady following hospitalisation is not particularly great, and diagnosis therefore purely academic. Instead, the majority of these diagnostic techniques are

adopted by research groups, particularly within diagnostic units, who are interested in studying sporadic outbreaks of infection, and in the evolutionary drift of these viruses. For such studies, the use of oligonucleotide probes would be of considerable benefit, particularly ones that could distinguish between strains. The hexon gene provides an ample source of information from which a battery of probes could be designed, however since each one would need to be assessed for strain-specificity, it is not a task to be undertaken lightly. This would be simplified, if more information became available on the location of type- and strain-specific antigenic determinants, thus revealing the sequences most likely to vary.

To make this approach more desirable for general use, the replacement of ^{32}P with a non-radioactive tag would be extremely advantageous and at this point it would be appropriate to mention work performed by Peter Bates, an undergraduate student in this laboratory who, for his Senior Honours project, attempted to develop a non-radioactive probing system. His approach is discussed below with only brief references to the details of the methods used. Each oligonucleotide was resynthesised to incorporate an Aminolink™ at the 5' end. This constitutes an active amine group coupled to a 2-carbon spacer arm and provides a means for chemically crosslinking oligonucleotides to proteins or certain haptens. The Aminolink oligonucleotides were subsequently coupled to a biotin derivative (N-hydroxysuccinimido biotin), to yield biotinylated probes. Dot-blots of plasmid DNA similar to those described in figure 17, were probed with the biotinylated oligonucleotides, which were in turn detected with streptavidin-conjugated horseradish peroxidase (HRP) in the

presence of hydrogen peroxidase and diaminobenzidine. This method gave a maximum sensitivity of 10ng plasmid DNA. Gold enhancement of the signal was performed, however was found not to increase the sensitivity, although it did increase signal contrast. Since this level of sensitivity was far below that of the radioactive probes, attempts were made to improve upon it. Several workers have reported alkaline phosphatase (AP) to be a more sensitive detector enzyme than HRP (e.g. Matthews and Kricka, 1988). With this in mind, the dot-blots were repeated substituting streptavidin and biotinylated AP for the streptavidin-conjugated HRP. In addition to the biotinylated monomeric AP, a polymeric form was synthesised with the idea that this would yield a higher signal for each molecule that bound. Although the sensitivity obtained with the polymeric AP was threefold greater than with the monomeric form, it still was not as high as when HRP was used as a detector. To further amplify the signal, attempts were made to build up larger enzyme complexes on the filter by performing repeated washes with streptavidin and biotinylated AP, however despite being slightly more sensitive this was found to increase background interference. Biotinylated ß-galactosidase was substituted for AP, again using a streptavidin bridge, however this approach was the least sensitive, requiring 300ng plasmid DNA to obtain a signal. From this limited study, it would appear that the substitution of a single biotin molecule for ^{32}P as a method of labelling, produces probes which are far too insensitive for use with clinical samples and would probably fail to detect virus even in infected cell extracts. However there remain several avenues to explore in trying to improve probe sensitivity. The incorporation of several biotins per molecule of probe could serve to increase

the signal, however it is also feasible that this would interfere with binding due to steric hindrance. Longer spacer arms are available which may overcome this problem. Rather than trying to increase the signal-generating capacity of the probe itself, an increase in the sensitivity of the detector system would seem less fraught with problems. Such was the reasoning behind the attempt to build large streptavidin-biotinylated AP complexes. Although Peter Bates was not successful at this in his brief attempt, with a lot more time available to optimise conditions this method could form the basis of a reasonably sensitive test. Another approach worth exploring is the use of enhanced-chemiluminescence which is reputed to be at least as sensitive as radioactive labelling. This requires the presence of horseradish peroxidase which could either be directly conjugated to the oligonucleotide, or could be incorporated into a biotinylated probe complex via streptavidin in a similar manner to that described above.

One of the reasons for considering the development of non-radioactive labelling techniques was with a view to producing probes suitable for use in developing countries. In the West where we enjoy the luxury of good health care, the risk of serious illness due to enteric adenovirus infection is minimal. However in developing countries the situation is a lot more serious, with a huge number of infant mortalities due to viral diarrhoea every year. In this sort of environment, the ability to rapidly diagnose adenovirus infection would be advantageous both in determining a suitable treatment for the patient, and in the study and management of epidemics. Several factors must be considered when designing a diagnostic test: the cost, whether the neccessary equipment

is readily available in Third World countries and whether the reagents are stable. All these criteria can be met by non-radioactively labelled oligonucleotide probes. Oligonucleotides can be made in very large quantities reasonably cheaply, and the requirement for specialised equipment for their synthesis can be circumvented by ordering from commercial companies. Once obtained these can be conjugated to biotin or a reported enzyme and when stored in a lyophilised form, are stable. This kind of test requires few chemicals and no specialised equipment, providing that it is sensitive enough to be performed directly on clinical samples. The results obtained so far indicate that a lot more work needs to be done to improve the sensitivity of these probes for diagnosing Ad40 and 41 infection, however the sensitivities reported to have been obtained by other groups in similar tests suggest that this approach could be successful.

The polymerase chain reaction has recently been shown to be extremely successful in diagnosing adenovirus infection using clinical specimens (Allard *et al*, 1990). In this study PCR was compared to the more conventional methods of latex agglutination, restriction enzyme digest and cytopathicity, and found to be by far the most sensitive. Like dot-blot hybridisation using oligonucleotide probes, PCR requires prior knowledge of nucleotide sequence to facilitate design of specific primers. Depending on the sequences chosen, primers may be group-, subgroup-, type- or strain-specific. In the study by Allard, a group-specific region was identified within the P1 ß-barrel of the hexon. Primers corresponding to this region were tested for their

ability to amplify DNA from 18 different serotypes. and found to be successful in each case. Subgroup F- and type-specific sequences were identified within the Ad40 and 41 E1a and E1b regions, and primers derived from here were found to be able to distinguish the enteric adenoviruses from other serotypes, and from each other.

The oligonucleotides derived from the Ad40 and 41 hexon genes which form the basis of the dot-blot assay reported here, were shown to be type-specific. Ad41/1 and Ad41/2 correspond to the same strand of the Ad41 genome and are thus not suitable for use in a PCR reaction. However, were the complimentary sequence to be used in place of Ad41/2, this pair of primers would amplifiy a specific region of the Ad41 hexon gene. Since Ad41/1 was designed from a type-specific region of the hexon gene, these primers should provide a means of distinguishing Ad41 DNA from that of other serotypes in clinical samples using the PCR technique. As well as containing information for the design of a multitude of type-specific primers, the hexon gene also has the potential for providing strain-specific sequences. Hence PCR could be useful in the study of the antigenic drift of these viruses. As well as being able to amplify specific sequences in stool samples, PCR can be performed on water supplies and food sources which would be of considerable benefit in monitoring and controlling the spread of epidemics.

In view of the enormous sensitivity of PCR, coupled with the advantages of speed and low cost, it is likely that this technique will play a vital role in future developments in the field of diagnostic Virology.

REFERENCES

Adam, S, and Dreyfuss, G. (1987). Adenovirus proteins associated with mRNA and hnRNA in infected HeLa cells. Journal of Virology **61** 3276-3283.

Agrawal, S.; Christodoulou, C. and Gait, M.J. (1986). Efficient methods for attaching non-radioactive labels to the 5' ends of synthetic oligo-deoxyribonucleotides. Nucleic Acids Research **14** 6227-6245.

Akusjärvi, G.; Aleström, P.; Pettersson, M.; Lager, M.; Jörnvall, H. and Pettersson, U. (1984). The gene for the adenovirus 2 hexon polypeptide. Journal of Biological Chemistry **259** 13976-13979.

Albiges-Rizo, C. and Chroboczek, J. (1990) Adenovirus serotype 3 fibre protein is expressed as a trimer in *Escherichia Coli*. Journal of Molecular Biology **212** 247-252.

Allard, A.; Girones, R.; Juto, P. and Wadell, G. (1990). Polymerase chain reaction for detection of adenoviruses in stool samples. Journal of Clinical Microbiology **28** 2659-2667.

Anderson, C.W.; Young, M.E.; and Flint, S.J. (1989) Characterisation of the adenovirus 2 virion protein, Mu. Virology **172** 506-512.

Babiss, L.E. and Ginsberg, H.S. (1984). Adenovirus type 5 early region 1B gene product is required for efficient shut-off of host protein synthesis. Journal of Virology 50 202-212.

Babiss, L.E.; Ginsberg, H.S. and Darnell, J.E. (1985). Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport. Molecular and Cellular Biology **5**

2552-2558.

Biggin, M.D.; Gibson, T.J. and Hong, G.F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proceedings of the National Acadamy of Sciences, U.S.A. **80** 3963-3965.

Brandt, C.D.; Rodriguez, W.J.; Kim, H.W.; Arrobio, J.O.; Jeffries, B.C. and Parrott, R.H. (1984). Rapid presumptive recognition of diarrhoea-associated adenoviruses. Journal of Clinical Microbiology **20** 1008-1009.

Brandt, C.D.; Kim, H.W.; Rodriguez, W.J.; Arrobio, J.O.; Jeffries, B.C.; Stalling, E.P.; Lewis, C.; Miles, A.J.; Gardner, M.K. and Parrott, R.H. (1985). Adenovirus and paediatric gastroenteritis. Journal of Infectious Diseases **151** 437-443.

Bridge, E. and Ketner, G. (1990). Interaction of viral E4 and E1b products in late gene expression. Virology **174** 345-353.

Brown, M.; Petric, M. and Middleton, P.J. (1984). Diagnosis of fastidious enteric adenoviruses 40 and 41 in stool specimens. Journal of Clinical Microbiology **20** 334-338.

Burnett. R.M. (1985). The structure of the adenovirus capsid. II. The packing symmetry of hexon and its implications for viral architecture. Journal of Molecular Biology, **185** 125-143.

Cepko, C.L. and Sharp, P.A. (1982). Assembly of adenovirus major capsid protein is mediated by a non virion protein. Cell **31** 407-415.

Cepko, C.L.; Whetstone, C.A. and Sharp, P.A. (1983). Adenovirus hexon monoclonal antibody that is group specific and potentially useful as a

diagnostic reagent. Journal of Clinical Microbiology **17** 360-364.

Colby, W.W. and Shenk, T. (1981). Adenovirus type 5 virions can be assembled *in vivo* in the absence of detectable polypeptide IX. Journal of Virology **39** 977-980.

de Jong, J.C.; Wigand, R.; Kidd, A.H.; Wadell, G.; Kapsenberg, J.G.; Muzeric, C.J.; Wermenbol, A.G. and Firtzlaff, R.G. (1983). Candidate adenoviruses 40 and 41: fastidious adenoviruses from human infant stool. Journal of Medical Virology **11** 215-231.

Devaux, C; Zulauf, M.; Boulanger, P. and Jacrot, B. (1982). Molecular weight of adenovirus serotype 2 capsomers: a new characterisation. Journal of Molecular Biology **156** 927-939.

Devaux, C; Adrian, M; Berthet-Colominas, C; Cusack, S. and Jacrot, B. (1990). Structure of adenovirus fibre. I. Analysis of crystals of fibre from adenovirus serotypes 2 and 5 by electron microscopy and X-ray crystallography. Journal of Molecular Biology **215** 567-588.

Dowling, J.M. and Wynne, H. (1981). Role of enteric adenoviruses and rotaviruses in infantile gastroenteritis. Lancet **ii** 305-306.

Emini, E.A.; Ostapchuk, P. and Wimmer, E. (1983). Bivalent attachment of antibody onto poliovirus leads to conformational alteration and neutralisation. Journal of Virology **48** 547-550.

Everitt, E.; Sundquist, B.; Pettersson, U. and Philipson, L. (1973). Structural proteins of adenoviruses. X. Isolation and topography of low molecular weight antigens from the virion of adenovirus type 2. Virology **52** 130-147.

Everitt, E.; Lutter, L. and Philipson, L. (1975). Structural proteins of adenovirus. XII. Location and neighbour relationaship among proteins of adenovirion type 2 as revealed by enzymatic iodination, immunoprecipitation and chemical cross-linking. Virology 67 197-208.

Everitt, E.; Persson, M.J. and Wohlfart, C. (1988). pH-dependant exposure of endoproteolytic sites on the adenovirus 2 hexon protein. FEMS Microbiology Letters **49** 229-233.

Flewett, T.H.; Burden, A.S.; Davis, H. and Morris, C.A. (1975). Epidemic viral enteritis in a long-stay children's ward. Lancet **i** 4-5.

Forster, A.C.; McInnes, J.L.; Skingle, D.C. and Symons, R.H. (1985). Non-radioactive hybridisation probes prepared by the chemical labelling of DNA and RNA with a novel reagent, photobiotin. Nucleic Acids Research **13** 745-761.

Fukui, Y.; Shiroki, K.; Saito,I. and Shimojo, H. (1984). Characterisation of a host-range mutant of human adenovirus 12 defective in early region 1B. Journal of Virology **50** 132-136.

Furcinitti, P.S.; van Oostrum, J.; Wall, J.S. and Burnett, R.M. (1987). Visualization of a small polypeptide within an adenovirus capsid fragment by subtracting the hexon X-ray structure from STEM images. Proceedings of the 45th Annual Meeting of the Electron Microscope Society of America. San Fransisco Press. 738-739.

Garnier, J.; Osguthorpe, D.J. and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. Journal of Molecular Biology **120** 97-120.

Gary, G.W.; Hierholzer, J.C. and Black, R.E. (1979). Characteristics of non-cultivatable adenoviruses associated with diarrhoea in infants: a new subgroup of human adenoviruses. Journal of Clinical Microbiology **10** 96-103.

Ginsberg, H.S. (1979). Adenovirus structural proteins. In "Comparative Virology" (H. Fraenkel-Conrat and R.R. Wagner, Eds.) **vol 13** 409-457. Plenum, New York.

Ginsberg, H.S. (1984). The Adenoviruses. Plenum Press, New York.

Ginsberg, H.S.; Pereira, H.G.; Valentine, R.C. and Wilcox, W.C. (1966). A proposed terminology for the adenovirus antigens and virion morphological subunits. Virology **28** 782-783.

Graham, F.L.; Smiley, J.; Russell, W.C. and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. Journal of General Virology **36** 59-72.

Grandlen, M.; Pettersson, C-A.; Svensson, L. and Uhnoo, I. (1987). Latex agglutination test for adenovirus diagnosis in diarrhoeal disease. Journal of Medical Virology **23** 311-316.

Green, N.M.; Wrigley, N.G.; Russell, W.C.; Martin, S.R. and McLachlan, A.D. (1983). Evidence for a repeating cross-ß sheet structure in the adenovirus fibre. EMBO Journal **2** 1357-1365.

Grütter, M. and Franklin, R.M. (1974). Studies on the molecular weight of the adenovirus type 2 hexon and its subunit. Journal of Molecular Biology **89** 163-178.

Halonen, P.; Sarkkinen, H.; Arstila, P.; Hjertsson, E. and Torfason,

E. (1980). Four-layer radioimmunoassay for detection of adenovirus in stool. Journal of Clinical Microbiology **11** 614-617.

Hammond, G.W.; Mauthe, G.; Joshua, J. and Hannan, C.K. (1985). Examination of uncommon clinical isolates of human adenoviruses by restriction endonuclease analysis. Journal of Clinical Microbiology **21** 611-616.

Hammond, G.; Hannan, C.; Yeh, T.; Fischer, K.; Mauthe, G. and Straus, S.E. (1987). DNA hybridisation for diagnosis of enteric adenovirus infection from directly spotted human faecal specimens. Journal of Clinical Microbiology **25** 1881-1885.

Hanahan, D. (1983). Studies on transformation of Escherichia Coli with plasmids. Journal of Molecular Biology **166** 557-580.

Hay, R.T.; Stow, N.D. and McDougall, I.M. (1984). Replication of adenovirus mini-chromosomes. Journal of Molecular Biology **175** 493-510.

Hayes, B.W.; Telling, G.C.; Myat, M.M.; Williams, J.F. and Flint, S.J. (1990). The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species. Journal of Virology **64** 2732-2742.

Herrmann, J.E.; Perron-Henry, D.M.; Stobbs-Walro, D. and Blacklow, N.R. (1987). Preparation and characterisation of monoclonal antibodies to enteric adenovirus types 40 and 41. Archives of Virology **94** 259-265.

Hogle, J.M.; Chow, M. and Filman, D.J. (1985). Three-dimensional structure of poliovirus at 2.9Å resolution. Science **229** 1358-1365.

Horne. R.W.; Brenner, S.; Waterson, A.P. and Wildy, P. (1959). The

icosahedral form of an adenovirus. Journal of Molecular Biology **1** 84-86.

Hu, S-L.; Hays, W.W. and Potts, D.E. (1984). Sequence homology between bovine and human adenoviruses. Journal of Virology **49** 604-608.

Ishino,M.; Ohashi, Y.; Emoto, T.; Sawada, Y. and Fujinaga, K. (1988). Characterization of the adenovirus type 40 E1 region. Virology **165** 95-102.

Jacobsson, P.A.; Johansson, M.E. and Wadell, G. (1979). Identification of an enteric adenovirus by immunoelectro-osmophoresis (IEOP) technique. Journal of Medical Virology **3** 307-312.

Jeffreys, A.J.; Wilson, V. and Thein, S.L. (1985). Hypervariable 'minisatellite' regions in human DNA. Nature, London **314** 67-73.

Johansson, M.E.; Uhnoo, I.; Kidd, A.H.; Madeley, C.R. and Wadell, G. (1980). Direct identification of enteric adenovirus, a candidate new serotype, associated with infantile gastroenteritis. Journal of Clinical Microbiology **12** 95-100.

Johansson, M.E.; Uhnoo, I.; Svensson, L.; Pettersson, C-A. and Wadell, G.(1985). Enzyme-linked immunosorbent assay for detection of enteric adenovirus 41. Journal of Medical Virology **17** 19-27.

Kidd, A.H. (1984). Genome variants of adenovirus 41 (subgroup G) from children with diarrhoea in South Africa. Journal of Medical Virology **14** 49-59.

Kidd, A.H.; Berkowitz, F.E.; Blaskovic, P.J. and Schoub, B.D. (1984). Genome variants of human adenovirus 40 (subgroup F). Journal of Medical Virology **14** 235-246.

Kidd, A.H.; Harley, E.H. and Erasmus, M.J. (1985). Specific detection and typing of adenovirus types 40 and 41 in stool specimens by dot-blot hybridisation. Journal of Clinical Microbiology **22** 934-939.

Kidd, A.H. and Erasmus, M.J. (1989). Sequence characterisation of the adenovirus 40 fibre gene. Virology **172** 134-144.

Kinloch, R.; Mackay, N. and Mautner, V. (1984). Adenovirus hexon. Sequence comparison of subgroup C serotypes 2 and 5. Journal of Biological Chemistry **259** 6431-6436.

Kjellén, L. and Pereira, H.G. (1968). Role of adenovirus antigens in the induction of virus neutralising antibody. Journal of General Virology **2** 177-185.

Leite, J.P.G.; Pereira, H.G.; Azeredo, R.S. and Schatzmayr, H.G. (1985). Adenoviruses in faeces of children with acute gastroenteritis in Rio de Janeiro, Brazil. Journal of Medical Virology **15** 203-209.

Leith, I.R.; Hay, R.T. and Russell, W.C. (1989). Adenovirus subviral particles and cores can support limited DNA replication. Journal of General Virology **70** 3235-3248.

Liljas, L. (1986). The structure of spherical viruses. Progress in Biophysics and Molecular Biology **48** 1-36.

Madeley, C.R.; Cosgrave, B.P.; Bell, E.J. and Fallon, R.J. (1977). Stool viruses in babies in Glasgow. 1. Hospital admissions with diarrhoea. Journal of Hygiene **78** 261-273.

Matthews, J.A. and Kricka, L.J. (1988). Analytical strategies for the use of DNA probes. Analytical Biochemistry **169** 1-25.

Mautner, V.; Mackay, N. and Steinthorsdottir, V. (1989). Complementation of enteric adenovirus type 40 for lytic growth in tissue culture by E1B 55K function of adenovirus types 5 and 12. Virology **171** 619-622.

Mautner, V.; Mackay, N. and Morris, K. (1990). Enteric adenovirus type 40: expression of E1B mRNA and proteins in permissive and non-permissive cells. Virology **179** 129-138.

Messing, J.; Crea, R. and Seeburg, P.H. (1981). A system for shotgun sequencing. Nucleic Acids Research **9** 309-321.

Morin, N. and Boulanger, P. (1986). Hexon trimerization occurring in an assembly-defective, 100k temperature-sensitive mutant of adenovirus 2. Virology **152** 11-31.

Neumann, R.; Chroboczek, J. and Jacrot, B. (1988). Determination of the nucleotide sequence for the penton-base gene of human adenovirus type 5. Gene **69** 153-157.

Niel, C.; Gomes, S.A.; Leite, J.P.G. and Pereira, H.G. (1986). Direct detection and differentiation of fastidious and non-fastidious adenoviruses in stools by using a specific non-radioactive probe. Journal of Clinical Microbiology **24** 785-789.

Norrby, E. and Skaaret, P. (1967). The relationship between the soluble antigens and the virion of adenovirus type 3. III. Immunological identification of fibre antigen and isolated vertex capsomer antigen. Virology **32** 489-502.

Norrby, E. (1969). The relationship between the soluble antigens and the virion of adenovirus type 3. IV. Immunological characteristics. Virology **37**

565-576.

Norrby, E. and Wadell, G. (1969). Immunological relationship between hexons of certain human adenoviruses. Journal of Virology **4** 663-670.

Olson, A.J.; Bricogne, G. and Harrison, S.C. (1983). Structure of tomato bushy stunt virus. IV. The virus particle at 2.9Å resolution. Journal of Molecular Biology **171** 61-93.

Pereira, H.G.and Wrigley, N.G. (1974). *In vitro* reconstitution, hexon bonding and handedness of incomplete adenovirus capsid. Journal of Molecular Biology **85** 617-631.

Persson, H.; Mathisen, B.; Philipson, L. and Pettersson, U. (1979). A maturation protein in adenovirus morphogenesis. Virology **93** 198-208.

Pettersson, U. (1971). Structural proteins of adenoviruses. VI. On the antigenic determinants of the hexon. Virology **43** 123-136.

Pettersson, U. and Hyppia, T. (1985). Nucleic acid hybridisation- an alternative tool in diagnostic Microbiology. Immunology Today **6** 268-272.

Pieniazek, N.J.; Slemenda, S.B.; Pieniazek, D.; Velarde, J. and Luftig, R.B. (1989).Sequence of human enteric adenovirus type 41 Tak fibre protein gene. Nucleic Acids Research **17** 9474.

Pieniazek, N.J.; Slemenda, S.B.; Pieniazek, D.; Velarde, J. and Luftig, R.B. (1990). Human enteric adenovirus type 41 (Tak) contains a second fibre protein gene. Nucleic Acids Research **18** 1901.

Pieniazek, N.J.; Velarde, J.; Pieniazek, D. and Luftig, R.B. (1990). Nucleotide sequence of the region coding for the 100k and 33k proteins of human enteric adenovirus type 41 (Tak). Nucleic Acids Research **18**

3069.

Possee, R.D.; Schild, G.C. and Dimmock, N.J. (1982) Studies on the mechanism of neutralisation of influenza virus by antibody: evidence that neutralising antibody (anti-haemagglutinin) inactivates influenza virus in vivo by inhibiting virion transcriptase activity. Journal of General Virology **58** 373-386.

Retter, M.; Middleton, P.J.; Tam, J.S. and Petric, M. (1979). Enteric adenoviruses: detection, replication and significance. Journal of Clinical Microbiology **10** 574-578.

Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. Advances in Protein Chemistry **34** 167-339.

Rigby, P.W.J.; Dieckmann, M.; Rhodes, C. and Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. Journal of Molecular Biology **113** 237-251.

Roberts, M.M.; White, J.L.; Grütter, M.G. and Burnett, R.M. (1986). Three-dimensional structure of the adenovirus major coat protein hexon. Science **232** 1148-1151.

Rossmann, M.G.; Abad-Zapatero, C.and Murthy, M.R.N. (1983). Structural comparisons of some small spherical plant viruses. Journal of Molecular Biology **165** 711-736.

Rossmann, M.G.; Arnold, E.; Erickson, J.W.; Frankenberger, E.A.; Griffith, J.P.; Hecht, H-J.; Johnson, J.E.; Kamer, G.; Luo, M.; Mosser, A.G.; Rueckert, R.R.; Sherry, B. and Vriend, G. (1985).

Structure of a human common cold virus and functional relationship to other picornaviruses. Nature London **317** 145-153.

Ruigrok, R.W.; Barge, A.; Albiges-Rizo, C. and Dayan, S. (1990). Structure of adenovirus fibre. II. Morphology of single fibres. Journal of Molecular Biology **215** 589-596.

Russell, W.C.; McIntosh, K. and Skehel, J.J. (1971). The preparation and properties of adenovirus cores. Journal of General Virology **11** 35-46.

Russell, W.C.; Patel, G.; Precious, B.; Sharp, I. and Gardner, P.S. (1981). Monoclonal antibodies against adenovirus type 5: preparation and preliminary characterisation. Journal of General Virology **56** 393-405.

Russell, W.C. and Precious, B. (1982). Nucleic acid binding properties of adenovirus structural polypeptides. Journal of General Virology **63** 69-79.

Saiki, R.K.; Gelfand, D.H.; Stoffel, S.; Scharf, S.J.; Higudu, R.; Horn, G.T.; Mullis, K.B. and Ehlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239 487-491.

Sambrook, J.; Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratory Press. Second Edition.

Sanger,F.; Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, U.S.A. **74** 5463-5467.

Sanger, F.; Coulson, A.R.; Barrell, B.G.; Smith,A.J.H. and Roe,B.A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA

sequencing. Journal of Molecular Biology **143** 161-178.

Shiroki, K.; Ohshima, K.; Fukui, Y. and Ariga, H. (1986). The adenovirus type 12 early region 1B 58,000-Mr gene product is required for viral DNA synthesis and for initiation of cell transformation. Journal of Virology **57** 792-801.

Singh-Naz, N. and Naz, R.K. (1986).Development and application of monoclonal antibodies for specific detection of human enteric adenoviruses. Journal of Clinical Microbiology **23** 840-842.

Slemenda, S.B.; Pieniazek, N.J.; Velarde, J.; Pieniazek, D. and Luftig, R. (1990). Nucleotide sequence of the region coding for 100k and 33k proteins of human enteric adenovirus type 41 (Tak). Nucleic Acids Research **18** 3069.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology **98** 503-517.

Staden, R. (1982a). Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Research **10** 4731-4751.

Staden, R. (1982b). An interactive graphics programme for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Research **10** 2951-2961.

Subramanian, T.; Kuppuswamy, M.; Gysbers, J.; Mak, S. and Chinnadurai, G. (1984). 19kDa tumour antigen coded by early region E1B of adenovirus 2 is required for efficient synthesis and for protection of

viral DNA. Journal of Biological Chemistry **259** 11,777-11,783.

Svensson, L.; Wadell, G.; Uhnoo, I. Johansson, M. and von Bonsdorff, C-H. (1983). Cross-reactivity between enteric adenoviruses and adenovirus type 4: analysis of epitopes by solid-phase immune electron microscopy. Journal of General Virology **64** 2517-2520.

Svensson, U. (1985). Role of vesicles during adenovirus 2 internalisation into HeLa cells. Journal of Virology **55** 442-449.

Takiff, H.E.; Straus, S.E. and Garon, C.F. (1981). Propagation and in vitro studies of previously non-cultivatable enteral adenoviruses in 293 cells. Lancet **ii** 832-834.

Takiff, H.E.; Reinhold, W.; Garon, C.F. and Straus, S.E. (1984). Cloning and physical mapping of enteric adenoviruses (candidate types 40 and 41). Journal of Virology **51** 131-136.

Tang, J.Y. and Agrawal, S. (1990). Incorporation of multiple reporter groups on synthetic oligonucleotides. Nucleic Acids Research **18** 6461.

Uhnoo, I.; Wadell, G.; Svensson, L. and Johansson, M. (1984). Importance of enteric adenoviruses 40 and 41 in acute gastroenteritis in infants and young children. Journal of Clinical Microbiology **20** 365-372.

van Loon, A.E.; Rozijn, T.H.; de Jong, J.C. and Sussenbach, J.S. (1985). Physicochemical properties of the DNAs of the fastidious adenovirus species 40 and 41. Virology **140** 197-200.

van Loon, A.E.; Ligtenberg, M.; Reemst, A.M.C.B.; Sussenbach, J.S. and Rozijn, T.H. (1987a). Structure and organisation of the left-terminal DNA regions of fastidious adenovirus types 40 and 41. Gene

58 109-126.

van Loon, A.E.; Gilardi, P.; Perricaudet, M.; Rozijn, T.H. and Sussenbach, J.S. (1987b). Transcriptional activation by the E1A regions of adenovirus types 40 and 41. Virology **160** 305-307.

van Oostrum, J. and Burnett, R.M. (1985). Molecular composition of the adenovirus type 2 virion. Journal of Virology **56** 439-448.

van Oostrum, J.; Smith, P.R.; Mohraz, M. and Burnett, R.M. (1987). The structure of the adenovirus capsid. III. Hexon packing determined from electron micrographs of capsid fragments. Journal of Molecular Biology **198** 73-89.

Varga, M.J.; Bergman, T. and Everitt, E. (1990). Antibodies with specificities against a dispase-produced 15-kilodalton hexon fragment neutralise adenovirus type 2 infectivity. Journal of Virology **64** 4217-4225.

Vayda, M.E.; Rogers, A.E. and Flint, S.J. (1983). The structure of nucleoprotein cores released from adenovirions. Nucleic Acids Research **11** 441-460.

von Bahr-Lindström, H.; Jörnvall, H.; Althin, S. and Philipson, L. (1982). Structural differences between hexons from adenovirus types 2 and 5: correlation with difference in size and immunological properties. Virology **118** 353-362.

Vos, H.L.; van der Lee, F.M.; Reemst, A.M.C.B.; van Loon, A.E. and Sussenbach, J.S. (1988). The genes encoding the DNA binding protein and the 23K protease of adenovirus types 40 and 41. Virology **163** 1-10.

Wadell, G. and Norrby, E. (1969). Immunological and other biological

characteristics of pentons of human adenoviruses. Journal of Virology **4** 671-680.

Weber, J. (1976). Genetic analysis of adenovirus type 2. III. Temperature-sensitivity of processing of viral proteins. Journal of Virology **17** 462-471.

Webster, A.; Russell, W.C. and Kemp, G.D. (1989a). Characterisation of the adenovirus proteinase: Development and use of a specific peptide assay. Journal of General Virology **70** 3215-3223.

Webster, A.; Russell, S.; Talbot, P.; Russell, W.C. and Kemp, G.D. (1989b). Characterisation of the adenovirus proteinase: substrate specificity. Journal of General Virology **70** 3225-3234.

White, E; Denton, A. and Stillman, B. (1988). Role of adenovirus E1B 19,000-dalton tumour antigen in regulating early gene expression. Journal of Virology **62** 3445-3454.

Whitelaw, A.; Davies, H. and Parry, J. (1977). Electron microscopy of fatal adenovirus gastroenteritis. Lancet **i** 361.

Wilcox, W.C. and Ginsberg, H.S. (1963). Production of specific neutralising antibody with soluble antigens of type 5 adenovirus. Proceedings of Society for Experimental Medicine **114** 37-42.

Willcox, N. and Mautner, V. (1976). Antigenic determinants of adenovirus capsids. II. Homogeneity of hexons, and accessibility of their determinants, in the virion. Journal of Immunology **116** 25-29.

Williams, J.; Karger, B.D.; Ho, Y.S.; Castiglia, C.L.; Mann, T. and Flint, J. (1986). The adenovirus E1B 495R protein plays a role in

regulating the transport and stability of the viral late messages. Cancer Cells **4** 275-284.

Wohlfart, C. (1988). Neutralisation of adenoviruses: kinetics, stoichiometry and mechanisms. Journal of Virology **62** 2321-2328.

Wohlfart, C.E.G.; Svensson, U.K. and Everitt, E. (1985). Interaction between HeLa cells and adenovirus type 2 virions neutralised by different antisera. Journal of Virology **56** 896-903.

Wong, M-L. and Hsu, M-T. (1989).Linear adenovirus DNA is organised into supercoiled domains in virus particles. Nucleic Acids Research **17** 3535-3550.

Wood, D.J.; de Jong, J.C.; Bijlsma, K. and van der Avoort, H.G.A.M. (1989) Development and evaluation of monoclonal antibody-based immune electron microscopy for diagnosis of adenovirus types 40 and 41. Journal of Virological Methods **25** 241-250.