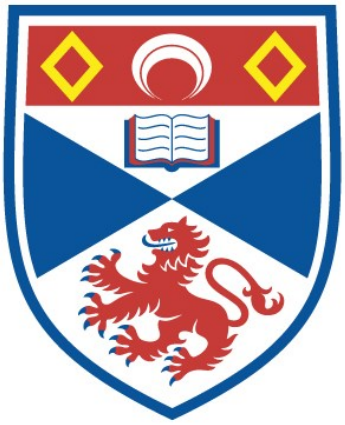
CHARACTERISATION OF THE ADENOVIRUS PROTEASE

Ailsa Webster

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1992

CHARACTERISATION OF THE ADENOVIRUS PROTEASE

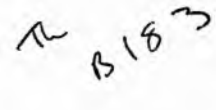
By Ailsa Webster

Department of Biochemistry and Microbiology

University of St. Andrews

A thesis presented for the degree of Doctor of Philosophy

at the University of St. Andrews, March 1992


B183

DECLARATION

(a) I, Ailsa Webster hereby certify that this thesis has been composed by myself, that it is a record of my work and that it has not been accepted in partial or complete fulfillment of any other degree of professional qualification.

Signed .. Date 27 March 92

(b) I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 1st October 1988 and as a candidate for the degree of Ph.D. on 1st October 1989.

Signed .. Date 27 March 92

(c) I hereby certify that the candidate has fulfilled the conditions of the Resolutions appropriate to the Degree of Ph.D.

Signature of Supervisor ..... Date 27 March 1992

(d) In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

ABBREVIATIONS

A	Adenine
AcNPV	Autographa Californica nuclear polyhedrosis virus
Ad	Adenovirus
AIDS	Acquired immunodeficiency syndrome
Amp	Ampicillin
AMV	Avian myeloblastosis virus
ASF	African swine fever (virus)
ASLV	Avian sarcoma leukosis virus
BSA	Bovine serum albumin
BVDV	Bovine viral diarrhoea virus
C	Cytosine
CIAP	Calf intestinal alkaline phosphatase
CPMV	Cow pea mosaic virus
CRB	Cambridge Research Biochemicals
DBP	DNA binding protein
DCI	3,4 dichloroisocoumarin
dCMP	Deoxycytosine monophosphate
dd	deionised distilled
DEAE	Diethyl-aminoethyl
DEN	Dengue (virus)
DFP	Diisopropylfluorophosphate
Dhbt	3,4 dihydro-4-oxobenzotriazinyl-3-yl
DMAP	Dimethylaminopyridine
DMF	Dimethylformamide
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
ds	Double stranded
DTDP	Dithiodipyridine
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EEO	Electroendosmosis
E1-E4	Early (genes or proteins) 1-4
eIF	Eukaryotic initiation factor
EM	Electron Microscopy
EMC	Encephalomyocarditis (virus)
FDNB	Dinitrofluorobenzene
FCS	Fetal calf serum
FMDV	Foot and mouth disease virus
FMOC	Fluorenylmethoxycarbonyl
FPLC	Fast Protein Liquid Chromatography
G	Guanine
GST	Glutathione-S-transferase
G-MEM	Glasgow modified Earle's minimal essential medium
GON	Group of nine
HA	Haemagglutinin
HIV	Human immunodeficiency virus
HOBt	Hydroxybenzotriazole

HPLC	High performance liquid chromatography
HSV	Herpes simplex virus
IPTG	Isopropyl-b-D-thiogalactopyranoside
L1-L5	Late (genes or proteins) 1-5
LB	Luria broth
LMP	Low melting point
MEM	Eagle's minimal essential medium
MLP	Major late promoter
mRNA	Messenger ribonucleic acid
MW	Molecular weight
Mtr	4-methoxy 2,3,6 trimethylbenzene sulphonyl
NEM	N-ethyl maleimide
NF	Nuclear factor
NP-40	Nonidet P-40
OPA	o-phthaldialdehyde
PCMB	p-chloromercuribenzoate
PCR	Polymerase chain reaction
p.f.u.	Plaque forming units
PBS	Phosphate buffered saline
PMSF	Phenylmethylsulphonylfluoride
pol	Adenovirus DNA polymerase
PPV	Plum pox virus
pRB	Retinoblastoma susceptibility gene product
pTP	Pre-terminal protein
PVY	Potato virus Y
RNA	Ribonucleic acid
RSV	Rous sarcoma virus
SBTI	Soya bean trypsin inhibitor
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SFV	Semliki Forest virus
S-MEM	Earle's minimal essential medium for suspension culture
SPDP	N-succinimidyl 3,2(pyridyl-dithio)propionate
ss	Single stranded
SSC	Salt/sodium citrate buffer
T	Thymine
TBS	Tris buffered saline
TEMED	N,N,N',N' tetramethylethylenediamine
TEV	Tobacco etch virus
TF	Transcription factor
TFA	Trifluoroacetic acid
TLCK	Tosyl-lysine-chloromethyl ketone
TLE	Thin layer electrophoresis
TP	Terminal protein
TPCK	Tosyl-phenylalanine-chloromethyl ketone
ts	Temperature sensitive
TTBS	Tris buffered saline containing Tween 20
TVMV	Tobacco vein mottling virus
uv	Ultra-violet

VARNA	Viral associated ribonucleic acid
VV	Vaccinia virus
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactoside
YFV	Yellow fever virus

Single Letter code and standard abbreviations for amino acids:

A	Alanine	Ala
C	Cysteine	Cys
D	Aspartic acid	Asp
E	Glutamic acid	Glu
F	Phenylalanine	Phe
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
K	Lysine	Lys
L	Leucine	Leu
M	Methionine	Met
N	Asparagine	Asn
P	Proline	Pro
Q	Glutamine	Gln
R	Arginine	Arg
S	Serine	Ser
T	Threonine	Thr
V	Valine	Val
W	Tryptophan	Trp
Y	Tyrosine	Tyr

ACKNOWLEDGEMENTS

I am indebted to all the members of the academic and technical staff of the late Department of Biochemistry and Microbiology for many useful discussions and for providing expertise in a wide range of aspects of biochemistry and virology. In particular I would like to thank my supervisor Graham Kemp for his ideas, continual encouragement and advice throughout the practical and written components of the project. I am also very grateful to Paul Talbot for all his help, particularly in tissue culture and for his good humour.

I thank Carl Anderson (Brookhaven National Laboratory, New York) for giving me the opportunity to work in his laboratory for a month, for supplying the Ad2ts1 stock and for providing unpublished information. I also acknowledge the time and patience of his technician, Ronni Greene, who gave me an excellent introduction to the basic techniques of cloning. I am indebted to Ron Hay for his many useful suggestions, for supervising the baculovirus expression component of the project and for his patience during the writing up stages. I thank Ian Leith for providing the components for the *in vitro* replication assays and for being a willing and valuable source of information throughout the project.

I am grateful to Ian Armitt and John Nicholson for preparing media and for synthesising the oligonucleotides, Julia Bosher for providing the pTP antiserum, Lars Rogge (University of Munich) for providing a recombinant baculovirus expressing the adenovirus pre-terminal protein, Suzanne Russell for constructing most of the CHEM-X models and Richard Nicholson for advice and help on the use of the Wisconsin computing package. I thank Bill Blythe for his photography and Willie Russell for his lively discussions and critical reading of the manuscript. Finally I acknowledge the encouragement and support of my parents throughout the project.

ABSTRACT

The substrate specificity, classification, expression and control of the adenovirus encoded protease, the activity of which is required for the production of virions, is described. The use of synthetic peptides has shown that the protease cleaves sites of the form (M,L,I)XGG-X or (M,L,I)XGX-G and these consensus sequences have been used to identify potential cleavage sites in all the known substrates of the protease. Putative cleavage sites have also been found in a number of other adenovirus proteins including the major coat proteins, the hexon and the penton, that had not previously been considered as substrates of the protease. The octapeptide, MSGGAFSW, has been used to develop a specific assay for the protease where digestion is monitored by HPLC reverse phase chromatography.

Purification of the protease from adenovirus particles and the preparation of antipeptide sera against the L3 23-kDa protein have been used to confirm that the latter is the protease and that it is not proteolytically activated. Inhibitor studies reveal that the enzyme is inhibited by the thiol directed reagents iodoacetate, PCMB and DTDP, and activated by DTT or cysteine suggesting that it is a member of the cysteine class of proteases. Analysis of the sequence of the enzyme, however, shows that it is not a papain-like enzyme and it is suggested that it might be a member of the recently identified subclass of cysteine proteases that are related to trypsin.

The 23-kDa protease has been expressed using *E.coli* and baculovirus expression systems and a purification schedule for the protein was developed using anion exchange and hydrophobic interaction chromatography. The purified enzyme, however, was not able to cleave the peptide substrate MSGGAFSW and the conclusion was drawn that the protein is probably synthesised as an apoenzyme.

One of the substrates of the protease, the pre-terminal (pTP) protein was expressed in insect cells using a recombinant baculovirus. Coinfections of cells with recombinant 23-kDa and pTP baculoviruses resulted in efficient processing of the pTP to the intermediate terminal protein (iTP) *in situ*. The partially purified baculovirus expressed 23-kDa protein was not, however, found to digest the pTP *in vitro*, supplying further evidence that the adenovirus protease requires a cofactor.

CONTENTS

CHAPTER 2

CHAPTER 3

INTRODUCTION

The broad aim of this study is to further characterise the adenovirus protease and to investigate its mode of action at the molecular level. In particular, the substrate specificity, mechanism and expression of the virus encoded enzyme will be considered. Active protease is an essential requirement for the production of virions and so the information presented here provides a basis for future studies aimed at designing and synthesising specific inhibitors as anti-viral agents. It is hoped further that the work on the adenovirus protease will be extended to provide a model for the study of proteolysis in other DNA viruses and that the similarities and differences between RNA and DNA viral proteases can be highlighted. Finally it is envisaged that, through its substrates, the protease could provide the key to unravelling some of the complex protein chemistry of the adenovirus and perhaps give an insight into control mechanisms in general.

The work will be presented in 4 chapters. The aim of chapter 1 is to use synthetic peptides to examine the substrate specificity of the protease while chapter 2 is primarily concerned with the classification of the enzyme. Attempts to express and purify active protease from *E.Coli* and baculovirus systems will be discussed in chapter 3 along with possible mechanisms for the control of its activity. The final chapter will be devoted to one of the substrates of the protease, the pre-terminal protein, and the central role that it plays in DNA replication. Before considering these topics in detail it is important to put the project into context and this will be done from 2 viewpoints. First, the epidemiology and replicative cycle of the adenoviruses will be outlined, with particular reference being made to the structure and function of the viral proteins. The second part of the introduction aims to underline the essential role that proteolysis plays in virus infections. Specific examples, including the well characterised proteases from HIV and poliovirus, will be used in an attempt to demonstrate the scope and diversity of viral proteases.

1 ADENOVIRUSES

1.1 EPIDEMIOLOGY AND CLASSIFICATION

The adenovirus was first recognised as an etiological agent of acute respiratory disease in the early 1950s (Rowe *et al.*,1953). Subsequently over 100 distinct serotypes have been identified that infect many different animal species (Horwitz,1990b). A number of criteria have been used to subdivide the adenoviruses into related groups and as a result the **Adenoviridae** family is made up of 2 genera, the **Aviadenoviruses** and the **Mastadenoviruses** (Ishibashi and Yasue,1984). At least 15 different serotypes of Aviadenoviruses have been isolated from a variety of avian species including hens, ducks, turkeys and pheasants and one such virus, FAV-1, has been shown to be responsible for a condition known as egg drop syndrome (McFerran *et al.*,1977). The disease causes hens to lay abnormal, soft shelled eggs for 3-4 days and is economically significant in the poultry industry (Ishibashi and Yasue,1984).

Most studies, however, have concentrated on the Mastadenoviruses which infect a range of mammalian species including foxes, sheep, cows, pigs, horses, dogs, monkeys, mice and humans. These viruses share a common antigenic determinant which is located on the major viral coat protein, the hexon (Norrby *et al.*,1976). To date, 47 serotypes have been shown to infect man and these are responsible for a number of diseases (Hierholzer *et al.*,1991). Different serotypes are specific for different tissues with the respiratory tract, gastrointestinal tract and ocular tissue being the favoured targets (**Table 1a**). Epidemics of acute respiratory disease (Ad4, 7) and keratoconjuntivitis (Ad8, 19, 37) are well documented and can have severe consequences for the individuals concerned (Huebner *et al.*,1958; Guyer *et al.*,1975 and Kemp *et al.*,1983). Less severe respiratory and eye infections are common amongst infants and most children acquire antibodies to one or more of the adenovirus serotypes before entering adolescence (reviewed in Straus,1984). Adenovirus types 40 and 41 are enteric viruses that replicate in the gastrointestinal tract and are a major cause of death from diarrhoea in children of the third world (Hammond *et al.*,1987).

The observation that Ad12 induces tumours in rodents created much interest and raised questions as to whether adenoviruses are responsible for some human cancers (Trentin *et al.*, 1962). Thirty years on and there is still no direct evidence to link any of the adenoviruses with

Associated Disease	Principal Serotypes
Acute febrile pharyngitis	1, 2, 3, 5, 6,7
Pharyngoconjunctival fever	3, 7, 14
Acute respiratory disease	3, 4, 7, 14, 21
Pneumonia	1, 2, 3, 4, 7
Epidemic keratoconjunctivitis	8, 11, 19, 37
Pertussis-like syndrome	5
Acute hemorrhagic cystitis	11, 21
Gastroenteritis	40, 41
Hepatitis	1, 2, 5
Persistence of virus in urinary tract	34,35

Table 1a Adenovirus serotypes and their associated diseases
A list of the human adenovirus serotypes and the conditions with which they are normally associated.

Subgroup	Serotypes	Oncogenic	% GC in DNA
A	12, 18, 31	High	48-49
B	3, 7, 11, 14, 16, 21, 34, 35	Moderate	50-52
C	1, 2, 5, 6	Low or none	57-59
D	8, 9, 19, 37, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36, 37, 38, 39, 42-47	Low or none	57-61
E	4	Low or none	57-59
F	40, 41	Unknown	-

Table 1b Classification of Human Adenoviruses
Classification of the human adenoviruses based on their DNA homologies and their ability to induce tumours in animals. Modified from Horwitz (1990a).

human tumours; but their ability to transform rodent cells has served as a useful animal model to study the molecular basis of cell transformation (Graham,1984).

The 47 human serotypes have been classified into 6 groups (A-F) based on DNA sequence homologies, common antigenic determinants, hemagglutination reactions and their oncogenic properties (Horwitz,1990a) and these are listed in **Table 1b**. DNA sequence homologies between members of the same group are usually very high (48-85%) with the most significant differences occurring in the coding regions for the major coat proteins. In contrast, less than 20% DNA sequence homology has been demonstrated between viruses from different groups (Garon *et al.*,1973). Despite their differences all the adenoviruses have the same basic structure and genome organisation, with the sequences of many functional proteins being highly conserved between viruses from different subgroups. Most reference will be made to Ad2 when considering the structure and organisation of the virus since its entire genome has been sequenced (Roberts *et al.*,1986) and Ad2 was the serotype used throughout this project. The general principles and structural features, however, are equally applicable to the other mastadenoviruses and aviadenoviruses that have been characterised.

1.2 INFECTIOUS CYCLE

Individual cells generally have 2 fates following their infection by an adenovirus. They can (1) be transformed by the integration of part or all of the viral DNA into their genome or (2) play host to a full lytic infection and act as a factory for the production of thousands of progeny virions.

Transformation takes place in semi- or non-permissive cells and, depending on the adenovirus serotype and the nature of the host, it may be a malignant event. All of the human adenoviruses are able to transform new born rodent cells, but only serotypes from groups A and B are known to be oncogenic. The transformation and oncogenic properties of adenoviruses will be considered further in section 1.5b. The primary concern of this project, however, is the adenovirus protease which has an essential role in the development of the infectious cycle in permissive cells. The adenovirus lytic cycle can be divided into 6 stages as follows:

(a) Uptake of the virus by the cell.

(b) Transcription of early viral mRNAs; translation of early proteins.

(c) Replication of viral DNA.

(d) Transcription of late viral mRNAs; translation of late proteins.

(e) Assembly of virus particles.

(f) Maturation of virions.

The adenovirus protease has been shown to be responsible for the maturation of virions (f); but its involvement at earlier stages in the infectious cycle cannot be ruled out. In particular, there is a suggestion that the enzyme may play a role in the uncoating of virus particles soon after they enter the cell (Hannan *et al.*,1983). Before considering each of the stages of the lytic cycle in turn, the structural organisation of the adenovirus particle will be described.

1.3 STRUCTURE OF THE ADENOVIRUS

Adenoviruses are double stranded DNA viruses with a 34-36K base pair genome. The genome contains open reading frames with the potential to code for at least 50 distinct polypeptides and is encapsidated by a protein shell (Green *et al.*,1967; Nermut,1984; Sussenbach,1984). Analysis of purified virions by SDS polyacrylamide gel electrophoresis revealed 12 distinct polypeptides and these were termed II-XII according to their relative mobilities, with protein II having the largest apparent MW and protein XII the smallest (Maizel *et al.*,1968a). **Fig.1** shows a schematic representation of the virus and an SDS-PAGE separation of its major structural proteins. Seven of these, II, III, IIIa, IV, VI, VIII and IX have been shown to be associated with the viral coat and, as in many other non-enveloped DNA and RNA viruses, including the polyoma- and picorna- viruses, the adenovirus coat proteins adopt an icosahedral conformation (Harrison,1991).

Disruption of the virion by treatment with acetone, 5M Urea, 10% pyridine or 0.5% deoxycholate removes the coat proteins leaving the viral cores (Russell *et al.*,1968; Maizel *et al.*,1968b; Prage *et al.*,1970; Nermut,1984). The cores are made up of the viral DNA and a number of proteins including V, VII, mu and the terminal protein (Chatterjee *et al.*,1986).

The external structure of the adenovirus has been studied using a variety of techniques

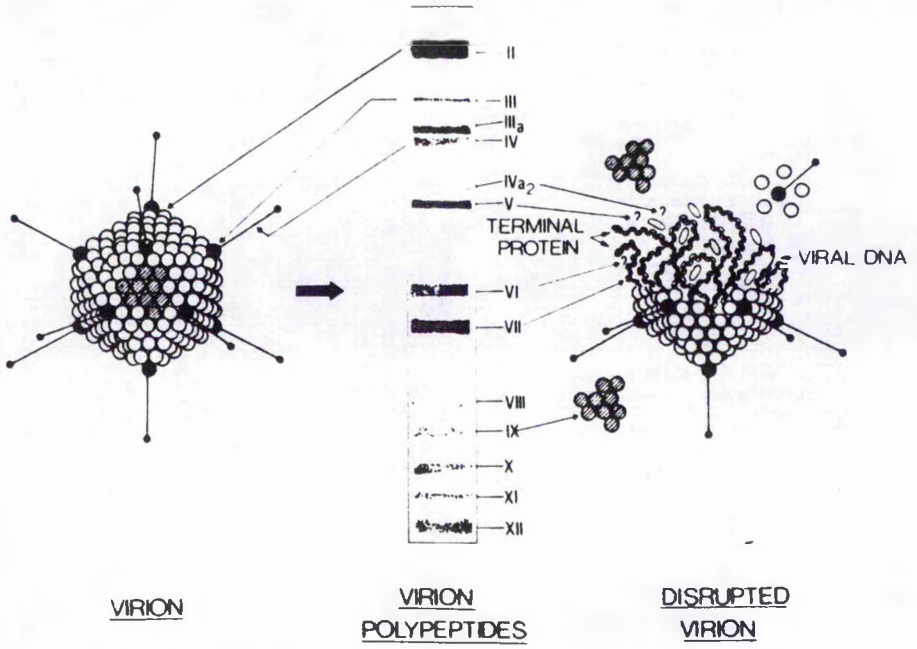


Fig. 1. Structure of the Adenovirus.
The diagrammatic representation of the Ad2 virion on the left shows the locations of the hexon (II), penton (III) and fibre (IV) proteins in the outer capsid. Alongside is a polyacrylamide gel showing the relative mobilities of the Ad2 structural proteins. Disruption of the virion by dialysis against a hypotonic buffer removes the coat proteins to reveal the double stranded DNA genome and associated core proteins (from Precious and Russell,1982).

including electron microscopy, cryo-electron microscopy and X-ray crystallography and there is general agreement with respect to its organisation and protein composition. Much less certainty, however, exists with regards to the structural organisation of the DNA and associated proteins that make up the viral cores.

1.3a The Adenovirus Coat

As mentioned above, the adenovirus has an outer shell made up entirely of proteins. This serves as a permeability barrier to protect the viral DNA; but must also be able to undergo a number of morphological changes during the infectious cycle. This combination of strength and versatility relies upon the formation of many specific protein-protein interactions. Electron microscopy studies in the mid 1960s revealed that the adenovirus protein coat is made up of 252 subunits or capsomers and, based on their relative geometric positions, these subunits were termed hexons and pentons (Valentine and Pereira,1965; Ginsberg *et al.*,1966). Each particle is made up of 240 hexons and 12 pentons with a penton occupying each of the 12 vertex positions in the icosahedron. Projecting away from each penton is a fibrous structure, termed the fibre, which can be seen in the diagrammatic representation of the virion shown in

fig.1. Of the 7 proteins that form the viral coat, II, III and IV make up the capsomers, which constitute the main body of the coat, whilst IIIa, VI, VIII and IX are thought to stabilise the overall structure and so are often referred to as cementing proteins (Ginsberg,1979; Everitt *et al.*,1973; Boulanger *et al.*,1979).

Hexon (Protein II)

The Ad2 hexon monomer (II) comprises 967 amino acids and has an apparent MW of 110-kDa (Akusjarvi *et al.*,1984). Crystallisation of the hexon (Pereira and Russell,1969; Franklin *et al.*,1971) and the solution of its 3D structure from X-ray diffraction data were major breakthroughs and make it the best characterised of the viral structural proteins (Roberts *et al.*,1986). The most striking features of secondary structure are 2 very similar 8 stranded β-barrels present in each hexon monomer. Projecting away from these are 4 loop structures designated l_1, l_2, l_3 and l_4 (see **fig.2a**). In the virion coat 3 hexon monomers come together to form each of the 240 hexon capsomers (Grutter and Franklin,1974). The 3 polypeptide chains are folded together by a novel pathway which is catalysed by another virus encoded protein, the L4 100-kDa protein (Cepko and Sharp,1983). The resulting trimer has 2 main structural regions, a pseudohexagonal base and a tower region as is shown in **fig.2b** (Burnett,1985). The base includes the 2 β-barrel domains from each monomer with the direction of the β-barrels being perpendicular to the surface of the virion. The hexon bases are tightly packed to form a protective protein shell that is 33A° thick (Roberts *et al.*,1986). Projecting away from each base, on the outside of the virion, are 3 towers and it is in this region that most of the sequence variation exists between hexons of different serotypes (Toogood and Hay,1988; Toogood *et al.*,1989). Each tower is 73A° high and is made up of 3 loops (l_1, l_2 and l_4) with one loop being donated by each of the hexon monomers (Roberts *et al.*,1986).

Penton (proteins III and IV)

Each penton capsomer is made up of 2 viral proteins, III and IV, which are often referred to as the penton base and fibre respectively. Five copies of protein III associate to form the base of each capsomer, whilst 3 copies of protein IV form the fibrous structure, on the exterior of the virion, which contains the viral receptor binding site (**fig.3**).

In Ad2, protein III contains 571 amino acids and has an apparent MW of 85-kDa. Five

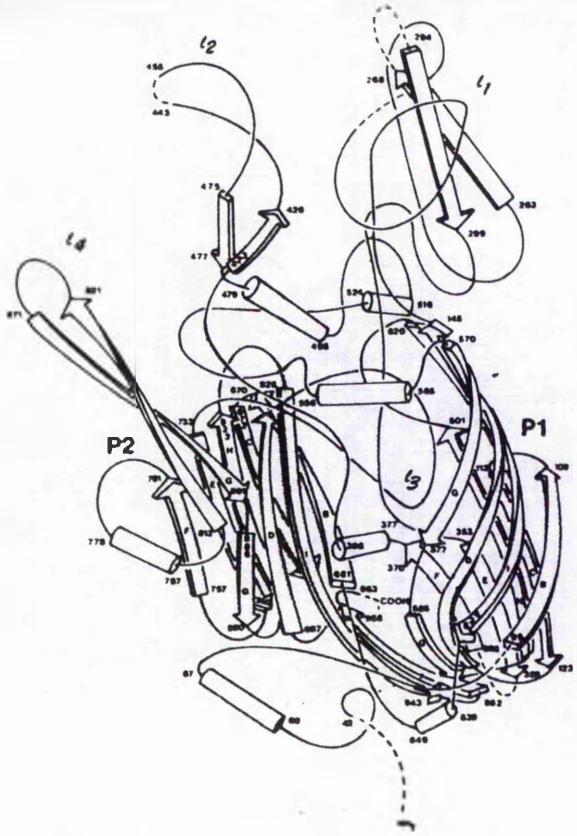


Fig. 2a Structure of the Ad2 Hexon
Alpha carbon trace of an Ad2 hexon subunit (from Roberts *et al.*,1986). The loops l1, l2 and l4 protrude from the 8 strand β-barrels (P1 and P2) and form 3 towers (T) which project from the main body of the virion. The l3 loop is sandwiched between the P1 and P2 domains. The dotted lines indicate regions where the structure has not been solved. The orientation of the trace is looking from the central cavity of the trimeric molecule toward the B hexon: hexon contact face (see below).

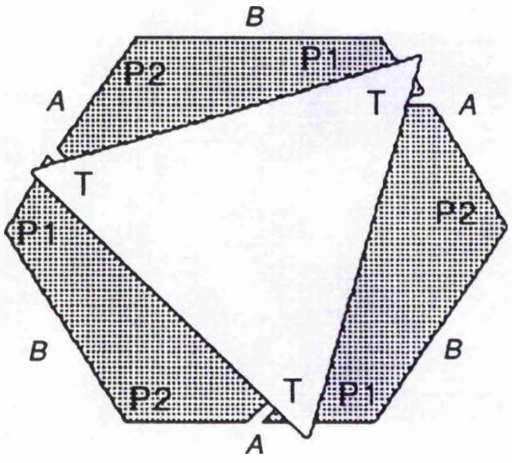


Fig. 2b The Hexon Trimer
The hexon is made up of three copies of protein II and contains two distinct regions: the base and the tower. The base of the hexon is shaded and the arrangement of the P1 and P2 β- barrels means that each hexon has two distinct faces (A and B). Projecting out from the base are three towers, with each A face lying under an apex of the triangular tower.

copies of III are non-covalently bound to give a ring like pentamer, which appears less dense in its centre (Stewart *et al.*,1991). Projecting out from the centre of each penton base is the fibre (protein IV) which is made up of 3 distinct domains, an N-terminal region that attaches the protein to the penton base, a central section comprising 60% of the protein that forms a long flexible shaft and a C-terminal knob like domain (Devaux *et al.*,1987; Stewart *et al.*,1991). The N-terminal 49 amino acids are relatively well conserved between adenovirus serotypes and resemble a signal sequence in that a negative stretch of amino acids is followed by a hydrophobic region and then by a second negative region (Ruigrok *et al.*,1990). It has been shown conclusively that the N-terminus of the protein is responsible for locking the Ad2 fibre into the penton base and that the removal of the N-terminal 17 amino acids, by proteolysis, is sufficient to prevent the association (Deveaux *et al.*,1987). Although the amino acid sequence of the remainder of protein IV is not well conserved, its overall structure is. In Ad2, the shaft of protein IV comprises 360 of the 582 amino acids and includes a 15 residue motif that is repeated 22 times. In the Ad3 protein IV there are only 6 such repeats in the shaft, a difference which explains why the length of the fibre varies significantly between serotypes (Norrby,1969). At the end of the fibre the C-terminal domain appears as a globular β-barrel and is thought to contain the receptor binding site (Green *et al.*,1983).

Although it has been well established that each penton capsomer contains 5 copies of protein III, the number of copies of protein IV making up the fibre has been the subject of controversy, with 2 and 3 copies being reported (Green *et al.*,1983; van Oostrum and Burnett,1985). Recent chemical crosslinking studies prove, however, that each fibre is made up of 3 copies of protein IV and suggest that the shaft is a triple helix and not an extended β-pleat structure as suggested previously (Deveux *et al.*,1990; Ruigrok *et al.*,1990). In contrast to the hexon capsomers which require the L4 100-kDa protein for assembly, proteins III and IV are able to associate *in vitro* and *in vivo*, without assistance, to form penton capsomers (Boudin *et al.*,1979).

Proteins IIIa, VI, VIII and IX

Four other viral proteins appear to be associated with the viral capsid and are thought to stabilise the virion by reinforcing existing intermolecular contacts.

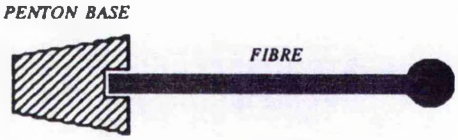


Fig. 3 The Ad2 Penton

Diagram of the Ad2 penton capsomer, comprising the penton base and the fibre. The penton base is made up of five copies of protein III and the fiber of three copies of protein IV.

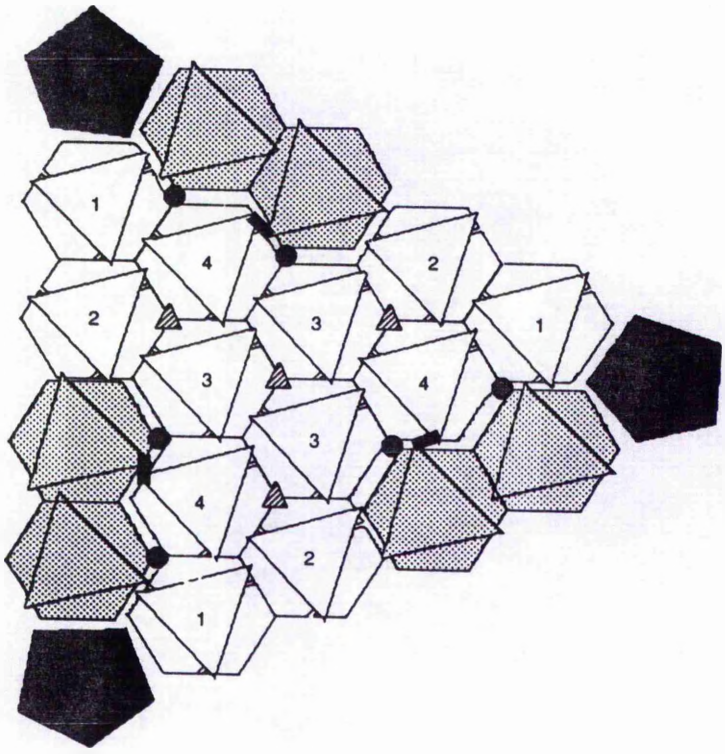


Fig. 4 Capsid Face.

Schematic diagram showing the arrangement of capsomers and associated proteins on one face of the icosahedral capsid of the adenovirus. The hexons are shown as hexagons, with a superimposed triangle representing the towers, and the pentons as pentagons. The 4 non-equivalent hexons are numbered with hexons 2, 3 and 4 making up the group-of-nine (GON). The shaded hexons are from adjacent faces of the capsid. The positions thought to be occupied by proteins VI (●), VIII (▬) and IX (◬) are also shown. Adapted from Stewart *et al*(1991).

IIIa is a phosphoprotein with an apparent MW of 66-kDa in Ad2. Dissociation of virus particles suggests that it is associated with the peripentonal hexons and labelling studies indicate that there are 5-6 copies of IIIa present at each vertex. Additional work has revealed associations between protein IIIa and the core proteins V and VII, leading to the proposal that it may be responsible for stabilising the interactions between the viral core and coat proteins (Boudin *et al.*,1980). The significance of the phosphorylation of IIIa is not known, although it has been suggested that it is carried out by a virion associated kinase and that it takes place early in infection (Blair and Russell,1978; Tsuzuki and Luftig,1983).

Proteins VI and VIII, with apparent MWs of 24-kDa and 14-kDa respectively, are both thought to be associated with hexon capsomers on the internal surface of the virion coat (Everitt *et al.*,1975; Philipson,1983; Horwitz,1990a). Possible locations for these proteins have been proposed on the basis of cryo-EM studies and the assumption is that proteins VI and VIII exist as trimers and dimers respectively (Stewart *et al.*,1991; van Oostrum and Burnett,1985). Reports that protein VI may also bind to the viral DNA have led to the suggestion that, like IIIa, VI may play some role in linking the ordered capsid to the viral cores (Russell and Precious,1982).

The association of protein IX with hexons on the outer surface of the virus coat has been well established using a range of biochemical and physical techniques (Boulanger *et al.*,1979; van Oostrum and Burnett,1985; Furcinitti *et al.*,1989; Stewart *et al.*,1991). Protein IX, which comprises 140 amino acids, has an apparent MW of 14-kDa and trimers of the polypeptide are located between specific hexon capsomers, with 4 trimers being associated with each of the 20 faces. The exact position of the IX trimers has been shown to be just above the pseudohexagonal base at the contact point between 3 hexon capsomers.

Overall Geometry of the Virus Coat

The asymmetrical nature of the hexon capsomers (**fig.2b**) and the presence of a penton at each vertex means that the hexons occupy 4 non-equivalent positions in the virion capsid. A diagram of the capsomers making up one face is included to illustrate this point (**fig.4**). Also shown are the proposed locations of proteins IIIa, VI, VIII and IX (Stewart *et al.*,1991). The contacts formed between adjacent hexons are much stronger than those between hexons and

pentons. This is illustrated by the fact that groups of 9 hexons (GONs), peripentonal hexons and pentons can be readily dissociated from the virions by a number of mild treatments including dialysis against a hypotonic buffer (Prage *et al.*,1970). Penton-less virions are referred to as sub-viral particles and their genome is much more susceptible to nucleases than the complete virion. The relative ease with which the pentons can be removed is thought to be an important step in the uncoating stage of the lytic cycle (see section 1.4).

1.3b Adenovirus Cores

DNA and RNA viruses have adopted a number of strategies for neutralising the charge of their packaged nucleic acids. For example, the coat proteins of a number of simple plant viruses contain a stretch of basic amino acids, close to their N-terminus, which interacts with the phosphates of the nucleic acids, inside the virion, whilst animal picornaviruses co-package polyamines with their RNA. Generally speaking, it appears that viral nucleic acids are not arranged in an ordered way within the virion, although X-ray crystallography has been used to show that the RNA of bean pod mottle virus and the single stranded DNA of canine parvovirus are packaged in an ordered icosahedral structure within their respective coats (Johnson,1991). Even in these relatively simple viruses much more information is required before there is a clear picture of how they package their nucleic acids.

When more complex viruses, such as the adenovirus, are considered surprisingly little is known about the structural organisation of their nucleic acids. Several theories have been put forward including the idea that the adenovirus DNA adopts a chromatin like structure, that it is packaged in 12 loops with one interacting with each of the 12 vertices in the outer capsid and that it forms an icosahedral cage within the outer capsid (Mirza and Weber,1982; Brown *et al.*,1975; Nermut,1979). Recently, the use of the restriction enzyme, Bal31, which cleaves supercoiled DNA, in parallel with electron microscopy, has led Wong and Hsu (1989) to suggest a continuous helix model for adenovirus DNA (**fig.5**). They propose that in the virion the DNA forms 8 supercoiled loops which are wrapped round a central protein filament. Furthermore, they have demonstrated that the promoters of E1a and E4, along with the 2 origins of replication, are insensitive to Bal31 and so include these regions outside the supercoiled domains in their model. The significance of this will be discussed in section 1.6.

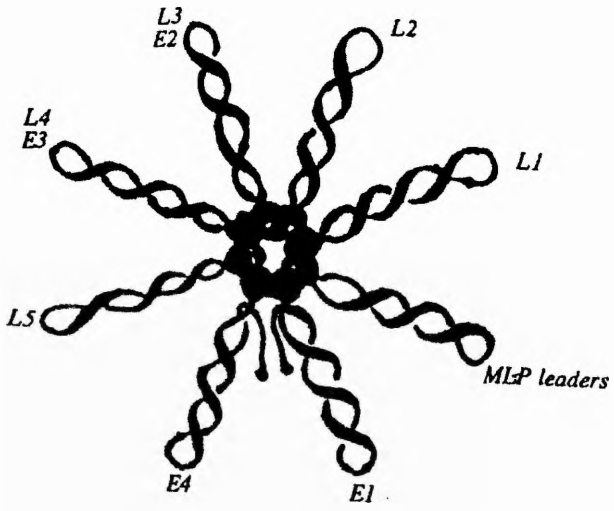


Fig. 5 Model for the structure of adenovirus DNA in virions.
The above diagram is modified from Wong and Hsu (1989). They propose an 8 loop domain structure for the adenovirus DNA with each loop being anchored to a central protein core. The locations of the early genes (E1-E4), the late genes (L1-L5) and the major late promoter (MLP) are indicated. The 2 termini of the genome, containing the origins of replication and the E1A and E4 promoters, are proposed to be located outwith the supercoiled domains.

Although it is well established that the adenovirus relies on a number of basic histone-like proteins to neutralise its DNA and to package it within the virion, the nature of the DNA-protein interactions remains uncertain. Two of these proteins V and VII, present at 180 and 1070 copies per virion, are termed the minor and major core proteins, respectively (Russell *et al.*,1968). Protein VII has an apparent MW of 18-kDa, is arginine rich and binds particularly tightly to the DNA, with 3M NaCl being required to dissociate it (Cupo *et al.*,1987). Protein V, on the other hand, can be removed from the viral DNA by relatively mild treatments leading to the suggestion that it forms a protein shell on the surface of the cores (Vayda *et al.*,1983). An alternative proposal is that 3 dimers of VII associate with 150 base pairs of DNA to form nucleosome like structures which are separated by variable lengths of DNA associated with the minor core protein, V (Corden *et al.*,1976). Neither of these models, however, is consistent with the results of Wong and Hsu (1989) and neither accounts for the presence of a third core protein termed mu (μ). Mu is a 19 amino acid protein, including 9 arginines and 3 histidines, which binds very tightly to DNA and is estimated to be present at 180-340 copies per virion (Weber and Anderson,1988). In addition to V, VII and mu, other

proteins are present in the viral cores at low copy number. One of these, the terminal protein (TP), is a hydrophobic protein with apparent molecular weight 55-kDa, which is covalently linked to the dCMP residue at the 5' end of each DNA strand. This protein plays a key role in viral replication and will be discussed further in section 1.6 and also in chapter 4.

Finally, with regard to the adenovirus cores, recent cryo-EM work, examining intact Ad2 particles, produced density maps suggesting that the centre of the virion is less dense than the outer core; but that no highly ordered inner structure exists (Stewart *et al.*,1991).

Six of the structural viral polypeptides described above, namely IIIa, VI, VIII, VII, mu and the TP are the products of the digestions of larger precursor molecules termed pIIIa, pVI, pVIII, pVII, 11K and the pTP respectively (Anderson *et al.*, 1973; Boudin *et al.*,1980; Tremblay *et al.*, 1983). The digestions are all thought to be carried out by the virus encoded protease, in a final maturation step, which is essential for the production of infectious virions (see section 1.8). The effect that the limited proteolysis has on the individual proteins and, indeed, on the overall structure of the virus remains to be elucidated.

1.4 UPTAKE OF THE ADENOVIRUS BY THE CELL

The knob at the end of each fibre is thought to contain the sequence required for the binding of the adenovirus to a specific cell surface receptor. The presence of 12 fibres per virion enables multivalent attachments and EM studies show that the virus is taken into the cell by receptor mediated endocytosis; a process involving the formation of clathrin coated pits, which bud off into the cells (Pastan *et al.*,1987). The resulting clathrin coated vesicles fuse with endosomes to give receptosomes, which have an internal pH of 5 and the low pH is thought to induce a change in morphology of the viral coat releasing the penton capsomers. It has been suggested that the low pH exposes hydrophobic epitopes in the hexon that can interact with the lipid bilayer and so allow the viral particles to penetrate the membrane (Everitt *et al.*,1988). The virus is then transported to the nucleus possibly through a specific interaction between the hexon and the microtubules (Dales and Chardonnet,1973). At the nuclear membrane the uncoating of the virions is completed and the viral DNA passes into the nucleus along with closely associated proteins. The remainder of the viral proteins are left behind in the

cytoplasm.

As is evident from this description of the uptake of the virions by the cell, the molecular mechanisms involved are poorly understood and much more work is required in this area to characterise the cell surface receptor and to define the binding epitopes on the fibre. In addition the transport pathway of the virion through the cell remains to be confirmed and the mechanisms and degree of uncoating of the virus particle are anything but clear. One of the viral proteins, IIIa, has been shown to be phosphorylated soon after the uptake of the virus by the cell (Tsuzuki and Luftig,1983). It is possible that changes in the phosphorylation state of IIIa or other viral proteins could cause morphological changes in the particles. Another phenomenon which is puzzling, but may be of relevance, is that purified penton bases are able to induce cytopathic changes when added to cell monolayers in tissue culture (Boudin *et al.*,1979). The mechanism behind this is not known, but it is possible that the adenovirus recognises 2 cell surface receptors, with the penton base binding to a receptor distinct from that of the fibre. With this in mind, it may be significant that the penton base (protein III) contains the amino acid sequence RGD, which has been shown to be responsible for the binding of a number of proteins including fibronectin to cell surface receptors (Ruoslahti and Pierschbacher,1986). Finally, it has also been noted that viruses lacking the protease are defective in uncoating, so specific cleavages by the adenovirus protease might have an important role to play in uncoating (Hannan *et al.*,1983).

1.5 ADENOVIRUS TRANSCRIPTION

The study of adenovirus transcription is of great interest not least because of the relationship between viral transcription regulatory proteins and malignant transformation, which will be discussed in section 1.5b (Heintz,1991). First, however, consider the complex organisation and control of transcription during the infectious cycle.

1.5a The Lytic Cycle

On entering the nucleus, the viral DNA is transcribed in a precise and ordered manner. Ten primary RNA transcripts have been identified that are each differentially spliced to give families of related mRNAs. The capped, polyadenylated mRNAs are then transported to the cytoplasm where translation takes place and at least 50 viral proteins are synthesised. The

regulation of gene expression is fundamental to the progression of the lytic cycle and is exerted at a number of levels with transcriptional control being paramount. In the first instance, the order and amounts in which the primary transcripts are synthesised is critical and then each must be processed to give its progeny mRNAs at the appropriate levels. The transportation of the mRNAs out of the nucleus and their stability in the cytoplasm are also subject to differential control (Sharp,1984).

Early infection is defined as the time prior to the onset of viral DNA replication. At this stage there are only 1-3 copies of the viral genome in the nucleus and transcription of the early genes takes place. Broadly speaking the products of these genes are required to modify the host cell to the advantage of the virus, to control the transcription of viral and cellular genes and to carry out viral DNA replication.

DNA replication begins 6-8 hours post infection with maximum rates being achieved by 18 hours post infection. By 24 hours post infection 10 000 -100 000 copies of viral DNA are present in the nucleus of each cell (Green *et. al.*,1970). Transcription of all the early genes continues after the onset of DNA replication; but, in addition, intermediate and late genes are transcribed at increasing levels as the lytic cycle progresses. The maximum levels of transcription from the late genes are achieved at 18 hours post infection and are maintained for at least 10 hours (Lucas and Ginsberg,1971). The intermediate and late genes nearly all code for structural proteins or proteins involved in virion assembly and maturation which, by definition, are not required until late stages of infection.

The factors governing the switch between early and late transcription are of great interest and a number of models have been proposed. An attractive theory is that the DNA from the original infectious virion, which is the template for early transcription, adopts a different conformation from its progeny DNA, which are synthesised during replication and are the templates for late transcription.

Before describing the events of early and late transcription in detail consider the structure and organisation of the viral DNA.

Adenovirus DNA and Organisation of the Genome

Adenovirus DNA varies in size from 34-36Kbase pairs in different human serotypes

and is asymmetric with respect to its distribution of A-T and G-C base pairs. By definition the right hand side is the A-T rich part and the right or positive strand is that which is transcribed from left to right (Doerfler and Kleinschmidt,1970). The complementary strand is transcribed from right to left and is termed the left or negative strand. The DNA from all adenovirus serotypes contain inverted terminal repeats ranging in length from 63-165 base pairs (Alestrom *et al.*,1982). In other words the 5' and 3' ends of the DNA are complementary to each other. These repeats include specific sequences necessary for DNA replication and transcription (see section 1.6) and are also important in terms of the structure of the DNA (Sussenbach,1984). A distinguishing feature of the adenovirus DNA is that it has a protein covalently attached to the 5' end of each of its strands. In the mature virion the protein is the 55-kDa terminal protein (TP), which is the digestion product of the virally encoded 87-kDa pre-terminal protein (pTP). The pTP is processed, by the adenovirus encoded protease, to the TP via an intermediate and the significance of this will be discussed in chapter 4 (Smart and Stillman,1982). **Fig.6** shows the transcriptional organisation of the Ad2 genome. Although there is substantial

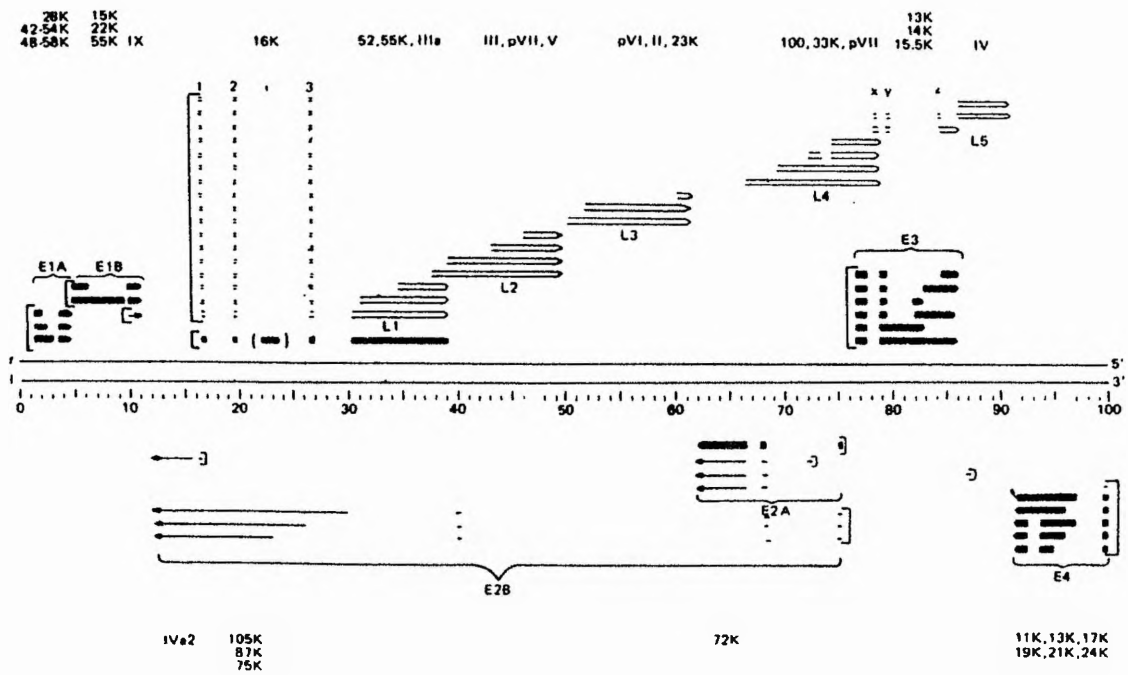


Fig. 6 Transcriptional organisation of the Ad2 genome.
Schematic representation of the adenovirus type 2 genome, divided into 100 map units; showing the locations of the major transcripts including their leader sequences and the protein coding regions. Early, intermediate and late transcripts are represented by solid, narrow and open arrows respectively (from Sussenbach,1984).

sequence divergence between some serotypes, most of the adenovirus genomes are similar with respect to the overall organisation of their transcriptional units. As can be seen transcription takes place off the right and left strands of the 36K base pair genome, which has been divided into 100 map units so that the locations of the individual genes can be readily identified (Gingeras *et al.*,1982).

Detailed examinations of the appearance of the various Ad2 mRNAs during infection has revealed further divisions of early and late transcription, namely immediate early (E1a), delayed early (E1b, E2a, E2b, E3, E4, L1), intermediate (IX, IVa2) and late (L1-L5) with synthesis of all these transcripts being carried out by the cellular enzyme, RNA polymerase II. In addition, there are 2 non-coding viral associated RNAs, VA-RNA_1 and VA-RNA_2, which are transcribed by RNA polymerase III. These are coded for by less than 1% of the viral genome and have been shown to play a key role in translational control (Sussenbach,1984; Mathews and Shenk,1991).

Immediate Early Transcription - E1A

The immediate early genes (E1) are switched on at 30 minutes post infection and their products are required for cellular transformation and for the transcriptional control of both viral and cellular genes. The E1A region is located at the 5' end of the right strand and gives rise to a complex transcriptional unit that generates 4 distinct mRNAs termed 13S, 12S, 11S and 10S, based on their sedimentation values (Nevins,1987). These are translated in the same reading frame to give 289, 243, 217 and 171 amino acid proteins respectively. The proteins all share the same amino and carboxy termini, with the differences in length resulting from RNA splicing. The 289 and 243 amino acid proteins, termed $E1A_{13S}$ and $E1A_{12S}$, are identical apart from 46 residues present at positions 140-185 in the former (Perricaudet *et al.*,1979). They are nuclear localised phosphoproteins and thus far all the functions attributed to the E1A region have been assigned to them.

$E1A_{13S}$ contains 3 conserved domains, the first 2 of which are shared by $E1A_{12S}$ (**fig.7**). Extensive mutagenesis studies by a number of groups have enabled functional assignments to be made to the various domains of the proteins (reviewed in Moran and Mathews,1987). Regions 1 and 2, present in both proteins, are primarily responsible for

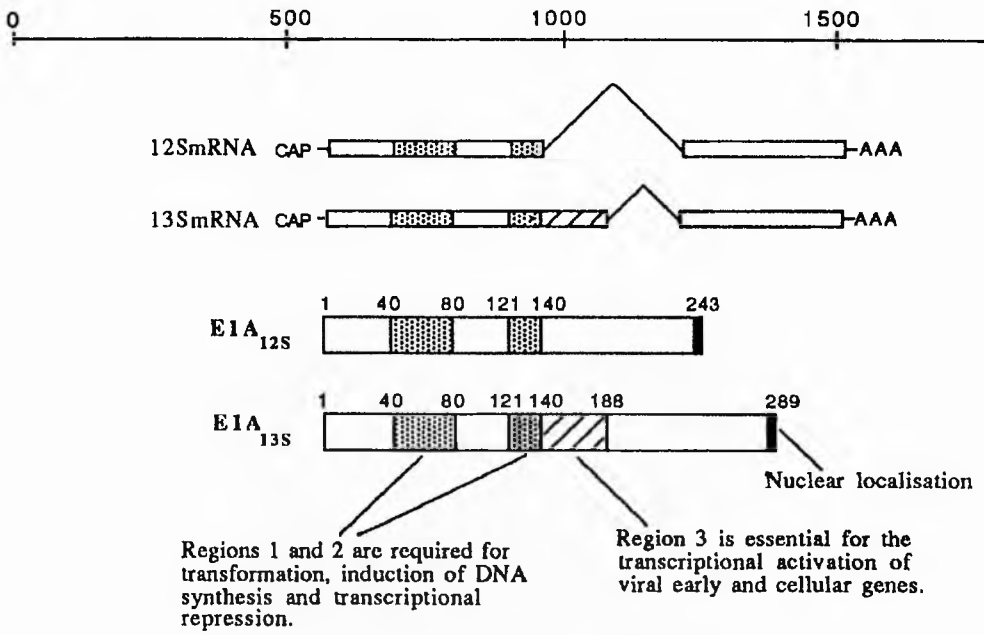


Fig. 7. Functional domains of the E1A proteins.
Diagram showing the origin and organisation of the E1A 12S and 13S mRNAs and their respective protein products. The E1A region is located at the 5' end of the right strand of the adenovirus genome and the map positions are included at the top of the figure. The important functional domains of the proteins, and their coding regions in the mRNAs, are shaded and the amino acid position numbers are shown above the bars.

transcriptional repression, transformation and DNA synthesis induction, whilst region 3, which is unique to $E1A_{13S}$, is essential for the transcriptional activation of viral and cellular genes. Indeed, a synthetic peptide comprising the 46 amino acids of region 3, along with 3 adjacent conserved residues, was found to be sufficient to stimulate transcription from several adenovirus promoters (Green *et al.*,1988). $E1A_{13S}$ switches on the transcription of all the viral early genes (E1B, E2A, E2B, E3 and E4) and also activates the major late promoter (MLP); but has not been shown to bind directly to the DNA (Berk,1986). Instead, it appears to modulate transcription by an indirect means. The current theory is that $E1A_{13S}$ operates by modifying existing cellular transcription factors including E2F, E4F, AP-1 and TFIIIC (reviewed in Nevins,1991). Two mechanisms for this activation have been put forward. The first proposal is that the cellular transcription factors are phosphorylated as a result of their association with $E1A_{13S}$. It is suggested that the kinase is a cellular protein complexed with $E1A_{13S}$ (Hoeffler *et al.*,1988; Herrman *et al.*,1991). Alternatively, there is some evidence to suggest that $E1A_{13S}$ binds directly to the transcription factors, which serve to escort it to the appropriate promoters

(Lillie and Green,1989). The full range of these interactions and the nature of the modifications remain to be established, but what is certain is that the E1A proteins direct the transcription of the other early viral genes and tailor the cells' machinery to suit its own requirements. In addition to their role in the lytic cycle, the adenovirus E1A proteins are also essential for cellular transformation and this will be discussed further in section 1.5b.

Delayed Early Transcription - E1B, E2A, E2B, E3, E4 and L1.

The delayed early transcripts are termed E1B, E2A, E2B, E3, E4 and L1 and, as with E1A, the individual families of mRNAs are created by differential splicing of the primary transcripts. All the delayed early genes are activated by $E1A_{13S}$ and are switched on from 1.5 to 7 hours post infection. Transcription of the L1 region continues until late in infection and will be discussed with the other products of the major late promoter in the late transcription section (see below).

The E1B region maps between 4.8 and 11.2 map units and differential splicing of the primary E1B transcript produces 2 major mRNA species with sedimentation coefficients of 13S and 22S. In addition, 14S and 14.5S mRNAs have been detected, but the functional significance of these is not known (Virtanen and Pettersson,1985). The 22S message carries 2 initiation codons in different, but overlapping reading frames, resulting in the synthesis of 2 unrelated proteins with apparent molecular weights of 55-kDa and 19-kDa (Bos *et al.*,1981). Thus far, these are the only well characterised E1B proteins and both are required for cellular transformation (see section 1.5b). In addition to its essential role in transformation, the E1B 55-kDa protein is required, in the infectious cycle, for the shut-off of cellular DNA synthesis, efficient DNA replication and the transport of late viral mRNAs across the nuclear membrane into the cytoplasm. It is thought that it performs some, if not all of these functions, in association with the E4 34-kDa protein (Sarnow *et al.*,1984; Bridge and Ketner,1990). The E1B 19-kDa protein is also required for the lytic cycle and appears to play an important role in stabilising cellular and viral DNA during replication. It has also been shown to trans-activate the other delayed early regions (White *et al.*,1988).

E2A is located at 61.5-75.4 map units on the left strand and codes for the 72-kDa single stranded DNA binding protein (DBP), that is essential for DNA replication (see section

1.6 and chapter 4). The DBP also has a role to play in transcriptional control and has been shown to repress transcription of the E4 region (Handa *et al.*,1983). The transcription of E2A continues late in infection, but whilst the promoter used at early times, at map position 75, is activated by $E1A_{13S}$, the late promoter at map position 72 is E1A independent. It has been suggested that the N-terminal domain of the DBP is involved in the splicing of the late viral mRNAs and that, through this role, it is able to influence the host range specificity of the virus (Anderson and Klessig,1984). The other 2 viral proteins required for viral DNA replication, the 80-kDa pre-terminal protein and the 140-kDa DNA polymerase, are coded for by the E2B region at position 75.4-14.2 map units on the left strand. Although the E2B transcript shares the same promoters and leader sequences as E2A, it is present at much lower concentrations. In Ad2 a sequence at 39 map units coding for Met-Ala-Leu is joined by RNA splicing to the 5' end of the main coding regions of both the pTP and polymerase proteins (**fig.8**). This exon includes the initiation codon for both the pTP and polymerase and effectively adds 18 and 142 amino acids to the predicted protein sequences (Stunnenberg *et al.*,1988). The structures and functions of the E2 proteins will be discussed in full in section 1.6 and chapter 4.

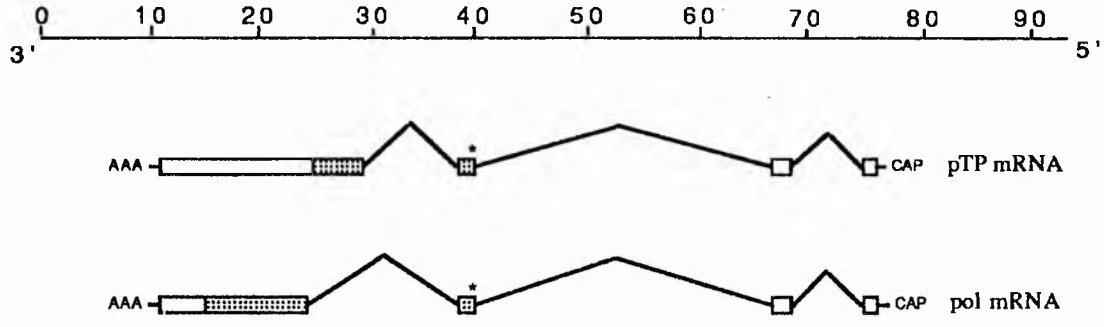


Fig. 8. Organisation of Ad2 pTP and pol mRNAs
Schematic diagram showing the organisation of the pTP and pol coding regions within the E2B region. The E2B region is transcribed from the reverse strand which is shown with the map units indicated. Three short exons are spliced to the main reading frames, at coordinates 28.9 and 24.1, to give the pTP and pol mRNAs respectively. The sections of the mRNAs coding for each protein are shaded. pTP and pol share the same initiation codon which is marked with an asterisk.

The E3 region is located at map position 75-86 on the right strand and deletion of these genes has been shown to have no effect on the progression of the lytic cycle in tissue culture. It does, however, contain a number of open reading frames encoding proteins that appear to

reduce the effects of the host's immune response. These include a 19-kDa glycoprotein that has been shown to bind to the MHC heavy chain preventing its transport to the cell surface and so lowering the cytotoxic T-cell response (Rawle *et al.*,1989). It has also been proposed that a 14.7-kDa E3 protein inhibits lysis of adenovirus cells by the tumour necrosis factor and that a 10.4-kDa E3 protein binds to the epidermal growth factor receptor (Gooding *et al.*,1988; Tollefson *et al.*,1990).

The E4 region is essential for the progression of the infectious cycle and is required for viral DNA replication and the synthesis of late viral proteins. It is located at the extreme right of the genome on the left strand at map units 99-91.3. The E4 transcript is processed to give a large family of mRNAs with the potential to produce at least 7 proteins (Virtanen *et al.*,1984). The E4 34-kDa protein carries out most of the functions assigned to the E4 region. In deletion mutants, lacking the 34-kDa protein, an 11-kDa protein, also encoded by E4, is able to deputise for the larger protein albeit not with quite the same efficiency. Mutant viruses lacking the genes for both proteins were found to be non-viable (Huang and Hearing,1989). As mentioned above the 34-kDa protein binds to the 55-kDa E1B protein and mutational analysis suggests that they perform some of their regulatory functions jointly. A third E4 protein, 19-kDa, transcriptionally activates the E2A gene by binding to the cellular factor E2F and increasing its affinity for the E2A promoter (Neill *et al.*,1990). The binding of the E4 19-kDa protein to E2F is enhanced by the presence of $E1A_{12S}$. It appears that the E1A protein is required to release the cellular factor from interactions with other proteins leaving it free to bind the 19-kDa protein.

The interaction between E2F, $E1A_{12S}$ and the 19-kDa E4 protein serves to emphasise the complex associations and interactions taking place between cellular and viral proteins during the early stages of adenovirus infection in preparation for viral replication and late transcription.

Intermediate Transcription - IX and IVa2

Transcription of the IX and IVa2 genes takes place at 6-8 hours post infection, from separate promoters. The gene for protein IX is contained within the E1B region, but is switched on independently to produce a 9S mRNA and is silent early in infection. Protein IX,

has already been described as a structural component of the virus capsid (section 1.3) and the reason why it is synthesised before the other structural proteins is not apparent.

The IVa2 gene is located on the left strand between 11.3 and 16 map units and the primary transcript is spliced to join 2 exons in the IVa2 mRNA (Crossland and Raskas,1983). IVa2 is a 449 amino acid protein with apparent molecular weight 50-kDa, but its function remains to be established. Originally it was thought to be a scaffolding protein required for assembly and not present in the mature virion. It has recently been shown to be a minor component of the mature virus and it has been suggested that its function is not confined to virus assembly (Winter and D'Halluin,1991). The observations that it contains a potential nucleotide binding site, that it is highly conserved in different adenovirus serotypes and that it has an affinity for viral DNA would appear to add weight to this argument (Gorbalenya and Koonin,1989; Kring and Spindler,1990; Russell and Precious,1982).

Late Transcription

The major late transcription unit is present on the right strand from 16.5 -99 map units, with the primary transcript being synthesised from the major late promoter. This is then spliced to give at least 18 distinct mRNAs which all share the same tripartite leader sequence at their 5' end. The mRNAs are divided into 5 families (L1-L5) according to their different poly-adenylation sites (**fig.6**). All the late mRNAs have 150-200 adenine bases at their 3' end and this appears to be required for their transportation to the cytoplasm and stability. The proteins encoded by the late transcription unit are either structural components of the virus or are required for assembly and are expressed maximally at 20 hours post infection (Fraser *et al.*,1979).

The L1 family of mRNAs encode 3 proteins, the 52-kDa, 55-kDa and pIIIa proteins. pIIIa is the precursor of IIIa which is a structural component of the viral capsid whilst the related 52-kDa and 55-kDa phosphoproteins are required for virion assembly (sections 1.3 and 1.7). In contrast to the other late mRNAs, the messages for the 52 and 55-kDa proteins have been detected in the cytoplasm early in infection, with the proteins themselves appearing at 6-7 hours post infection. This results from transcription from the major late promoter early in infection which is terminated midway through the late transcription unit. At early times

preferential poly-adenylation takes place at the L1 site and interestingly the other L1 protein pIIIa does not appear to be expressed. The role of the L1 52/55-kDa proteins at this stage in the infectious cycle is not known, but synthesis of the proteins continues late in infection (Shaw and Ziff,1980; Hasson *et al.*,1989).

The L2 family of mRNAs codes for III (penton base), V, pVII and an 11-kDa protein. pVII and 11K are the precursors of the basic core proteins VII and mu respectively. Protein V is also a core protein required for packaging the viral DNA, while the penton base is a major structural component of the virus capsid. All these proteins have been described previously (section 1.3).

L3 includes the genes for protein II (hexon), pVI and a 23-kDa protein. The 23-kDa gene product is required to carry out maturation cleavages of other viral proteins and is assumed to be the virally encoded protease, whilst the hexon and VI are structural components of the virion (sections 1.8 and 1.3). The genes for pVIII, the 100-kDa chaperone protein and a 33-kDa protein are located in L4 whilst the coding region for protein IV (penton fibre) is in L5. The functions of all these proteins are described elsewhere (sections 1.3 and 1.7).

It should be noted that the late transcription unit includes a number of other open reading frames that have not been included in this report and the significance of these remains to be established.

Viral Associated mRNAs (VA-RNAs)

The genes encoding the viral associated mRNAs are contained within the major late transcription unit, at map position 28.8-29.5, and are transcribed by RNA polymerase III to give two 160 nucleotide transcripts. Synthesis of these non-coding RNAs, termed VA-RNA_1 and VA-RNA_2, begins early after infection; but their transcription is increased dramatically late in infection, with up to 10^8 copies of VA-RNA_1 being present in the cytoplasm. Mutant adenoviruses lacking the genes for the VA-RNAs grow less well than the wild type virus, apparently as a result of defective initiation of viral protein synthesis in the mutants (Thimmappaya *et al.*,1982). The molecular basis for this defect has not been completely elucidated, but there is a strong suggestion that VA-RNA_1 is able to block one of the main

anti-viral activities induced by interferon. Interferon activates a double stranded RNA dependent protein kinase (DAI), which phosphorylates a serine residue on the α-subunit of eIF-2. This has the effect of blocking the initiation of protein synthesis during viral infection. The VA-RNAs are able to prevent this process by binding to the kinase, blocking its activation by interferon and so allowing initiation of protein synthesis to proceed (Schneider *et al.*,1984; Mathews and Shenk,1991). Other mechanisms to control translation also operate late in infection. The cap binding complex (eIF-4F) is inactivated by dephosphorylation which confers a selective advantage for the translation of late viral mRNAs. This is because, unlike most host cell mRNAs, their poly(A) extensions enable them to bind to ribosomes in the absence of active eIF-4F (Huang and Schneider,1991). Recently, the L4, 100-kDa, chaperone protein has been shown to bind to RNA and there is a suggestion that it plays a role in translational selectivity of late viral proteins (Hayes *et al.*,1990).

5.2 Transformation and Adenovirus Oncogenes.

Adenovirus transformed cells contain a portion of the viral DNA integrated into their DNA which includes the E1A and E1B regions. They constitutively express the E1A and E1B proteins; but interestingly the IX transcription unit which is contained within E1B is inactive (Sharp,1984; Boulanger and Blair,1991). All adenovirus serotypes are able to transform cells, but not all are oncogenic. Comparisons between the sequences of the E1A regions of human Ad12 and simian SA7 viruses, which are both oncogenic, reveals a conserved stretch of 7 amino acids in $E1A_{13S}$ not present in the non-oncogenic viruses. The suggestion is that this sequence may be involved in tumorigenicity (Kimelman *et al.*, 1985).

The binding of E1A proteins to a number of cellular proteins including cyclin A and the retinoblastoma susceptibility gene product (pRB) takes place through domains that are essential for transformation and so has created much interest (**fig.7**). pRB is an anti-oncogene that binds to the cellular transcription factor E2F and so controls its activity. E1A proteins are able to interact with pRB in a way that nullifies its anti-oncogenic properties (Whyte *et al.*,1988). The association between $E1A_{12S}$ and cyclin A, along with evidence that a serine/threonine protein kinase forms a complex with $E1A_{12S}$, has led to the suggestion that a cell cycle regulated kinase such as $p34^{cdc2}$ may be important in E1A mediated transformation (Herrmann

et al.,1991). The E1A region is necessary; but not sufficient, for transformation and requires the E1B 55-kDa protein for the cells to become fully transformed (van den Elsen *et al.*,1983). The cooperative role of E1B can also be carried out by a number of oncogenes including *ras* (Ruley,1983). The function of the 55-kDa protein in transformation is thought to involve its interaction with the cellular anti-oncogene product p53 and may serve to release the cells from controlled growth. The E1B 19-kDa protein has been shown to associate with vimentin-containing intermediate filaments and with the nuclear lamina. The resulting disruption of the cell is thought to be an important part of the transformation process (Boulanger and Blair, 1991).

1.6 ADENOVIRUS DNA REPLICATION

The onset of viral DNA replication takes place at 6 hours post infection, is coincidental with a virtual shutdown of cellular DNA synthesis and forms the natural division between the early and late transcription events described in section 1.5a. The process continues until 24 hours post infection when up to 100 000 copies of the viral DNA are present in each cell. Only about 20% of these genomes are, however, ultimately packaged into progeny virions (Pina and Green,1969; Green *et al.*,1970).

Adenovirus DNA replication has been studied extensively using both *in vivo* and *in vitro* systems (Pearson and Hanawalt,1971; Challberg and Kelly,1979; Hay and Russell,1989). Electron microscopy studies of replication intermediates, isolated from Ad2 infected cells, led to the proposal of a strand displacement mechanism for replication (Flint *et al.*,1976). **Fig.9** shows a model for one cycle of adenovirus replication which operates by a protein primer mechanism and can be divided into 2 stages. The first stage, termed initiation, involves the formation of an ester linkage between the phosphate of a dCMP base and the β-hydroxyl group of Ser_{580} in the pre-terminal protein (pTP). The 3' hydroxyl group of the resulting pTP-dCMP then acts as a primer for the synthesis of a new strand of DNA using one strand of the parent DNA as a template. This process is termed elongation. The non-template strand of the parent DNA is displaced and forms a 'pan handle' circular structure, resulting from the annealing of the inverted terminal repeats present at each end of the single strand of DNA (Challberg *et al.*,1980; Hay *et al.*,1984).

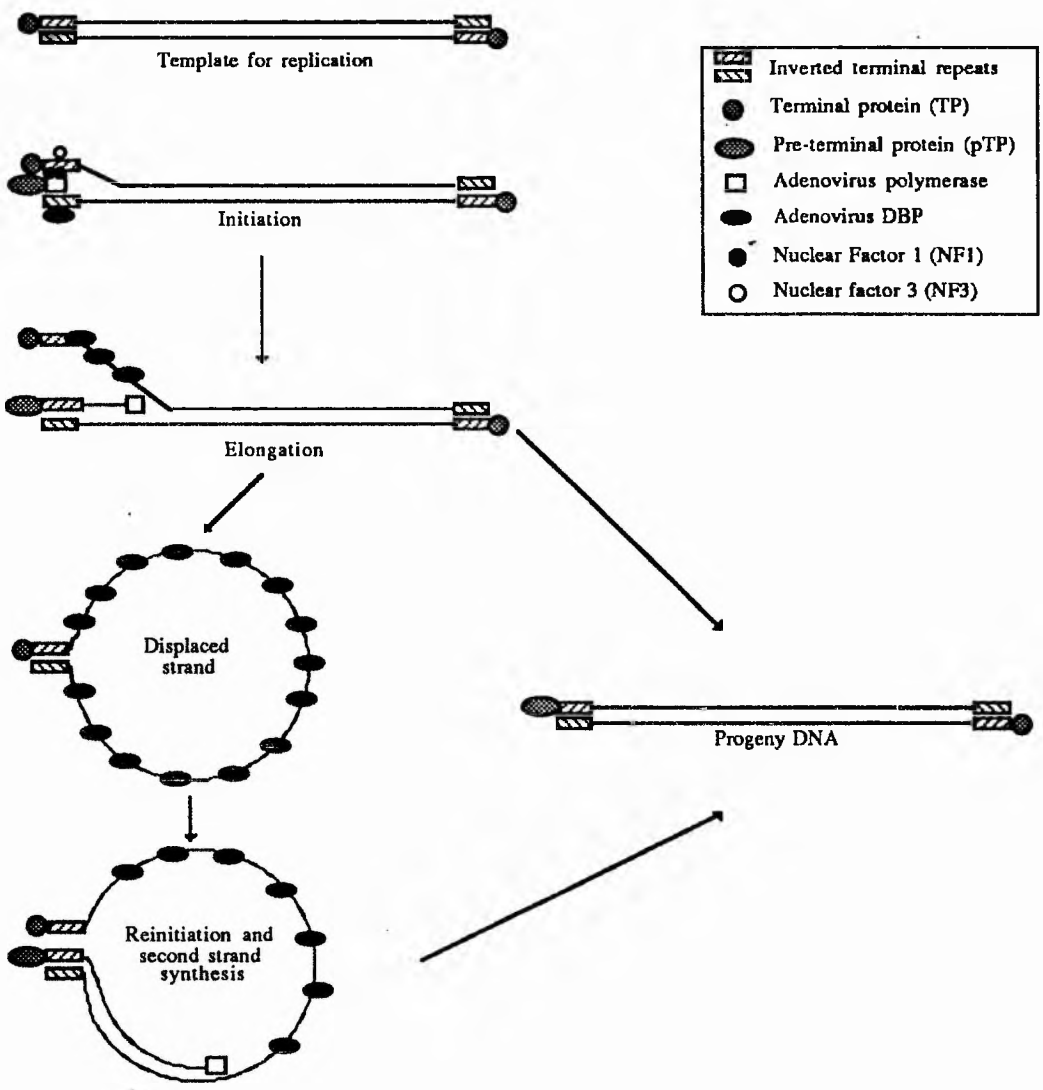


Fig. 9 Model for Adenovirus DNA Replication
A model for the replication of Ad2 DNA with schematic representations of pTP, TP, polymerase, DBP, NF1, NF3 and the inverted terminal repeats.

The inverted terminal repeats play a central role in replication with bases 1-18 being absolutely required for initiation. The structure of the DNA template *in vivo* is the subject of much controversy, but templates effective in *in vitro* assays include linearised plasmids containing bases 1-18, deproteinised adenovirus DNA, adenovirus DNA with the terminal proteins attached to each 5' end, viral cores and subviral particles (Challberg and Kelly,1979; Harris and Hay,1988; Leith *et al.*,1989). Interestingly, the efficiencies of these templates vary considerably in initiation reactions and although bases 1-18 are sufficient for initiation *in vitro*, the process is stimulated up to 200 fold by the presence of bases 19-50 (Hay *et al.*,1984). Furthermore, the presence of the terminal protein at the 5' end of the DNA considerably enhances the process (Tamanoi and Stillman,1982). The observation that subviral particles are able to act as efficient templates for initiation and elongation is noteworthy since subviral particles, prepared by dialysis against a hypotonic buffer, lack only the apical penton capsomers from their coat. This means that Ad2 DNA can act as the template for replication even in the presence of bound viral proteins (Leith *et al.*,1989). Indeed this work lends support to the model of the viral DNA proposed by Wong and Hsu (1989) shown in **fig.5**. This model suggests that the termini of the DNA are not supercoiled in the virus particle and so presumably are free to participate in DNA replication. It is also worth noting that the E1 and E4 regions are free and an extension of the work by Leith *et al.*(1989) is that subviral particles could act as templates for early transcription.

Two proteins are essential for the initiation of Ad2 DNA replication *in vitro*, the virus encoded pTP and polymerase. The process is significantly enhanced, however, in the presence of the 72-kDa DNA binding protein (DBP) and 2 cellular factors termed NFI and NFIII. The latter also serve as transcription factors and bind to sequences at 22-39 and 40-51 in the Ad2 DNA (reviewed in Kelly,1984; Hay and Russell,1989; Salas,1991). The involvement of these factors explains the importance of bases 19-50 in Ad2 DNA replication, as discussed above. In contrast, studies carried out on the replication of Ad4 DNA reveal that it requires only bases 1-18 at the origin of replication and that the presence of the NFI and NFIII binding sites have no effect on the efficiency of initiation of replication (Harris and Hay,1988).

The proteins involved in adenovirus DNA replication will be described in more detail in

chapter 4, which is concerned with the processing of the pTP by the virus encoded protease.

1.7 VIRION ASSEMBLY

Adenovirus assembly is a complex process the details of which are poorly understood. The large number of proteins involved means that it is difficult to establish whether a particular protein plays a direct or an indirect role in assembly. A protein may be indirectly involved if it is required for the transcriptional, translational or post-translational activation of a structural or scaffolding protein. A further problem is that most studies have involved the analysis of the protein/DNA composition of incomplete or mutant virus particles. Some of these intermediate particles will be incorrectly assembled and, therefore, not true precursors of the mature virus. Furthermore, the isolation and analysis of intermediate or mutant particles frequently involves harsh treatments such as high salt, detergents, urea, centrifugation and chemical crosslinking. These procedures may shear DNA and remove proteins associated with the virus or, alternatively, create artifactual associations.

The viral proteins are synthesised in the cytoplasm before being transported to the nucleus where assembly takes place. The uptake of proteins by the nucleus is extremely selective and it is believed that nuclear proteins contain a sequence that allows their specific accumulation in the nucleus (Dingwall and Laskey,1991). These sequences are rich in the positively charged amino acids, arginine and lysine, and the N-terminal sequence from the Ad2 fibre, MKKARPSEDTF, has been shown to be sufficient for nuclear localisation (Hong and Engler,1991). Nuclear localisation signals have been identified in a number of other adenovirus proteins including KRPRP in $E1A_{13S}$ and RLPVRRRRRRVP in the pre-terminal protein (Lyons *et al.*,1987; Zhao and Padmanabhan,1988). The nuclear transport of viral proteins lacking a nuclear targeting sequence may be facilitated by their association with a nuclear protein containing such a sequence. This has been shown in the case of the Ad2 polymerase, which is escorted to the nucleus by the pre-terminal protein (Zhao and Padmanabhan,1988).

The first step in adenovirus assembly is the formation of the hexon capsomers, which requires protein II (hexon monomer) and the L4 100-kDa protein. The 100-kDa protein binds to II in stoichiometric amounts while the latter is still attached to the polyribosomes. The

resulting II/100K heterodimers aggregate and trimers of II are folded together. The trimers (hexon capsomers) dissociate from the 100-kDa protein and are transported to the nucleus. Mutation studies suggest that the 100-kDa protein is involved in the transport of the hexon to the nucleus, but it is possible that this could be an indirect consequence of incorrect folding of the hexon by the mutant proteins (Cepko and Sharp,1983). It has also been suggested that the structural protein, pVI, is required to chaperone the hexon to the nucleus (Morin and Boulanger,1986). This is based on an established association between pVI and the hexon and the characterisation of a temperature sensitive mutant, H5ts147, the defect of which apparently maps to pVI (Everitt *et al.*,1973; Kauffman and Ginsberg,1976). On reaching the nucleus, the hexon capsomers appear to be able to self assemble to form empty icosahedral shells. In Ad2 infections only 10% of the hexons synthesised are incorporated into mature virions and free hexons; but not virion associated hexons, have been found to contain a reactive serine residue that binds to diisopropylfluorophosphate (DFP). The suggestion is that the reactive serine might be required for the folding process (Devaux and Boulanger,1980).

After the construction of the hexon shells it is suggested that pVI, pVIII and pIIIa associate with the virions along with 50-kDa and 39-kDa proteins (D'Halluin *et al.*,1978). The 50-kDa and 39-kDa proteins do not appear to be present in the mature virions and it is thought that they are scaffolding proteins required for correct assembly. The L4 100-kDa protein has been isolated from intermediate particles and is also believed to function as a scaffolding protein. Furthermore, a 90-kDa protein, related to the 100-kDa protein, has been found in intermediate particles and there is a suggestion that it is the product of limited proteolysis of the 100-kDa species, perhaps by the adenovirus protease (Morin and Boulanger,1986).

The next step in assembly is the insertion of the viral DNA and core proteins into the empty capsids. Some reports suggest that the naked DNA enters first, followed by the core proteins, but it would seem more likely that the basic core proteins would be required to neutralise the charge of the DNA prior to its packaging in the virion. The left end of the viral genome including the E1A region has been shown to be inserted into the virions first and nucleotide sequences at 200-400 base pairs appear to be essential for the packaging of the genome (Hammarskjold and Winberg,1980). It is assumed that specific proteins must bind to

these sequences (Grable and Hearing,1990). Mutation studies have revealed that 2 proteins, the L1 52/55-kDa protein and IX are essential for the incorporation of the viral DNA into the particles. A temperature sensitive mutant in the 52/55-kDa phosphoproteins was isolated and shown to be normal with regards to viral replication; but defective in particle assembly at non-permissive temperatures. Although the 52/55-kDa proteins have not been detected in mature virus particles, this does not rule out the possibility that they are present at low copy numbers (Hasson *et al.*,1989). Protein IX which is a cementing protein located on the external side of the coat is required for stability of the mature virus. Initially it was thought that it was not required for assembly. More recent mutant studies, involving complementation assays, however, suggest that IX is required for the correct insertion of the viral genome (Colby and Shenk,1981; Ghosh-Choudhury *et al.*,1987). The role of the penton capsomers and particularly of the fibre in assembly is controversial. One theory is that proteins III and IV self assemble to form the pentons, which don't associate with the virus until after the insertion of the DNA and core proteins (Falgout and Ketner,1988). Other reports suggest, however, that the fibre is required at earlier stages in the assembly process (Chee-Sheung and Ginsberg, 1982). It has also been proposed that mutants in pTP, pol and E4 proteins are defective in assembly (Horwitz,1990a).

In conclusion, many proteins are involved in assembly; but in most cases their exact roles are poorly understood. Once assembled the virions undergo a morphological change that converts them from young to mature virions. The antigenicities of the young and mature virus particles differ considerably and although a virally encoded protease is thought to be solely responsible for the process it cannot be ruled out that other proteins are involved.

1.8 THE ADENOVIRUS PROTEASE AND MATURATION.

The final step in the production of infectious virus particles is carried out by the subject of this thesis, the adenovirus protease, and involves limited proteolysis of a number of viral proteins (Anderson *et al.*,1973). These include the pre-terminal protein (pTP), pVI, pVII, pVIII, IIIa and the L2 11-kDa protein (Boudin *et al.*,1980, Tremblay *et al.*,1983, Weber and Anderson,1988). It is possible that other viral, and perhaps cellular proteins, are cleaved by the adenovirus protease during the infectious cycle. Indeed, Morin and Boulanger (1986)

suggest that the L4 100-kDa scaffolding protein is processed by the protease at its N-terminus to give a 90-kDa protein (see section 1.7). The cleavage sites in pVI and pVII have been identified by amino acid sequencing and, based on this information, cleavage sites have been proposed in pVIII and the pTP (Akusjarvi and Persson,1981; Sung *et al.*,1983; Smart and Stillman,1983). Tremblay *et al.* (1983) suggest that the enzyme is specific for glycine, alanine containing bonds, but concede that there must be other determinants in specificity for such a selective enzyme. A detailed examination of the substrate specificity of the Ad2 protease will be carried out in chapter 1.

The isolation of a temperature sensitive mutant, Ad2ts1, defective in processing at non-permissive temperatures, proved a major discovery and suggested that a virus encoded enzyme was responsible (Weber,1976). Ad2ts1 accumulates the precursor proteins and although viral particles are assembled at non-permissive temperatures they are not infectious. The mutation in Ad2ts1 has been mapped to a gene in the L3 region, coding for a 23-kDa protein, and gives rise to a Pro -> Leu change at residue 136 (Yeh-Kai *et al.*,1983). It is, therefore, generally accepted that the L3 23-kDa protein is the protease (Krausslich and Wimmer,1988). The gene has been sequenced in 7 different human serotypes and is remarkably well conserved suggesting that it must have an important function in the virus (Houde and Weber,1990a). It bears no obvious resemblance, however, to any established viral or cellular proteases.

Limited inhibitor studies have shown that the enzyme is inhibited by PMSF and TPCK and based on these it has been reported that the 23-kDa protein is a chymotrypsin-like serine protease (Bhatti and Weber,1979a; Tremblay *et al.*,1983). It has been suggested further that His_{54}, Asp_{102} and Ser_{187} could form the catalytic triad of such an enyzme despite the fact that there are no homologies between the 23-kDa protein and chymotrypsin in the surrounding residues (Houde and Weber,1987). Chapter 2 aims to carry out a more rigorous examination of the classification of the adenovirus protease.

The observation by Chatterjee and Flint (1987), that a 23-kDa phosphoprotein is processed to give a 19-kDa protein led to the conclusion that the adenovirus protease is phosphorylated and autocatalytically cleaved. The possibility that the activity of the enzyme is

controlled by phosphorylation and/or proteolysis is an interesting one which will be explored further in chapters 3 and 4.

The location of the protease activity has been shown to be nuclear and activity can also be extracted from purified virions. The 23-kDa protein is, however, present at a low copy number and has not been visualised by SDS-PAGE (Anderson,1990). The timing/ coordination of the cleavages of the different substrates is another aspect that has not been addressed. Is it that the protease is not required before the final maturation step as suggested by Ad2ts1 studies or is it that in the absence of active protease defective particles are assembled? In other words in wild type infections do any or all of the cleavages take place before the formation of the virions? With the exception of the pTP, which is required for replication, all the other substrates have been assigned structural roles. Chapter 4 attempts to study in more detail the processing of the pTP and the role that it plays in the lytic cycle. A more detailed description of the properties of the adenovirus protease will be given in the introductions to chapters 1-4.

Like many other viruses the adenovirus codes for a protease the activity of which is essential for infectivity. In section 2 the importance of proteolysis in the replicative cycles of other viruses will be considered.

2 LIMITED PROTEOLYSIS OF VIRAL PROTEINS.

Most viruses rely on the action of specific proteases at some point in their infectious cycle. These proteases may be cellular, as in the case of influenza, but in many instances are coded for by the viruses themselves. The adenovirus is the only example of a DNA virus that is known to code for its own protease; but limited proteolysis has also been shown to take place in vaccinia, African swine fever and herpes viruses (section 2.6). In contrast, numerous RNA viruses code for their own proteases and some of these have been very well characterised.

The positive strand RNA viruses translate their genomes as polyproteins and require proteases to divide these up into functional units. This is usually a prerequisite for viral replication and particle assembly. Before considering the processing of specific viral proteins,

it is necessary to define some terminology that will be used when referring to the polyproteins of RNA viruses. *Cis* cleavages are those carried out when the protease and substrate are part of the same polypeptide, whilst *trans* cleavages take place when the protease and scissile bonds are in separate molecules. *Structural* proteins form part of the mature virus particles, whereas *non-structural* proteins are enzymes or other gene products that are not packaged in the virion.

2.1 ORTHO- AND PARAMYXOVIRUSES

Ortho- and paramyxoviruses are negative strand, enveloped viruses that include the influenza and respiratory syncytial viruses, respectively. Although these viruses have not been shown to code for their own proteases they are included as the best characterised examples of the processing of viral proteins by cellular proteases (Kingsbury,1990; Chanock and McIntosh,1990; Kawahara *et al.*,1992). The ortho- and paramyxoviruses share a number of common features, including an absolute requirement for the cleavage of a surface viral glycoprotein before membrane fusion can take place. The hemagglutinin (HA) protein of influenza A is by far the best defined of these glycoproteins and so its processing will be described as a typical example.

The HA protein mediates the initial attachment of the influenza A virus to its receptor and after uptake into the cell, initiates the fusion of the virus envelope with an intracellular membrane. The protein is synthesised as a 566 amino acid precursor and its insertion into the endoplasmic reticulum is directed by a 16 amino acid signal sequence, which is later removed by a signalase. After assembly of the virus, the HA protein is cleaved at the C-terminal side of Arg_{328} by a cellular trypsin-like enzyme to give HA_1 and HA_2, which are connected by a disulphide bridge (**fig.10**). The resulting conformational change exposes a hydrophobic

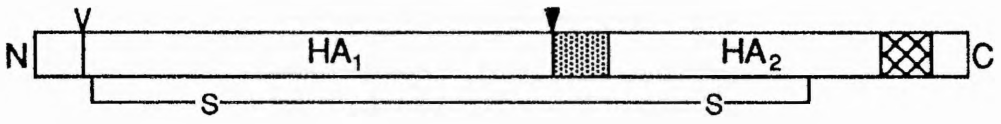


Fig. 10 Influenza Virus Haemagglutinin

Schematic diagram of the haemagglutinin glycoprotein from influenza virus showing the position of the cleavage necessary for the development of infectivity (▼). Cleavage at this site exposes a hydrophobic fusion domain (▨). The open arrow indicates the site of cleavage of a cellular signalase which removes the signal peptide. The shaded area close to the C-terminus represents the membrane spanning region.

sequence at the N-terminus of HA_2 which is capable of directing membrane fusion (Kingsbury,1990; Klenk *et al.*,1975; Lazarowitz *et al.*,1973). The cleavage sites in the other ortho- and paramyxovirus glycoproteins are relatively well conserved and are rich in basic amino acids. In some of the viruses the maturation cleavage of the surface glycoprotein is an extracellular event.

Interestingly, it has been noted that combined influenza/ *S.Aureus* pneumonia is approximately 3 times more common than primary viral pneumonia (Tashiro *et al.*,1987). The reason for this appears to be that certain strains of *S.Aureus* secrete proteases that accelerate the processing of the HA protein and so increase the amount of infectious virus present.

2.2 PICORNAVIRUSES

The picornaviruses are positive strand RNA viruses that have been divided into 4 genera termed rhino-, entero-, cardio- and aphtho- viruses. They are non-enveloped, icosahedral particles and are the causative agents of a range of diseases including polio, hepatitis A, the common cold, and foot and mouth disease. Their RNA genomes are very small, comprising less than 8000 bases, and contain a single open reading frame that codes for a large polyprotein of molecular weight 200-300-kDa. These polyproteins are then cleaved to release all the structural and non-structural proteins of the virus (Jacobson and Baltimore,1968; Wimmer,1989). Proteolytic processing was first observed in the enterovirus, polio and as a result all the cleavage sites in the poliovirus polyprotein and the functions of many of its individual proteins have been well defined (Kitamura *et al.*,1981; Nicklin *et al.*,1986). Subsequent work on the other picornaviruses has revealed that their genome organisations and processing sites are very similar to those found in poliovirus (reviewed in Krausslich and Wimmer,1988).

The organisation of the poliovirus polyprotein is shown in **fig.11a** and the positions of the cleavage sites are marked. The processing is carried out by 3 distinct proteases and 2 of these are the viral proteins 2A and 3C, which have been well characterised. The first processing event is the separation of the structural and non-structural proteins and this co-translational cleavage takes place at a Y-G bond at the C-terminus of VP1. This is an autocatalytic event carried out by the 17-kDa, 2A protein which cleaves at its own N-terminus.

In poliovirus this cleavage is a prerequisite for viral assembly and the formation of replication complexes. The only other site in the polyprotein that is cleaved by 2A is a second Y-G bond in the 3D polymerase protein. This is not essential for virus replication, but is thought to provides a means of regulating the activities of the 3C protease and 3D polymerase (Lee and Wimmer,1988). A further function performed by the 2A protease is the shutting off of host cell protein synthesis. The precise mechanism for this remains to be elucidated, but it has been shown that 2A and the eukaryotic initiation factor 3 (eIF3) act together to cleave the p220 subunit of eIF4 (Wyckoff *et al.*,1990). The latter is required for the binding of capped mRNA to the ribosome and its cleavage results in the inhibition of host cell protein synthesis. The poliovirus RNA is not capped at its 5' end and can be translated by a cap independent mechanism in the absence of p220 (Sonenberg,1991). Whether 2A or eIF3 carries out the cleavage of p220 is a matter of controversy and until recently it was thought that 2A played only an indirect role (Krausslich *et al.*,1987). Mutational analysis of the 2A protein, however, reveals that its proteolytic activity correlates with its ability to induce the cleavage of p220 suggesting that 2A is directly responsible (Hellen *et al.*, 1991). The function of eIF3 may be to bind to the p220 protein and to present it to the protease in the correct conformation for cleavage (Wyckoff *et al.*,1990). Thus it appears that the 2A protease of poliovirus plays 2 essential roles in the replicative cycle of the virus, one involving the separation of the structural and non-structural viral proteins and the other, the cleavage of the cellular protein p220.

Eight of the 11 cleavages of the poliovirus polyprotein are catalysed by the 3C protease (3C), either alone or in association with the 3D polymerase (3CD), as shown in **fig.11a**. Although all the cleavage sites are Q-G bonds, the separation of the capsid proteins VP0, VP3 and VP1 is carried out by 3CD and not by 3C. It has been proposed that elements in 3D are required to bind to the capsid protein precursor and so increase the affinity of 3C for its structural substrates. This may be necessary as a result of the initial separation of the structural and non-structural proteins catalysed by 2A. The 6 remaining Q-G sites that link the non-structural proteins are readily cleaved by the 3C protease. It has been suggested that the different affinities of the 3C protease for the various substrates may be important in the regulation of the lytic cycle (Hellen *et al.*,1989). Like 2A, the 3C protease has also been

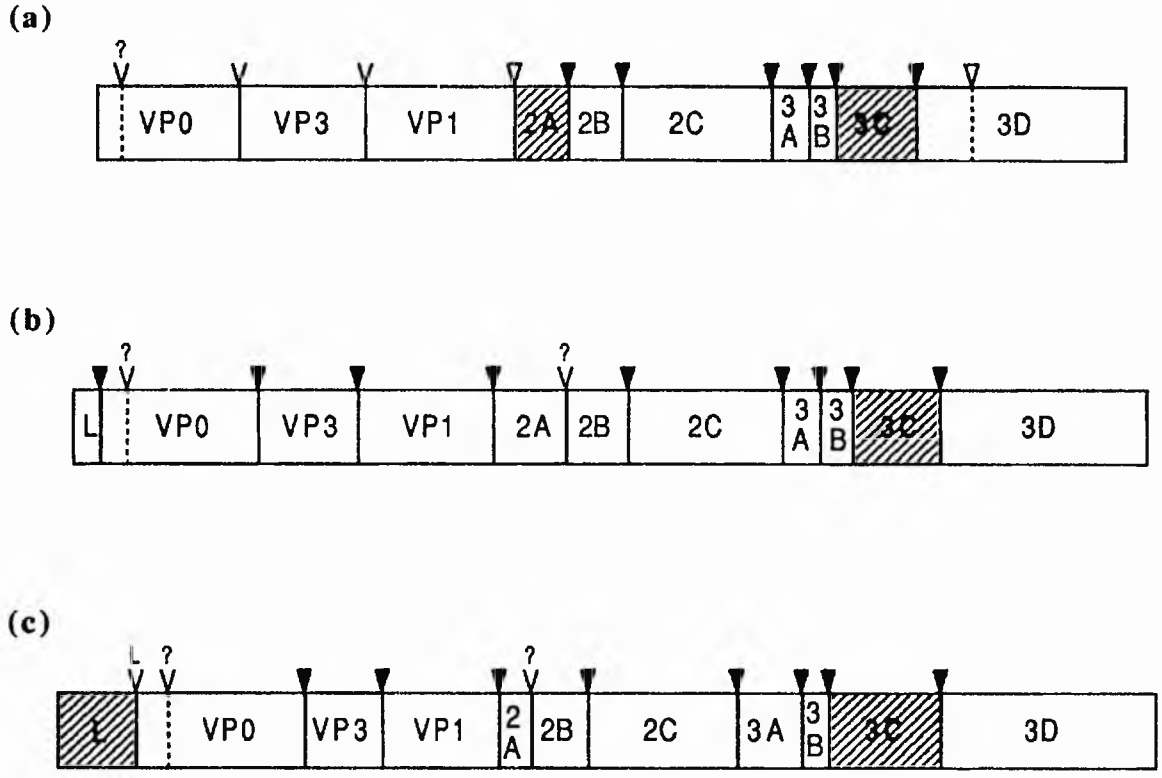


Fig. 11 Polyproteins of the Picornaviridae

Schematic diagrams showing the organisation of the polyproteins of (a) entero- and rhino-viruses; (b) cardioviruses; (c) aphthoviruses. The 2A, 3C and Leader proteases are shaded and arrows indicate the scissile bonds cleaved by 2A (▽) , 3C (▼) and 3CD (V). A question mark is shown above cleavage sites where the protease has not been identified while L indicates cleavage by the Leader protease.

associated with host cell control. Clark *et al.* (1991) have shown that 3C is required for the inhibition of host cell RNA polymerase III during poliovirus infection. The effect is mediated by the transcription factor, TFIIIC, which is converted from an active to an inactive form. The suggestion is that 3C is responsible for the cleavage of TFIIIC; but proof is still required.

The final processing event in polio is the cleavage of the structural protein VP0 at an N-S bond to give VP4 and VP2 and this takes place after assembly of the virus particles. At first it was assumed that this was an autocatalytic event with Ser_{10} of VP2 being an active site residue and the viral RNA playing a catalytic role (Arnold *et al.*,1987). Mutants with alanine or cysteine at position 10 of VP2, however, were able to process VP0 correctly and the mechanism by which the cleavage is carried out remains in doubt (Harber *et al.*,1991).

The organisation and processing of the other entero- and rhino- viruses are to all intents and purposes identical to those described for poliovirus, but there are interesting differences in the aphtho- and cardio- viruses (**fig.11b,c**). Viruses from these genera contain leader proteins at the N-termini of their polyproteins and do not contain a functional 2A protease. In the case of the cardiovirus, encephalomyocarditis (EMC), a 2A protein is present; but it is unrelated to the proteases from entero- and rhino- viruses and has no proteolytic activity. The primary co-translational cleavage that separates the structural from the non-structural proteins in EMC is between 2A and 2B, not at the N-terminus of 2A as in polio. The protein responsible for this cleavage remains to be identified and as yet there is no evidence that a viral protein is involved. The 3C protein of EMC is a protease and has been shown to cleave at least 9 bonds in the polyprotein. These include the bond between the leader protein and VP0 and the VP1/2A site, cleaved by 2A in polio (Palmenberg and Parks,1989). As with the other picornaviruses, the mechanism by which the final maturation cleavage of VP0 is carried out remains to be established.

In the aphthovirus, foot and mouth disease virus (FMDV) the 2A protein is absent and in its place is found a 16 amino acid peptide. The primary cleavage takes place at the N-terminus of 2B; but as with EMC the protease responsible has not been identified. One proposal is that the event is autocatalytic and is dependent on sequences at either side of the scissile bond (Krausslich and Wimmer,1988). Ryan *et al.* (1991) go further to propose that

the cleavage is catalysed by a 19 amino acid peptide spanning the 2A region of the FMDV polyprotein. Although this cannot be ruled out, perhaps more likely is the possibility that the 19 amino acid sequence contains a conserved structural domain that is cleaved by a host protease in a similar way to the HA protein of influenza (section 2.1). The removal of the leader protein from the FMDV polyprotein, on the other hand, does appear to be an autocatalytic event with the leader protein itself being the protease. Interestingly, it has been suggested that the L protein of FMDV, like the 2A protease of poliovirus, is responsible for the cleavage of the p220 eIF4 protein and the resulting shut off of host cell protein synthesis (Devaney *et al.*,1988).

The substrate specificities and classification of the picornavirus proteases will be discussed later in chapters 1 and 2 respectively.

2.3 POSITIVE STRAND RNA PLANT VIRUSES

A group of positive strand RNA viruses including the como-, poty and nepo viruses have very similar genome organisations to the picornaviruses and employ similar strategies with regards to polyprotein processing (Wellink and van Kammen,1988; Riechmann *et al.*, 1992). The comovirus, cowpea mosaic virus (CPMV), and the potyvirus, tobacco etch virus (TEV), have been particularly well studied and will be described here.

CPMV has a bipartite genome consisting of 2 positive strands of RNA termed B-RNA and M-RNA. Each has a 3-kDa protein, VPg, covalently attached to its 5' end, a poly(A) tail at its 3' end and contains a single, large open reading frame that is translated to give a polyprotein. The M-RNA codes for the capsid proteins, whilst the B-RNA codes for the non-structural proteins (Wellink *et al.*,1989). It is the non-structural polyprotein that is similar in organisation to the picornaviruses and the 24-kDa sequence in the corresponding position to 3C contains the proteolytic activity (**fig.12a**). Sequence alignments reveal that the 24-kDa protein is homologous to the 3C proteases of picornaviruses (Argos *et al.*,1984). The 24-kDa protein is responsible for 6 cleavages in the viral polyprotein, 2 in the M-protein and 4 in the B-protein. Whilst the separation of the non-structural B-proteins only requires the 24-kDa protein, complete processing of the M-protein also requires the 32-kDa non-structural protein (Vos *et al.*,1988). The 32-kDa protein is not thought to be a protease, but instead appears to

act as a cofactor in a similar fashion to 3D of polio (section 2.2).

TEV has a positive strand RNA genome of around 10 000 bases that includes a single open reading frame with the potential to code for a 350-kDa polyprotein and has a 6-kDa viral protein attached to its 5' end (Dougherty and Carrington, 1988).

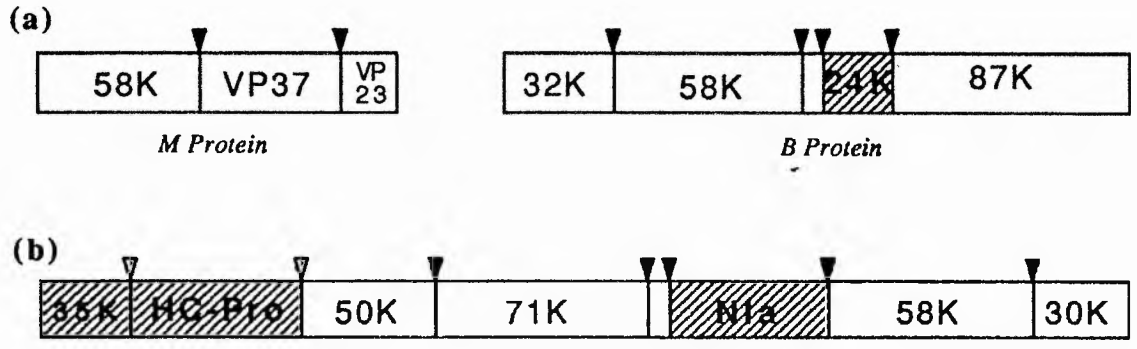


Fig. 12 Polyproteins of Plant Viruses
Schematic diagrams showing the organisation of the polyproteins from (a) cow pea mosaic virus (CPMV) and (b) tobacco etch virus (TEV). The proteases are shaded and the filled arrows indicate the scissile bonds cleaved by the 24K and NIa proteases. The TEV 35K and HC-Pro proteases carry out autocatalytic cleavages at their C-termini.

The organisation of the polyprotein is similar to that of the picornaviruses and is shown in **fig.12b**. It is processed by 3 distinct proteases, all of which are believed to be virus encoded and are termed NIa, HC-Pro and 35-kDa proteins. NIa is a 49-kDa multifunctional protein, with its C-terminal 241 amino acids containing the protease activity (Dougherty and Parks,1991). The N-terminal 189 amino acids are thought to be required for RNA replication and it has been proposed that NIa is cleaved *in vivo* to give 21-kDa and 27-kDa proteins. The NIa protein would appear, therefore, to be equivalent to 3BC in the picornavirus polyproteins (**fig.11**). The C-terminal domain of NIa is related to the 3C proteases and catalyses the cleavages at 5 sites in the TEV polyprotein. It has an extended substrate specificity recognising at least 7 amino acids (Dougherty and Parks,1989).

The 87-kDa N-terminal portion of the TEV polyprotein is removed by an autocatalytic event and is then further processed to give a 35-kDa protein and the 52-kDa HC-Pro. The HC-Pro domain of the 87-kDa protein is responsible for the initial cleavage, whilst the 35-kDa N-terminal domain is thought to carry out the second cleavage at its C-terminus (Carrington *et*

al.,1990; Verchot *et al.*,1991).

2.4 TOGAVIRUSES AND FLAVIVIRUSES

The toga- and flaviviruses are enveloped animal viruses containing a single positive strand RNA genome that is encapsidated by a viral protein shell. They have in common their ability to replicate in and be transmitted by arthropods such as the mosquito. The Togaviridae comprises 2 genera, namely the alpha- and rubiviruses whilst the Flaviviridae includes the flavi- and pestiviruses (Schlesinger and Schlesinger,1990). Proteolytic processing of the polyproteins of the alphaviruses, Sindbis and Semliki Forest virus (SFV), the flavivirus, yellow fever virus (YFV) and the pestivirus, bovine viral diarrhoea virus (BVDV) have all been shown to require virally encoded proteases and will be discussed in turn below. It is likely, however, that proteolytic processing is involved in the replication of all the members of the Toga- and Flaviviridae (Krausslich and Wimmer,1988).

Alphaviruses - Sindbis Virus.

The alphavirus, Sindbis, is by far the best characterised of the Toga- and Flaviviridae in terms of its processing. It has an 11 700 base RNA genome that is capped and polyadenylated and the organisation of the genome is shown in **fig.13.** It codes for 3 polyproteins, p230 and p250 which include the non-structural proteins and are translated directly from the genomic RNA and p130 which comprises the structural proteins and is translated from subgenomic mRNA. p230 and p250 are translated from the same initiation codon and in most cases translation is terminated at an opal codon just downstream from the nsP3 gene to give the p230 polyprotein. The p250 polyprotein which includes nsP4 results from read through of the termination codon (Strauss *et al.*,1984). The non-structural polyproteins p230 and p250 are cleaved by the virus encoded protease nsP2 to give nsP1, nsP2, nsP3 and nsP4. Although nsp2 carries out the processing, the other non-structural proteins are able to influence the order in which the different sites are cleaved. The regulation of processing is achieved by different affinities of the nsP2 protein for the 3 cleavage sites. It has been shown that whilst all nsP2 containing polyproteins are able to cleave the nsP1-nsP2 site in *trans*, the nsp2-nsP3 bond can only be cleaved if the nsP1 protein has first been removed. *Trans* cleavages at the nsP3-nsP4 site can only be carried out by nsP2 if nsP3 is attached to its C-terminus (de Groot *et al.*,

1990). This would explain why nsP4 is present at early times of infection, but that late in infection nsP3-nsP4 accumulates as a result of incomplete processing. It has been proposed that nsP4, the polymerase, is responsible for the synthesis of minus strand RNA early in infection. Four hours post infection this synthesis is switched off and the subgenomic RNA is synthesised, possibly by nsP3-nsP4. Thus in the case of the sindbis virus, the order in which the non-structural proteins are cleaved appears to play an important regulatory role in the lytic cycle of the virus (de Groot *et al.*,1990).

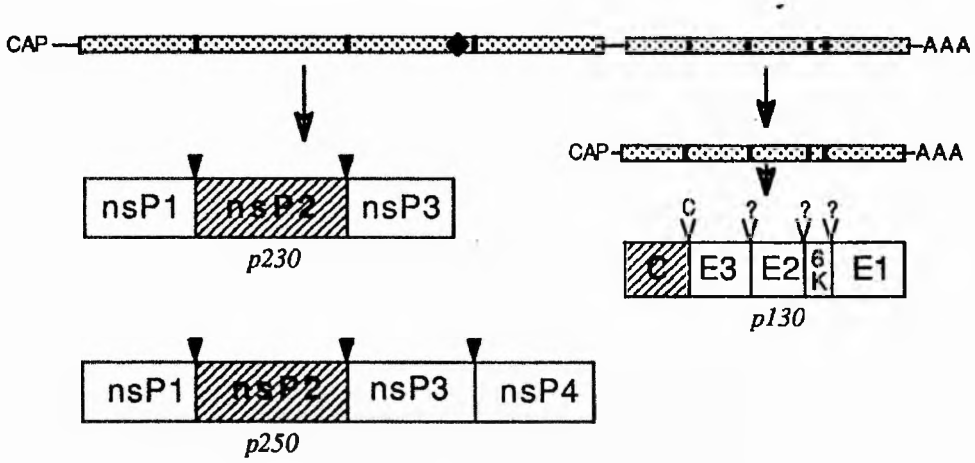


Fig. 13 Sindbis Virus Genome and Polyprotein Organisation

A schematic diagram showing the Sindbis virus genome and the organisation of the poyproteins containing the non-structural and structural proteins. The polyproteins coding for the nonstructural proteins (p230 and p250) are translated directly from the positive strand RNA genome while the structural proteins (p130) are translated from sub-genomic RNA. The nsP2 protease is shaded and its cleavage sites are indicated by the filled arrows. The capsid protein (C) carries out an autocatalytic cleavage at its C-terminus. Question marks indicate sites thought to be cleaved by cellular proteases.

The structural proteins of sindbis are also separated by proteases and the first event is the removal of the N-terminal capsid protein. This is an autocatalytic event, with the active site residues being identified as His_{141}, Asp_{163} and Ser_{215} in the capsid protein (Hahn and Strauss,1990). The structure of the C-terminal, protease-containing domain of the protein has been successfully elucidated by X-Ray crystallography and comparisons with cellular enzymes have confirmed that it is a chymotrypsin-like serine protease (Choi *et al.*,1991). This is in keeping with the fact that it cleaves on the C-terminal side of a tryptophan residue which has been shown to remain in the active site cleft after catalysis. This led Choi *et al.* (1991) to

propose that the cleaved capsid protein is devoid of activity after the single autoprocessing events. This, however, conflicts with the observation that mutants which fail to cleave the capsid protein can be complemented by mutant viruses of a different phenotype, suggesting that *trans* cleavages are possible (Scupham *et al.*,1977). In any case, after cleavage, the capsid protein forms an icosahedral shell round the viral RNA and adopts a structural role. The protein binds to the RNA, probably through sequences in its N-terminal half, outwith the protease domain (Choi *et al.*,1991). Removal of the capsid protein exposes an N-terminal signal sequence in the PE2 glycoprotein which is responsible for the insertion of the polyprotein in the endoplasmic reticulum. The remainder of the processing events of the structural proteins are thought to be carried out by cellular proteases. The last of these, the maturation cleavage of the membrane glycoprotein PE2 to give E2 and E3, appears to be very similar to the cleavage of the influenza HA protein described in section 2.1 (Strauss *et al.*,1987).

Flaviviruses - Yellow Fever Virus.

Processing in the flavivirus polyproteins is not as well understood as in the alphaviruses; but nevertheless yellow fever virus (YFV) and dengue virus (DEN) have been shown to code for their own proteases. YFV contains a single strand positive RNA genome of about 11 000 bases which includes a single open reading frame encoding a 350-kDa polyprotein (**fig.14a**). The 3 structural proteins; the capsid protein (C), glycoprotein precursor (PrM) and the envelope protein (E), are located at the N-terminal end of the polyprotein and are thought to be separated by a cellular signalase located at the endoplasmic reticulum. Unlike the capsid proteins of the alphaviruses, the capsid protein of YFV does not appear to be a protease; but instead is thought to contain a membrane-spanning anchor at its C-terminal end. This would be consistent with its cleavage by a signalase (Strauss *et al.*, 1987). The PrM glycoprotein undergoes a maturation cleavage late in infection in a similar way to PE2 from sindbis and HA from influenza. The processing of the non-structural proteins of YFV requires the N-terminal third of the 70-kDa viral protein NS3 (Chambers *et al.*,1991). This confirms the prediction that NS3 is a protease made on the basis of sequence comparisons with cellular and viral proteases (Bazan and Fletterick,1989). The NS3 protein is

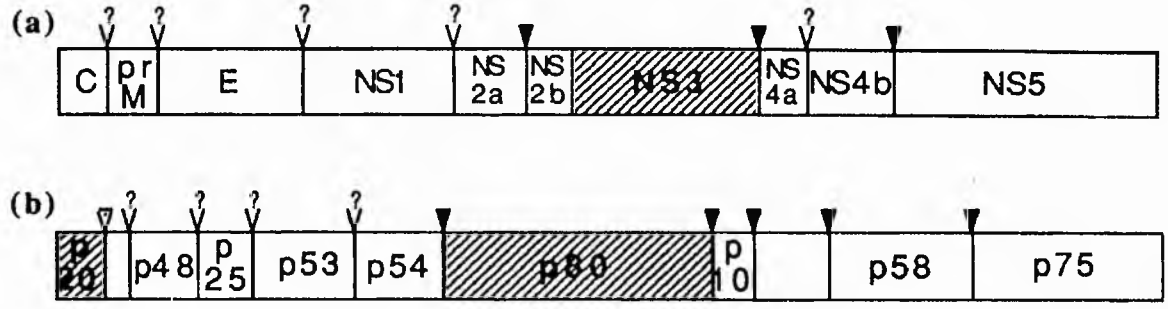


Fig. 14 Polyproteins of the Flavi- and Pesti-viruses

Schematic diagrams showing the organisation of the polyproteins in (a) Yellow Fever virus (YFV) and (b) Bovine Viral Diarrhoea virus (BVDV). The YFV NS3 protease and the BVDV p20 and p80 proteases are shaded and filled arrows indicate the scissile bonds cleaved by NS3 and p80. p20 carries out a single cleavage at its own C-terminus. A question mark indicates where cleavages are thought to be carried out by cellular proteases.

required for the separation of all the non-structural proteins of YFV with the exception of NS1-NS2a and NS4a-NS4b which are thought to be cleaved by cellular enzymes. The separation of the non-structural proteins also requires NS2 to be present either in *cis*, at the N-terminus of the NS3 protease, or in *trans*. It is postulated that the requirement for NS2 may be a means of regulating the processing at the different sites (Chambers *et al.*, 1991).

Pestiviruses - Bovine Viral Diarrhoea Virus.

Pestiviruses are responsible for a range of infections in cattle and sheep and contain a positive strand RNA genome of 12-13 000 bases with a single open reading frame coding for a polyprotein of around 350-kDa. The structural proteins are located at the N-terminal end of the polyprotein as can be seen in **fig.14b**. Sequence comparisons between the bovine viral diarrhoea virus (BVDV) polyprotein and other viral and cellular proteases revealed that 2 of the BVDV proteins contain putative protease sequences, the capsid protein (p20) and the non-structural p80 protein (Bazan and Fletterick,1989; Gorbalenya *et al.*,1989). Subsequent biochemical analysis has revealed that both these proteins are proteases responsible for the processing of the BVDV polyprotein. The p20 capsid protein is responsible for its own removal from the other structural proteins in a similar way to the capsid protein of Sindbis (Wiskerchen *et al.*,1991). The p80 protein carries out the 5 cleavages that separate the non-structural proteins including 2 *cis* cleavages that serve to cut itself out of the polyprotein. The remaining cleavages can be carried out in *trans* by p80, but the separation of p58-p75

appears to require an additional cofactor. Although the predicted catalytic triad is at the N-terminus of p80, the first two thirds of the protein are required for activity indicating that the central region of the molecule must play some role, perhaps a structural one (Wiskerchen and Collet, 1991). The remaining 4 processing sites, separating the structural proteins are thought to be cleaved by cellular proteases in a similar way to that discussed for YFV.

2.5 RETROVIRUSES

The retroviruses are enveloped single strand RNA viruses that are responsible for a wide range of serious diseases including acute T-cell leukemia and AIDS. They utilise both cellular and virus encoded proteases to separate their polyproteins into functional units. They are unusual amongst the RNA viruses in that the viral protease is not required for assembly of the virus particles; but instead carries out maturation cleavages after the virions have been formed. Nevertheless, activity of the retroviral proteases is a prerequisite for the formation of infectious particles.

On entering the cell the viral genome is reverse transcribed by a retroviral enzyme and the resulting double stranded DNA is integrated into the host cell chromosome. The genome organisations of the different retroviruses vary slightly and have been extensively reviewed, but basically speaking they all consist of 3 major genetic elements that are arranged in order 5'-*gag-pol-env*-3'. The *gag* genes code for the structural proteins present in the nuclear capsid, whilst *pol* codes for the viral replication enzymes and *env* for the membrane associated glycoproteins. Some of the retroviruses also include oncogenes in their genomes. For example Rous Sarcoma Virus (RSV) carries the *src* gene, but the oncogene products are not proteolytically processed so will not be included in this discussion. Following integration into the host cell chromosome, the viral DNA lies silent until the cell is activated. Upon activation the viral genes are transcribed by the host cell machinery to give at least 2 species of mRNA. The first is produced by splicing and contains the *env* gene which codes for the viral glycoproteins. These are translated as a polyprotein and are processed by cellular proteases in the Golgi apparatus as the polyprotein is transported to the plasma membrane. The second species approximates to the viral RNA and is translated to give 2 polyproteins termed *gag* and *gag-pol* (or *gag-pro-pol*). The *gag* and *gag-pol* polyproteins are translated from the same

initiation codon and whilst termination usually occurs downstream from *gag*, occasional read through takes place as a result of suppression of the termination codon. Alternatively in some retroviruses a frame shift is required to produce the *gag-pol* polyprotein (reviewed in Oroszlan,1989; Krausslich and Wimmer,1988).

The processing of the *gag* and *gag-pol* polyproteins is carried out by virally encoded enzymes and, as mentioned above, takes place after the assembly of the virion. Although protease genes have been identified in most of the retroviruses, the enzymes from Rous sarcoma virus (RSV), avian myeloblastosis virus (AMV) and human immunodeficiency virus types 1 and 2 (HIV-1, HIV-2) are by far the best characterised. For this reason, this discussion will concentrate on these viruses, with particular emphasis being put on recent work on the HIV-1 and HIV-2 proteases. Many of the properties attributed to these proteins are also likely to apply to their counterparts in other retroviruses.

Avian Sarcoma Leukosis Viruses.

The avian sarcoma leukosis viruses (ASLV), RSV and AMV are highly related and share the same genome organisation, with the protease (PR) being located at the C-terminal end of the *gag* polyprotein (**fig.15**). The inclusion of the protease in the nucleocapsid polyprotein is unusual among the retroviruses and means that in RSV and AMV the protease is synthesised in the same amounts as the structural proteins.

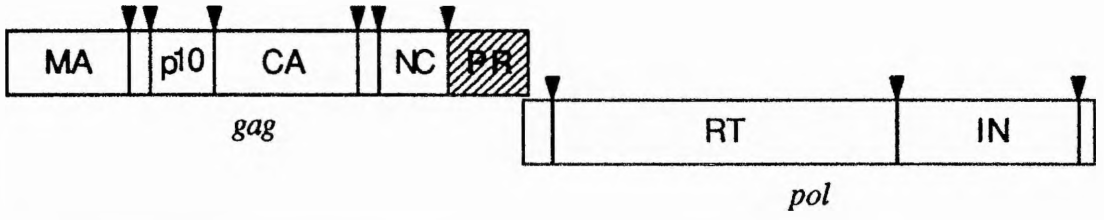


Fig. 15 RSV Polyproteins
Schematic diagram showing the organisation of the gag and gag-pol polyproteins of Rous sarcoma virus (RSV). The gag-pol is produced by readthrough of the stop codon at the end of the gag reading frame. The arrows indicate the bonds cleaved by the retroviral protease (shaded).

This property has proved useful and has enabled the purification of the p15 putative protease from virus particles. Incubation of the purified protein with precursor proteins *in vitro* provided proof that p15 was indeed the enzyme responsible for the proteolytic processing (von der Helm,1977). The p15 proteases from RSV and AMV have been expressed in *E.Coli*,

purified and crystallised in the presence and absence of substrate analogues (Miller *et al.*,1989a; Leis *et al.*,1989; Foundling *et al.*,1989). Elucidation of the crystal structures revealed that the enzymes are active as dimers and are members of the aspartic class of proteases. The requirement for dimerisation raises a number of interesting questions with regards to activation of the retroviral proteases during the lytic cycle. Most of the aspects of dimerisation will be discussed for the HIV proteases (see below); but the unusual genome organisation of RSV and AMV gives them some distinct properties that merit consideration. The inclusion of the protease coding region in the *gag* gene mean that it is present in a potentially active form in both the *gag* and *gag-pol* polyproteins. The *gag* and *gag-pol* polyproteins are synthesised at a ratio of 20:1, with the *gag-pol* protein being produced as the result of a ribosomal frame shift. Recently it has been shown that a recombinant AMV, producing only the *gag-pol* protein, is not capable of assembly thus revealing that the *gag* protein must direct the assembly process (Craven *et al.*,1991). Furthermore no proteolytic processing of the *gag-pol* proteins took place. It appears that only protease present at the C-terminus of the *gag* protein is active and that *gag-pol* must be cleaved in *trans*. Presumably the presence of *pol* in *cis* with the protease must put some conformational restraints on the protease that prevent activity (Stewart and Vogt,1991). It has also been shown that in the absence of proteolysis the viral reverse transcriptase (RT) is not active (Craven *et al.*,1991).

Human Immunodeficiency Viruses - HIV-1 and HIV-2.

As mentioned previously the proteases of HIV-1 and HIV-2 have been subject to a great deal of attention recently and, as a consequence, our understanding of retroviral proteases in general has been greatly increased. The organisation of the *gag* and *gag-pol* polyproteins is given in **fig.16**. In the case of HIV-1 and -2, the *gag-pol* protein is translated as a result of ribosomal frame shifting. Unlike RSV and AMV, the protease is not part of the *gag* polyprotein, but instead is at the N-terminal end of the *pol* region of *gag-pol*. This means that the protease is expressed at much lower levels and so cannot be isolated from virus particles in significant amounts. Nevertheless the 99 amino acid sequence making up an 11-kDa protein at the N-terminal end of *pol* has been shown to be the protease in HIV-1 and -2. It is responsible for 8 specific cleavages, 4 in each of the *gag* and *pol* regions. The primary processing event is

at the N-terminus of the protease sequence at (p6-PR) and since different sites are cleaved at different rates the cleavages may be carried out sequentially (Fitzgerald and Springer, 1991). Much sequence and structural homology exists between the 99 amino acid HIV-1 and -2 proteases and other retroviral proteases (Weber,1989).

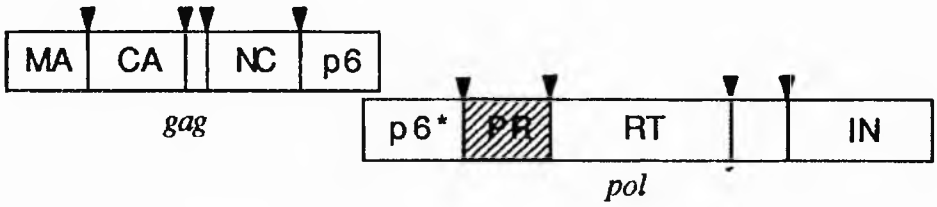


Fig. 16 HIV Polyproteins
Schematic diagram showing the organisation of the HIV gag and gag-pol polyproteins. Gag-pol results from a frameshift event within gag p6 resulting in p6*. The arrows indicate the bonds cleaved by the HIV protease (shaded).

The HIV proteases have been successfully cloned in an active form in *E.Coli* and when expressed as fusion proteins are capable of carrying out autocatalytic processing (Graves *et al.*,1988; Louis *et al.*,1991). Perhaps of even more significance is the fact that the HIV-1 protease has been chemically synthesised in its entirety by solid phase peptide synthesis and the resulting protein is active (Schneider and Kent,1988). The *E.Coli* expressed and chemically synthesised HIV-1 proteases have been successfully crystallised both alone and in the presence of a range of peptide substrates and inhibitors (Miller *et al.*,1989b; Navia *et al.*,1989; Wlodawer *et al.*,1989). The enzymes are functional as dimers and have structures which are remarkably similar to the RSV protease (**fig.17**). Prominent features include the active site cleft which contains the active site aspartyl residues and the 2 flaps which partially cover it. These flaps are mobile with a 7A^{o} shift taking place following the binding of the substrate (Gustchina and Weber,1990). The presence of the flaps may account for the extended substrate specificity of the protease. Complete mutagenesis of the HIV-1 protease has been carried out with each residue being substituted by at least one alternative amino acid and the mutants have been assessed for protease activity. The resulting activity profile is given in **fig.18** (Loeb *et al.*,1989). In addition to the active site residues a number of other residues are required for activity. These are involved in substrate binding or dimer formation or, alternatively, are essential for the structural integrity of the protein.

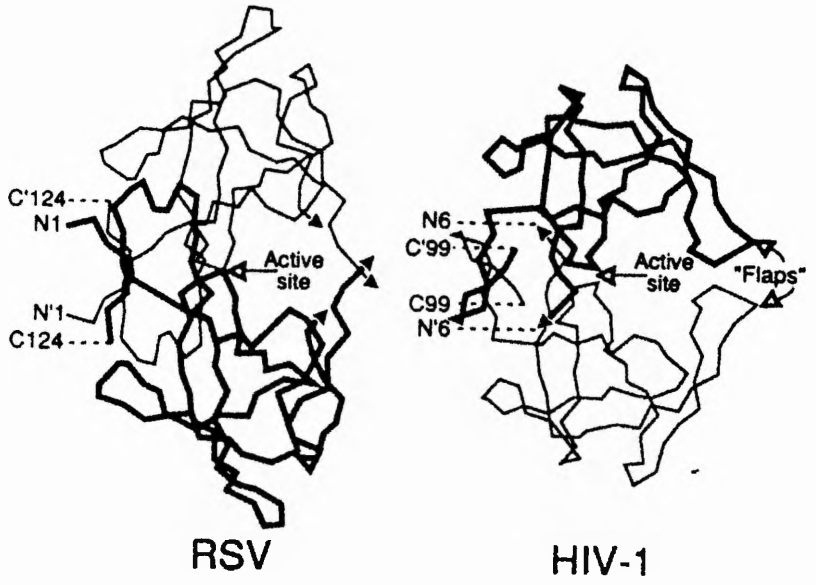
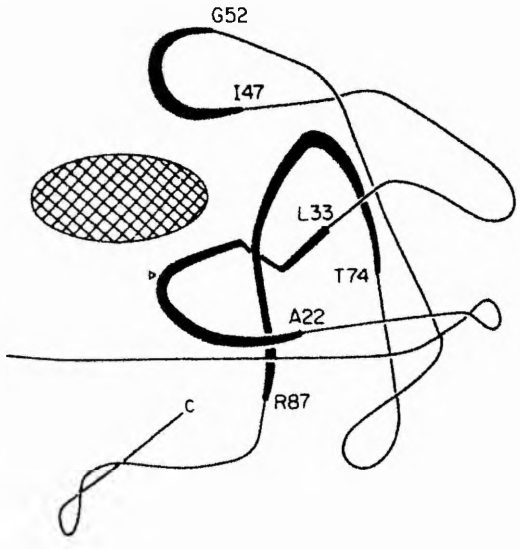


Fig.17. Structure of the RSV and HIV-1 proteases.
The above are α-carbon structures of the RSV and HIV-1 proteases were taken from Miller *et al.* (1989a) and Navia *et al.* (1989) respectively. The side views show the active site clefts and flaps. The N- and C- termini of the proteases are marked and one of the subunits is shown in bold in each case.

Fig. 18 Mutagenesis of the HIV-1 protease.

(b)



Complete mutagenesis of the HIV-1 protease was carried out by Loeb *et al.* (1989) and the diagrams in (a) and (b) were both taken from their paper. Site directed mutagenesis was used to change each residue of the enzyme to at least one other amino acid. The activity profile in (a) shows the effects of non-conservative amino acid substitutions at each position. The mutationally sensitive domains have been superimposed on the crystal structure of one subunit of the HIV-1 protease in (b). The diagram has been rotated 90° in an anti-clockwise direction when compared to the structure in fig.17 above.

The ultimate aim of the structural studies of viral proteases is to understand their mechanisms so that specific inhibitors can be designed and with HIV-1 and -2 such compounds have already been prepared. Two main approaches to developing inhibitors have been taken, with the conventional way being the use of peptide substrate analogues. This has resulted in the production of a number of successful inhibitors including compounds that inhibit the HIV proteases in the nanomolar range (Roberts *et al.*,1990). These compounds do not inhibit cellular aspartic proteases and so provide a good starting point for the production of an anti-viral agent. The second approach taken to find an inhibitor of the HIV proteases makes use of computer modelling and X-ray crystallography data bases to search for compounds that will fit into the substrate binding pocket of the enzymes. DesJarlais *et al.* (1990) used this method to predict that the antipsychotic agent, haloperidol, would bind to the HIV-1 protease and found that it was indeed an effective inhibitor. Haloperidol was found to inhibit in the micromolar range, concentrations which unfortunately would be too toxic for human use, but nevertheless, demonstrate the potential of this approach for searching for viral inhibitors.

Much interest surrounds the control of the protease activity in HIV and other retroviruses. The observations that chemically synthesised HIV protease is able to cleave peptide substrates and that fusion proteins purified from *E.Coli* undergo autocatalytic processing suggests that no other elements are required for activity (Schneider and Kent,1988; Boutelje *et al.*,1990). In *vivo*, however, it is imperative that cleavage of the *gag* and *gag-pol* polyproteins does not take place until after the virus particles are assembled. Premature processing would prevent the assembly of the virus particles as has been shown by the preparation of recombinant HIV-1 coding for a protease that is more active than the wild type enzyme (Krausslich,1991). So what controls the activity of the HIV proteases in the cell? Perhaps a natural inhibitor of the enzymes is present in the cytoplasm or alternatively the dimerisation could hold the key to the activation step. It has been proposed that in the cell the *gag-pol* polyproteins are too distal for dimerisation to take place, but that fusion to the plasma membrane through the N-terminal, myristylated MA protein serves to concentrate the polyproteins. This significantly increases the occurence of dimerisation. Furthermore, if

dimerisation takes place before budding of the viral particles from the plasma membrane then the protease would autocatalytically cleave at its N-terminus releasing itself and the other *pol* proteins into the cytoplasm (Navia *et al.*,1989). As a result the activity would be diluted in the cell. In contrast if dimerisation and autocatalysis take place after formation of the virus particles then the enzyme will be confined within the virion and a rapid cascade of maturation cleavages can ensue. In any case the role of dimerisation in the activation of retroviral proteases is necessary, if not sufficient, and so provides another target for a specific antiviral agent. The potential for such an approach has been explored by Babe *et al.* (1991) who prepared heterodimers of the HIV-1 and -2 proteases. Although the 2 enzymes are highly related both structurally and functionally they differ in 55% of their residues, a factor that is reflected in their surprisingly diverse isoelectric points; 9.9 for HIV-1 and 5.1 for HIV-2 (Gustchina and Weber,1991). The mixing of purified HIV-1 and -2 was found to cause severe inhibition of HIV protease activity against synthetic peptide substrates. Heterodimer formation was confirmed by the appearance of a species with intermediate isoelectric point. The inhibition results from the fact that around 50% of the amino acids from the HIV proteases reside at the dimer interface and some of these differ between types -1 and -2 (Wlodawer *et al.*,1989). As a result the inter-subunit contacts are not formed correctly and enzyme activity is lost. Babe *et al.* (1991) suggest further that heterodimers can be used to identify the critical contact points between monomers. The design of specific compounds that bind to these could lead to inhibition of dimer formation and thus to the production of an effective antiviral agent.

In addition to examining the processing of the viral polyproteins, some effort has also been devoted to searching for cellular substrates of the HIV proteases. Tomasselli *et al.* (1991) have shown that calcium free calmodulin is a substrate for both the HIV-1 and -2 proteases, but that the processing sites are different in the 2 enzymes. Shoeman *et al.* (1991) have concentrated on the intermediate filaments and shown that vimentin, α-actinin, α-spectrin, tropomysin, and actin are all cleaved at specific points by the HIV-1 protease. In a separate report fibronectin is described as a non-viral substrate for the HIV-1 protease (Oswald and von der Helm,1991). It could be that these *in vitro* observations are artifacts and not relevant to the situation *in vivo*. The possibility cannot be ruled out, however, that some of these proteins are

cleaved by the HIV proteases *in vivo*. Indeed the cleavage of specific cellular proteins could be an integral part of the virus lytic cycle and it has been proposed that cleavages of specific cytoskeletal proteins might be necessary for the release of virions from the cell (Shoeman *et al.*, 1991). Alternatively the cleavage of host proteins may simply be toxic to the cell and this point emphasises the need for strict control of viral protease activity within the cell.

2.6 DNA VIRUSES

Unlike the RNA viruses discussed above, DNA viruses produce monocistronic RNA and do not synthesise polyproteins, so they only require proteases for the processing or activation of precursor proteins to mature functional units. Proteolytic processing has been shown to play an essential role in the life cycles of a number of DNA viruses including the adeno-, vaccinia, African swine fever and herpes simplex viruses. The adenovirus protease is by far the best characterised of this group; but even it falls a long way behind its picorna- and retroviral counterparts. The adenoviruses have already been described in section 1, with particular reference being made to its protease in section 1.8. The Ad2 protease will also be discussed in more detail in the introductions to chapters 1-3, so for the moment attention will be turned to the other DNA viruses.

Vaccinia Virus.

Vaccinia virus is a member of the Poxviridae and is a large, enveloped virus containing a linear double strand DNA genome, comprising 192 000 base pairs and with the potential to code for at least 200 proteins (Sarov and Jolik,1972; Goebel *et al.*,1990).

Proteolytic processing of 3 of the vaccinia virus structural proteins, termed P4a, P4b and P25K is essential for the production of mature infectious virions. Limited proteolysis gives rise to the mature core proteins which are referred to as 4a, 4b and 25K respectively.

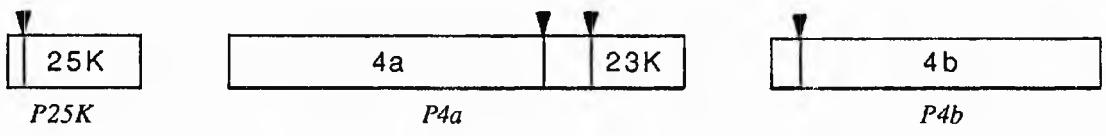


Fig. 19 Proteolysis of Vaccinia Virus Proteins
Schematic diagrams of the 3 Vaccinia virus proteins known to be proteolytically processed. Arrows indicate the sites of cleavage by an as yet unidentified protease.

Pulse labelling experiments using vaccinia infected cells have revealed that there is a

considerable lag period between the synthesis of the core protein precursors and their processing. VanSlyke *et al.*(1991a) suggest that this could be because the precursor proteins must assemble in immature virus particles before the cleavages can take place. Alternatively, they propose that post translational modifications of either the substrates, or the protease, may be required late in infection before processing can proceed.

Recent work identifying the cleavage sites in P4a, P4b and P25K will be discussed in chapter 1 (VanSlyke *et al.*,1991a,b). As yet, it is not known whether the protease responsible for the maturation cleavages in vaccinia is of viral or cellular origin.

African Swine Fever Virus.

African swine fever virus (ASF), like vaccinia, is a large enveloped double stranded DNA virus with a genome of 170 000 base pairs that is responsible for a highly contagious disease in domestic pigs. Three of the ASF structural proteins; p150, p37 and p34 are derived from larger precursor proteins termed pp220, pp60 and pp39. The protease responsible for these processing events has not been identified, but the cleavage sites have been established and will be discussed in chapter 1 (Lopez-Otin *et al.*,1989).

In the future it will be interesting to learn if any other, less abundant vaccinia and ASF viral proteins are also proteolytically processed and to establish the role of the proteolysis in their lytic cycles.

Herpes Simplex Virus.

Herpes simplex virus 1 (HSV-1) is a member of the double strand DNA Herpesviridae family and is an icosahedral enveloped virus (Booy *et al.*,1991). A 45-kDa HSV-1 protein, referred to as ICP35, has been shown to be proteolytically processed to give a 39-kDa protein via several intermediates (Braun *et al.*,1984). ICP35 is present in large amounts in intermediate capsids, but only trace amounts of the protein are present in mature virions. This has given rise to the suggestion that it plays an important role in virion assembly and DNA encapsidation. In a recent report by Preston *et al.* (1992) a second viral protein, termed UL26, has been shown to be essential for the proteolysis of ICP35. A temperature sensitive mutant, ts1201, defective in UL26, was isolated and shown to be unable to process ICP35 or to assemble mature virions at the non-permissive temperatures. The suggestion, therefore, is that

the 70-kDa UL26 protein is a protease, although the possibility that the UL26 plays an indirect role in the processing of ICP35 has not been ruled out. In any case, the ts1201 mutant could prove invaluable in future studies aimed at further characterising the role of proteolysis in HSV-1 infections.

The above discussion serves to emphasise the essential role that proteases play in the infectious cycles of many viruses. Further consideration will be given to the substrate specificities and mechanisms of the virus encoded enzymes in chapters 1 and 2 respectively. The focal point of this thesis, however, is the adenovirus protease and its further characterisation should prove useful from several viewpoints. It is hoped that it will enable us to learn more about the role that the protease and its substrates play in the complicated life cycle of the adenovirus described in section 1. It is also envisaged that the work might be extended to provide a model for the study of proteolysis in other viruses and in particular in the DNA viruses (section 2.6). In the longer term, the information obtained could provide the basis for future studies aimed at designing and synthesising specific inhibitors as anti-viral agents.

In an attempt to achieve these aims the substrate specificity, mechanism, expression and control of the adenovirus protease will be investigated in chapters 1-4 respectively. Chapter 4 will also consider the role that the processing of the adenovirus pre-terminal protein plays in the replicative cycle of the virus.

CHAPTER 1

Substrate Specificity of the Ad2 Protease

INTRODUCTION

The aim of this chapter is to define the precise requirements for the substrate specificity of the adenovirus protease. Based on observations by Tremblay *et al.* (1983) it has become generally accepted that the enzyme cleaves at glycine, alanine containing bonds (for reviews see: Krausslich and Wimmer,1988; Kay and Dunn,1990). Such a broad specificity, however, is not in keeping with the enzyme's selective nature and the recent identification of a glycine-methionine cleavage site in the 11-kDa protein (Weber and Anderson,1988) suggests that a rigorous examination of the Ad2 protease specificity is required. A definition of the enzyme's substrate specificity would not only enable the prediction of cleavage sites in all the viral substrates, but would also provide a basis for the rational design of specific inhibitors as anti-viral agents. Before examining the specificities of other virus encoded proteases a description of the established and predicted processing sites in the Ad2 proteins will be given.

SUBSTRATES OF THE AD2 PROTEASE

The Ad2 protease has previously been shown to cleave 6 viral proteins, namely, the pTP, IIIa, pVI, pVII, pVIII and an 11-kDa protein (Anderson *et al.*,1973; Boudin *et al.*,1980; Tremblay *et al.*,1983; Weber and Anderson,1988). The cleavage sites in pVI and pVII were identified by N-terminal sequencing of the mature proteins, VI and VII (Akusjarvi and Persson,1981; Sung *et al.*,1983). The amino acids in the vicinity of the cleavage sites in pVI and pVII are shown in **fig.1.1** along with the Berger and Schechter (1970) notation for referring to the surrounding residues. It was the similarity between these 2 sites that led Tremblay *et al.* (1983) to propose that the protease has a specificity for G-A bonds. Sequence analysis of the other Ad2 protease substrates gave rise to the proposal that A-G bonds might also be substrates of the enzyme (Tremblay *et al.*,1983; Hannan *et al.*,1983).

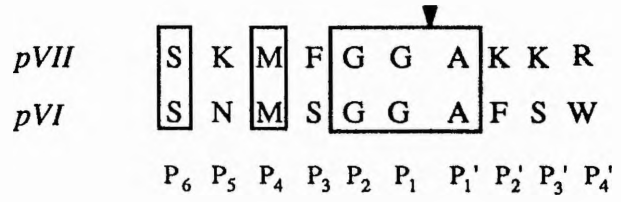


Fig. 1.1 Cleavage Sites in pVI and pVII
Alignment of the established pVI and pVII cleavage sites showing the Berger and Schechter (1970) notation for identifying surrounding residues. The arrow shows the cleaved bond and conserved residues are boxed.

Smart and Stillman (1982), on the other hand, took a broader look at the pVI and pVII sites and proposed that a similar sequence in the pTP (MTGG-V) could be the target for the action of the protease. More recently one of the cleavage sites in the 11-kDa protein, that gives rise to the basic core protein, mu, has been shown to be LTGG-M leading to the suggestion that the protease recognises the sequence GG-X (Weber and Anderson,1988). The possibility that the protease cleaves other viral or cellular proteins cannot be ruled out and indeed, there have been suggestions that the L4 100-kDa protein and the 23-kDa protease itself, are processed (Morin and Boulanger,1986; Chatterjee and Flint,1987). **Table 1.1** lists all the proposed and established cleavage sites of the Ad2 protease.

Protein	Putative Cleavage Site	Status
pVI	MSGG-AFSW	Confirmed
pVII	MFGG-AKKR	Confirmed
pVIII	TNSG-AQLA	
	LAAG-AAQD	
	QIGG-AGRS	
pTP	MTGG-VFQL	
IIIa	YTGG-FEVP	
11-kDa	LTGG-MRRA	Confirmed
23-kDa	VNTA-GRET	

Table 1.1 Putative Cleavage sites in Adenovirus Proteins
This table lists sequences in Ad2 proteins which have been proposed as cleavage sites for the virus encoded protease. Confirmed indicates that the sites have been established by sequencing studies. For references see text.

Consider next, the substrate specificities of other viral proteases.

SUBSTRATE SPECIFICITIES OF VIRAL PROTEASES

As discussed in section 2 of the main introduction, virus encoded proteases of restricted specificity play essential roles in the infectious cycles of many viruses. Much time has been devoted to investigating the substrate specificities of these enzymes, with the design of substrate analogues as inhibitors being the ultimate goal. In the first instance the cleavage sites in the natural substrates must be identified by N-terminal sequencing of the mature proteins. The precise requirements for specificity can then be established by varying the amino acids at each of the positions around the scissile bond and measuring the effects of the changes on the efficiency of processing. This can be achieved either by mutagenesis of the sites in the natural substrates or through the use of a range of synthetic peptides. Information is available concerning the bonds cleaved by the picorna-,como-, poty-, toga-, flavi- and retroviral proteases. Cleavage sites in proteins from the DNA viruses vaccinia and African swine fever have also been identified by N-terminal sequencing. Selected aspects of the substrate specificities of the proteases from these viruses will be considered in turn.

Picornaviruses

The picornaviruses code for 2 proteases, termed 2A and 3C, which are required for the separation of their polyproteins into functional units (see p35). The 2A and 3C cleavage sites in the poliovirus polyprotein are shown in **fig.1.2.**

Although the poliovirus polyprotein contains eight Y-G bonds, only two of these are cleaved by its 2A protease. This indicates that elements of the enzyme's specificity must reside outside the P_1 and P_1' positions. Comparisons of the VP1-2A cleavage sites from different strains of the polio-, rhino- and coxsackie viruses reveal that in each case the P_2 position is occupied by a threonine and that one of the hydrophobic amino acids methionine, leucine, isoleucine or valine is present at P_4. These observations led to the suggestion that the P_2 and P_4 residues provide additional determinants of specificity (Lee and Wimmer,1988). Mutagenesis at the P_2 position, however, has shown that its identity is not important for *cis* cleavages at the VP1-2A site in the poliovirus polyprotein (Hellen *et al.*,1990). Skern *et al.* (1991) investigated the specificity of the rhinovirus 2A protease by mutating the VP1-2A site

and studying the effects of the mutations on *cis* cleavages. Their studies indicate that processing is not affected by substitutions at P_1, but that all mutations altering the glycine at P_1' abolish the cleavage at the VP1-2A site. Changes of the amino acids at the P_2'-P_9' positions were also found to reduce the efficiency of the processing at the VP1-2A site. These residues, however, are equivalent to amino acids 2-9 of the 2A protease and so may be structurally or functionally important to the enzyme. Finally, there is a suggestion that the 2A proteases have different specificities for *cis* and *trans* cleavages, with the requirements for the *trans* cleavages being more stringent (Alvey *et al.*,1991). The elucidation of the specificities of the 2A proteases in *trans*, would be of particular interest in view of their role in the cleavage of p220 during the shut-off of host cell protein synthesis.

VP0-VP3	T L P R L Q ▼ G L P V M	**3CD**
VP3-VP1	Q K A L A Q ▼ G L G Q M	
2A-2B	E E A M E Q ▼ G I T N Y	
2B-2C	P Y V I K Q ▼ G D S W L	
2C-3A	M E A L F Q ▼ G P L Q Y	**3C or 3CD**
3A-3B	L F A G H Q ▼ G A Y T G	
3B-3C	R T A K V Q ▼ G P G F D	
3C-3D	Y F T Q S Q ▼ G E I Q W	
VP1-2A	K D L T T Y ▼ G F G H Q	**2A**
3C'-3CD'	K L L D T Y ▼ G I N L P	

Fig. 1.2 Cleavage sites in the poliovirus polyprotein.
The sites in the poliovirus type 1 (Mahoney) polyprotein cleaved by the 3CD, 3C and 2A proteases are listed above. The location of each site in the polyprotein is given in italics on the left (see fig.11a, p37) and the protease responsible for each cleavage is shown on the right in bold. The arrow indicates the cleaved bond in each case. The figure is adapted from Hellen *et al.*(1989).

In contrast to the 2A proteases, the substrate specificities of the polio- and rhinovirus 3C proteases have been extensively studied using both synthetic peptide substrates and site directed mutagenesis of the natural cleavage sites in the viral polyproteins. The 8 sites cleaved by the poliovirus 3C protease are listed in **fig.1.2** and it can be seen that in each case a Q-G bond is cleaved. Indeed, the poliovirus 3C protease has stringent requirements at P_1 and P_1' with the only permissible substitutions being those of serine and alanine at P_1' (Kean *et*

al.,1990). In addition to the P_1 and P_1' positions, the amino acid at P_4 has been shown to have a strong influence on the efficiency of cleavage, with alanine and valine being the preferred residues at P_4 for the polio- and rhinovirus 3C proteases respectively (Blair and Semler,1991; Cordingley *et al.*,1990). The 3CD site in the poliovirus polyprotein, which is processed slowly *in vivo* and *in vitro*, contains a threonine at P_4. Pallai *et al.* (1989) demonstrated that a single substitution of the threonine with an alanine at P_4 in the 3CD site increased its processing by 2 orders of magnitude, further demonstrating the importance of the P_4 position in defining the substrate specificity of the 3C protease. It has also been shown that the bonds cleaved most efficiently by the polio- and rhinovirus 3C proteases have a proline residue in the P_2' position (Pallai *et al.*,1989; Cordingley *et al.*,1989). This has led to speculation that the 3C proteases recognises Q-G bonds at β-turns and to the proposal that the preferred substrate specificity of the poliovirus 3C protease is AXXQ-GP. The strong influence of the neighbouring residues on the rates of cleavage of the different Q-G bonds in the poliovirus polyprotein is thought to be important in the regulation of gene expression during the infectious cycle (Hellen *et al.*,1989).

Positive Strand RNA Plant Viruses

The genome organisations and proteolytic processing strategies of the como-, poty- and nepoviruses are very similar to those of the picornaviruses (see p39). It is perhaps not surprising, then, that the substrate specificities of the cowpea mosaic virus (CPMV), 24-kDa protease and the potyvirus NIa proteases closely resemble those of the picornavirus 3C proteases.

The CPMV 24-kDa protease preferentially cleaves at Q-G bonds, but serine, alanine and methionine are also acceptable at P_1', albeit with a processing efficiency that is markedly reduced (Vos *et al.*,1988). Indeed the cleavage of the Q-M bond in the M-RNA polyprotein does not take place unless a 32-kDa virus protein is also present. The 24-kDa protease also shows a preference for alanine at P_4 and alanine or proline residues are found at P_2 in all of the enzymes natural cleavage sites (Wellink and van Kammen,1988).

The cleavage sites of the NIa and HC-Pro proteases from 4 different potyviruses are shown in **fig.1.3**. The NIa proteases have a 3C-like specificity recognising Q-G, Q-S or Q-A bonds (Riechmann *et al.*,1992). There is a strong preference for a hydrophobic residue at the P_4 position, with valine being present at P_4 in all the sites in the tobacco vein mottling virus (TVMV) polyprotein. The P_6 position is occupied by a glutamate residues in all the NIa cleavage sites in the tobacco etch virus (TEV) polyprotein.

	TEV	**TVMV**	**PPV**	**PVY**
HC-Pro-50K	YNVG-GM	YKVG-GL	YLVG-GL	YRVG-GV
50K-71K	EIIYTQ-S	NNVRFQ-S	QAVQHQ-S	YEVRHQ-S
71K-6K	ETIYLQ-S	EAVRFQ-S	ECVHHQ-T	QFVHHQ-A
6K-NIa	EPVYFQ-G	EQVKFQ-G	EEVVHQ-G	ETVSHQ-G
NIa-58K	ELVYSQ-G	DLVRTG-G	EFVYTQ-S	DVVVEQ-A
58K-30K	ENLYFQ-S	ETVRFQ-S	NVVVHQ-A	YEVHHQ-A

Fig.1.3 The HC-Pro and NIa cleavage sites from four potyviruses.
The amino acids surrounding the HC-Pro and NIa cleavage sites from tobacco etch virus (TEV), tobacco vein mottling virus (TVMV), plum pox virus (PPV) and potato virus Y (PVY) are listed. The positions of the cleavages are on the left in italics and refer to the nomenclature of the TEV polyprotein (see fig.12). The HC-Pro and 50K proteins are separated by HC-Pro, whilst all the other cleavages are catalysed by the NIa protease. In each case the position of the cleavage site is indicated by a dash (-). The figure has been adapted from Reichmann*et al.*(1992).

The consensus sequences for the TVMV and TEV NIa substrates have been described as XXVR(K)Q-G(S) and EXXYXQ-G(S) respectively (Dougherty and Carrington,1988). Interestingly, it has been noted that the TVMV and TEV NIa proteases are unable to recognise each other's cleavage sites. Parks and Dougherty (1991) took advantage of this observation and constructed hybrids of the TEV/TVMV NIa proteases in an attempt to identify the substrate binding regions in the enzymes. Their work suggests that the amino acids required for substrate binding are contained within the C-terminal 150 amino acids of NIa.

No investigations have been carried out to identify the substrate specificities of the potyvirus HC-Pro proteases, which carry out single processing events at their own C-termini. It is noted, however, that the HC-Pro cleavage sites identified in 5 different potyviruses are all of the form VG-GL(M,V) and are surrounded by additional hydrophobic amino acids **(fig.1.3)**.

Togaviruses and Flaviviruses

Alphaviruses are members of the Togaviridae family and code for an nsP2 protease,

which is required for the separation of their non-structural proteins (see p41). The 3 nsP2 cleavage sites have been identified in 6 different alphavirus polyproteins and are listed in **fig.1.4** (Strauss and Strauss,1990). The proteases appear to show a strong preference for a glycine residue at the P_2 position and amino acids with small side chains are present in the P_1 and P_1' positions, except at the nsP3-nsP4 cleavage site, where the P_1' residue is a tyrosine in all cases. Interestingly, it has been noted that the nsP2-nsP3 bond is cleaved first, followed by the nsP1-nsP2 one and that the nsP3-nsP4 link is processed very slowly.

	SIN	MID	SF	RR	ONN	VEE
nsP1-nsP2	DIGA-A	RAGA-G	HAGA-G	RAGA-G	RAGA-G	EAGA-G
nsP2-NsP3	GVGA-A	TAGC-A	TAGC-A	TAGC-A	RAGC-A	EAGC-A
NsP3-NsP4	GVGG-Y	RAGA-Y	RAGA-Y	RAGA-Y	RAGG-Y	DAGA-Y

Fig. 1.4 The nsP2 cleavage sites from six alphavirus polyproteins
Sequences around the cleavage sites in the non-structural proteins from Sindbis virus (SIN), Middelburg virus (MID), Semiliki Forest virus (SF), Ross River virus (RR), O'Nyong-nyong virus (ONN) and Venezuelan equine encephalitis virus (VEE). The locations of the cleavage sites are given in italics on the left (see fig.13) and the scissile bond is marked by a dash in each case. This figure is adapted from Strauss and Strauss (1990).

Thus the presence of the bulky tyrosine residue at P_1' in the latter could serve to regulate the complete processing of nsP3 and nsP4 (de Groot *et al.*,1990). Mutagenesis studies have shown that a glycine at P_2 is a key element in defining the nsP2 substrate specificity. Changes of glycine to valine at P_2 were found to eliminate processing, whilst glycine to alanine substitutions reduced it significantly (Shirako and Strauss,1990).

The capsid proteins of the alphaviruses have also been shown to possess protease activity. The sindbis virus 26-kDa capsid protein cleaves at the carboxyl side of the tryptophan residue at its own C-terminus, in keeping with its classification as a chymotrypsin-like serine protease (Choi *et al.*,1991).

The flavivirus NS3 proteases are required for the processing of the virus non-structural proteins (see p43). The 4 cleavage sites have been identified in several different flaviviruses including the Yellow fever and Dengue viruses. In every case the P_1 and P_2 positions are occupied by one of the basic amino acids, arginine or lysine. It has also been noted that amino acids with short side chains are frequently present at the other positions in the vicinity of the

NS3 cleavage sites (Strauss *et al.*,1987; Preugschat *et al.*, 1991; Chambers *et al.*,1991).

The substrate requirements of the pestivirus proteases, typified by the BVDV p20 and p80 proteins remain to be elucidated (Wiskerchen and Collet,1991).

Retroviruses.

The cleavage sites of the HIV-1 and RSV proteases in their respective *gag-pol* polyproteins are listed in **fig.1.5**. A striking feature of the sites is the variety of amino acids present at the various positions, including P_1 and P_1'.

	RSV Sites											HIV-1 Sites										
1	G	T	S	C	Y	-	H	C	G	T	A	**1**	V	S	Q	N	Y	-	P	I	V	Q	N
2	P	P	Y	V	G	-	S	G	L	Y	P	**2**	K	A	R	V	L	-	A	E	A	M	S
3	P	V	V	A	M	-	P	V	V	I	K	**3**	T	A	T	I	M	-	M	Q	R	G	N
4	I	A	A	A	M	-	S	S	A	I	Q	**4**	R	P	G	N	F	-	L	Q	S	R	P
5	Q	P	L	I	M	-	A	V	V	N	R	**5**	V	S	F	N	F	-	P	Q	I	T	L
6	P	P	A	V	S	-	L	A	M	T	M	**6**	C	T	L	N	F	-	P	I	S	P	I
7	R	A	T	V	L	-	T	V	A	L	H	**7**	G	A	Q	T	F	-	Y	V	N	L	R
8	T	F	Q	A	Y	-	P	L	R	E	A	**8**	I	R	K	I	L	-	F	L	D	G	I
9	S	P	L	F	A	-	G	I	S	D	W												

Fig. 1.5a Retroviral Protease Cleavage Sites
The sequences of the cleavage sites in the Rous sarcoma virus (RSV) and human immunodeficiency virus type 1 (HIV-1) gag-pol polyproteins. The numbers in bold to the left of each sequence refer to the positions of the sites in the respective polyproteins with site 1 being nearest to the N-terminus of the polyprotein in each case (see fig.15 and 16 for the organisation of the RSV and HIV polyproteins). The scissile bonds are indicated by a dash and the figure is adapted from Hellen*et al.*(1989).

The F-P bond occurs at a number of junctions in retroviral polyproteins, but other residues acceptable in the P_1 and P_1' positions include L-F, M-S, M-M, L-A and F-L. Although the principle features of substrate recognition are similar for the RSV, HIV-1 and HIV-2 proteases, there are subtle differences in the requirements at each of the positions (Tomasselli *et al.*,1990; Tozser *et al.*,1991). This means that they are not able to hydrolyse the same sets of substrates.

Most investigations have concentrated on the specificity of the HIV-1 protease and so it will be described here. Darke *et al.* (1988) showed that the HIV-1 protease was able to recognise and cleave the 7 amino acid peptide, SQNY-PIV, corresponding to the *gag* MA-CA site (see **fig.16**, p48). This observation revealed that the primary determinants of specificity for the HIV-1 protease must reside within the P_4 to P_3' positions. It was noted, however, that the K_m value for the cleavage was high, at 9mM and so the possibility exists that there are secondary binding sites, perhaps in the flaps of the retroviral protease (see **fig.17**, p49).

Alternatively, the high K_m value could reflect the inability of the peptide substrate to adopt the optimal conformation for a significant proportion of time (Kay and Dunn,1990).

The importance of each of the P_4-P_3' positions in defining the substrate specificity of the HIV-1 protease has been examined extensively using peptide substrates. Studies by Konvalinka *et al.* (1990) revealed that hydrophobic or aromatic residues are preferred at P_1 and the branched amino acids isoleucine and leucine at P_2. They found, however, that there are few restrictions at the P_3 position and suggest that P_3 can be occupied by any amino acid except proline. Margolin *et al.* (1990) examined the P_2' position and showed that peptides containing isoleucine, leucine or alanine at P_2' all bound to the HIV-1 protease with equal efficiencies. The hydrolysis of the peptide containing isoleucine at P_2', however, was 20 fold greater than for the alanine and leucine containing peptides. The processing of peptides containing glycine, phenylalanine and tryptophan at P_2' was negligible and these results led Margolin *et al.* (1990) to conclude that small hydrophobic amino acids are preferred at P_2'.

Other investigators have considered the effects of the mutagenesis of cleavage sites in the HIV-1 *gag-pol* polyprotein and have shown that hydrophobic residues are preferred at P_1 and P_1', that neither proline nor the branched chain amino acids, valine and isoleucine, are acceptable at P_1 and that small apolar amino acids are preferred at P_2 (Billich and Winkler,1991). Finally, Poorman *et al.* (1991) have examined all the known cleavage sites of the HIV-1 and HIV-2 proteases in viral and cellular proteins. The results of their studies are shown in **table 1.2** and in accord with the peptide and mutagenesis studies reveal that the HIV-1 protease shows the highest stringency for particular amino acids at P_2, P_1 and P_2', whilst the HIV-2 protease also has high stringency at P_1'. The lack of a primary consensus sequence to describe the substrate specificities of the HIV proteases serves to emphasise the importance of secondary structure in substrate recognition by viral proteases.

	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	P_4'
HIV-1	*0.8* P (2.8) A (2.4) S (2.2)	*1.2* Q (6.5) E (4.7)	*1.8* V (6.2) N (5.3) I (5.33)	*3.3* F (13.7) L (6.8) M (4.2) N (2.8) Y (2.2)	*1.2* F (4.5) P (4.0) Y (3.5) M (2.2)	*4.1* E (15.0) Q (6.5)	*1.1* F (3.2) T (2.9) E (2.1) R (2.1)	*0.8* F (3.2) P (2.8) M (2.2) S (2.2)
HIV-2	*1.8* P (4.6) G (3.5) R (3.5) Y (3.0) S(2.6)	*1.5* Q(11.1) F (2.7) E (2.5)	*2.4* N (11.4) I (3.7) V (2.7) A (2.6) E (2.5)	*2.3* M (8.9) L (7.4) Y (3.0) F (2.7)	*3.0* P (9.1) M (5.4) F (5.1) A (5.0) T (3.4)	*3.5* E (13.5) Q (8.4) V (2.7)	*0.5* Q (3.1) E (2.5)	*1.4* F (7.5) Q (5.7) P (2.3)

Table 1.2 Specificity of the HIV-1 and HIV-2 proteases
This table has been adapted from Poorman *et al.*(1991) who carried out statistical analysis of the substrate preferences of the HIV-1 and 2 proteases using all the known cellular and viral protein substrates. The positions of the amino acids around the scissile bonds are labelled in bold according to Berger and Schecter (1970) and the selectivity parameter for each position is given in italics. The parameter was calculated by Poorman *et al*(1991) and a high value indicates a specific requirement by the protease at that position. The amino acids occuring most frequently at each position are listed along with their relative abundances.

DNA Virus Proteases

Apart from the adenovirus, cleavage sites have been identified in proteins from 2 other DNA viruses; African swine fever (ASF) and vaccinia virus (VV).

Lopez-Otin *et al.* (1989) reported that 3 of the ASF structural proteins, p150, p37 and p34 are derived from the precursors pp220, pp60 and pp39 respectively (**fig.1.6a**). They showed further that the proteolytic cleavage, in each case, was at a GG-X sequence and proposed that GG-X is an important recognition sequence for the processing of a variety of viral and cellular proteins including those of the adenovirus.

Proteolytic maturation of at least 3 major structural proteins, p4a, p4b and p25K, is essential for the formation of infectious vaccinia progeny. In each of the substrates, the P_1 and P_2 positions are occupied by glycine and alanine residues, respectively (VanSlyke,1991a and b) and the cleavage sites are listed in **fig.1.6b**.

Summary

In conclusion the substrate specificities of viral proteases are complex and varied with secondary structure apparently playing an important part. This is particularly obvious for the

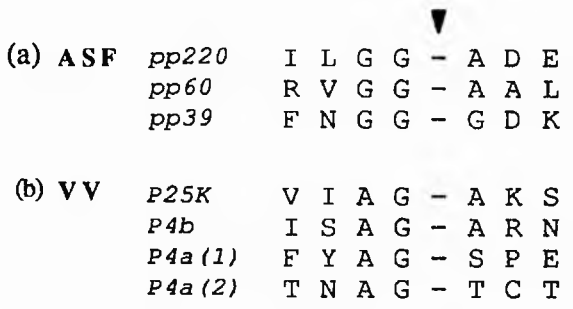


Fig. 1.6 Cleavage sites in DNA virus proteins
The amino acids surrounding the established cleavage sites in African swine fever virus (ASF) and vaccinia virus (VV) proteins. The data is from Lopez-Otin *et al.*(1989) and Vanslyke *et al*(1991b) and the precursor proteins are included in italics. There are two cleavage sites in P4a with site 1 being nearest to the N-terminus (see fig.19). The position of the cleavage sites is marked by the filled arrow.

picornavirus 3C and retrovirus proteases where it has been suggested that the scissile bonds form part of a β-turn (Kay and Dunn,1990).

The number of residues permitted in the P_1 and P_1' positions is considerably greater than for most cellular proteases; but there tends to be more stringency at other positions further away from the scissile bond and in particular at P_2 and P_4. Several of the proteases including those from alphaviruses, flaviviruses, the vaccinia and African swine fever viruses have been shown to have specific requirements at both P_1 and P_2.

Finally, Wellink and van Kammen (1988) made the observation that the N-termini created by viral proteases usually consist of glycine, serine, alanine, threonine, methionine or valine, giving them a longer half life in the cytoplasm than proteins with one of the other amino acids at their N-terminus. They go further to suggest that viruses employ proteases with a specificity towards these amino acids at the P_1' position and note that nsP4 of the alphaviruses is an exception to this, having a tyrosine at its N-terminus and also, perhaps significantly, a short half life.

The proposal in this chapter is to synthesise a series of peptides based on the known cleavage sites of pVI and pVII, to test them as substrates for the adenovirus protease and in so doing to define its substrate specificity. Using this information and that already available on the molecular weights of the precursor and mature proteins it should be possible to predict the cleavage sites in all the known Ad2 substrates.

METHODS

1 PEPTIDE SYNTHESIS

All peptides were synthesised by solid phase fluorenyl methoxycarbonyl (Fmoc) polyamide chemistry using a semi-automated CRB Pepsynthesiser II essentially as described by Atherton *et al.* (1988). The peptides synthesised were: MFGGAKKR, MFGGAKK, FGGAKK, MSGGAFSW, SGGAFSW, VSGGAFSW, YSGGAFSW, QSGGAFSW, ESGGAFSW, LSGGAFSW, TSGGAFSW, ISGGAFSW, MSAGAFSW, MSGSAFSW, MSGGGFSW, MSGAGFSW, MSGSGFSW, LAGGFRHR, MRGFGVTR, MGGRGRHL, SLGGGVPW and MTGGVFQ.

1.1 Reagent Grade and Purification.

The success of this technique is absolutely dependent on high grade solvents free from aqueous or amine contaminants. The Pepsyn KA resin (0.01 mequivalents/g), the Fmoc pentafluorophenyl ester derivatives of amino acids, o-rings and filters required for the synthesis were all obtained from CRB. Dimethylformamide (DMF), ethanol, methanol, acetonitrile and trifluoroacetic acid (TFA) were all HPLC grade from Rathburn. Piperidine was purchased from Applied Biosystems and ethanedithiol, anisole, tertiary amyl alcohol, 4-dimethylaminopyridine (DMAP) and 1-Hydroxybenzotriazole (HOBt) from Aldrich. Diethyl ether and acetic acid were obtained from May and Baker and the molecular sieves, difluoronitrobenzene (FDNB) and other chemicals from Sigma. All solid reagents were stored in the presence of silica gel as recommended by the manufacturers.

The primary solvent for peptide synthesis, DMF, was routinely tested for dimethylamine using FDNB. This is because dimethylamine can react with the Fmoc protecting groups and so when present, even at low concentrations, it can significantly reduce the final yield of peptide.

FDNB test for DMF

The DMF was tested by adding an equal volume of the solvent to 5mM FDNB in ethanol and leaving it at room temperature for 20 minutes. An absorbance of < 0.15 at 391nm, when read against a blank of 2.5mM FDNB in ethanol, indicated that the solvent was of a high

enough grade to use. Tests were carried out in duplicate.

Where the absorbance reading for the FDNB test was greater than 0.15, 50g of heat activated molecular sieves (2 hours at 200°C) were added to 2.5 litres of DMF and left at room temperature for several weeks, after which the DMF was retested and if satisfactory was ready for use. After storage on molecular sieves it is essential to filter the solvent through a 0.2 μ nylon filter (Anachem) to prevent blockage of the lines and valves in the Pepsynthesiser machine.

1.2 Synthesis of Peptides

0.5 - 1g of Pepsyn KA resin was washed with 3x 5ml of DMF to remove fines. The resin was packed into a 5 x 1cm column and the plunger adjusted to 2mm above the resin to remove air bubbles. All the lines in the Pepsynthesiser were flushed out with fresh DMF and the flow through was checked using the FDNB test. Bottles A and B were filled with DMF and 20% v/v piperidine in DMF, respectively, and the lines primed. The main lines were given a final 10 minute wash with DMF to remove all traces of piperidine and the flow rate was adjusted to 3 ml/minute before the reaction column was attached.

The synthesis of the peptides was controlled in part by CRB software using an Apple IIe computer. The programme used is shown in **fig.1.7** alongside a schematic diagram to represent the various steps in the cycle. The first step in each cycle was a 15 minute wash with DMF, carried out prior to the addition of derivatised amino acids in 5 fold excess. Each amino acid was recirculated for 1.5 hours through the column to allow complete coupling. This was verified by removing a few grains of resin and carrying out a ninhydrin or isatin test (see 1.3 below). The excess amino acids were then washed from the column and all the lines, prior to the deprotection step. This involved flowing 20% v/v piperidine in DMF (Reagent B) through the column for 5-10 minutes to remove the Fmoc protection group from the last amino acid attached. The lines were then washed for 15 minutes with DMF and the cycle was repeated until the N-terminal amino acid had been coupled and deprotected. The Fmoc amino acids used were all pentafluorophenyl esters except for serine and threonine, in which cases 3,4-dihydro-4-oxobenzotriazin-3-yl (Dhbt) esters were used. On the addition of these amino acids the resin immediately turned bright yellow, but as coupling progressed the colour

disappeared, giving an internal indicator for coupling and so making the ninhydrin test redundant.

Programme		Time (minutes)
1	FLOW A	5
2	FLOW B	10
3	FLOW A	15
4	CALL (load amino acid)	
5	RECIRCULATE	150
6	FLOW S	1
7	FLOW A	5
8	RECIRCULATE	0.5
9	FLOW A	5
10	CALL	
11	REPEAT	

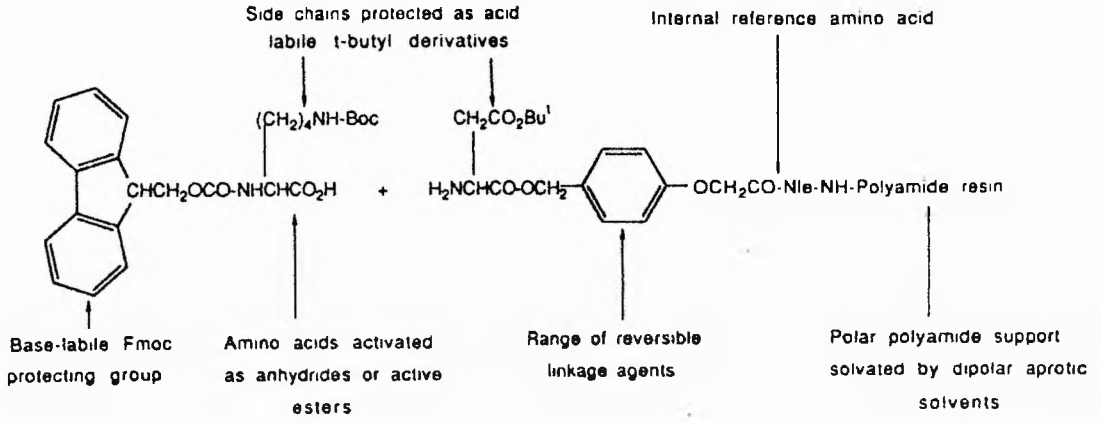


Fig. 1.7 Solid Phase Peptide Synthesis
The CRB programme used to control the synthesis of the peptides on an Apple IIe computer, is shown in (a) above. Amino acids were applied to the column at step 4 and allowed to recirculate round the reaction column for 1.5 hours (step 5). Steps 6-9 are wash steps and step 2 involves application of 20% v/v piperidine in DMF to the column in order to deprotect the N-terminal amino acid. The chemistry of Fmoc polyamide synthesis is shown in (b) and the diagram was taken from Atherton and Sheppard (1989).

All the derivatised amino acids were dissolved in a minimum volume of DMF (1- 3ml) and centrifuged for 3 minutes in a microcentrifuge to remove any insoluble particles that might block the valves in the machine. The coupling of the first amino acid to the resin was always done in the presence of the catalyst DMAP. 6-12mg DMAP were taken up in 200µl of DMF and loaded onto the column through the injection syringe immediately before the amino acid. Another step taken to ensure complete coupling of the first residue was to put it on twice, omitting the deprotection step in the first cycle. DMAP was used for both couplings. Generally speaking the subsequent amino acids were added in the absence of a catalyst unless

the ninhydrin test indicated incomplete coupling (see 1.3 below). In this case the deprotection step was bypassed and the catalyst HOBt (6-12mg taken up in 200µl of DMF) was added to the column immeditely before a second application of the same amino acid.

After the deprotection of the N-terminal amino acid was complete, the resin was removed from the column and placed in a sintered glass funnel. It was washed sequentially with 10ml each of DMF, t-amyl alcohol, glacial acetic acid and t-amyl alcohol followed by 3x 10ml diethyl ether. The resin was placed in a vacuum desiccator for 2-20 hours to dry before being cleaved (see 1.4 below).

1.3 Monitoring of the Synthesis

(a) Spectrophotometric Analysis.

The continuous flow synthesis and the uv-absorbing Fmoc protecting groups enable the acylation and deprotection reactions to be monitored spectrophotometrically. An 8µl flow cell was inserted and aligned in a Cecil spectrophotometer attached immediately downstream from the reaction column. The synthesis was monitored at a wavelength of 320nm and a full scale deflection of 2. The changes in absorbance were recorded on a chart recorder (chart speed 0.1mm/minute). A typical trace for a single coupling/ deprotection cycle of an amino acid is shown in **fig.1.8a.** Occasionally the deprotection was slow, as indicated by a more gradual decline of the trailing edge of the deprotection peak (**fig.1.8b**).

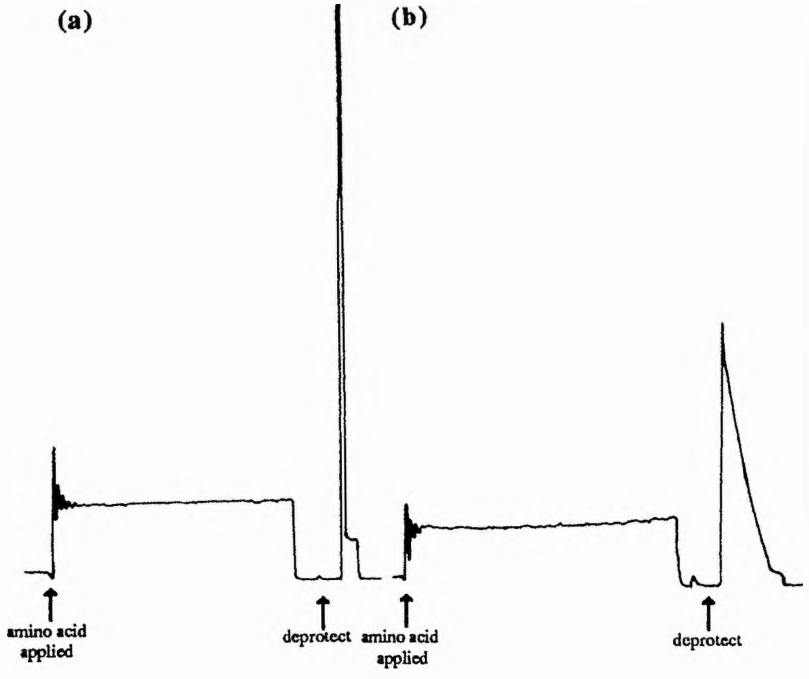


Fig. 1.8 Monitoring of Peptide Synthesis.
Traces (a) and (b) each show the coupling of a single Fmoc amino acid during solid phase peptide synthesis. The synthesis was monitored at 320nm with a full scale deflection of 2 absorbance units and a chart speed of 0.1mm/minute. Trace (a) shows a normal deprotection peak, whilst trace (b) shows a shorter, broader peak. The latter indicates that the N-terminus of the peptide is not readily accessible to the solvent.

This gave an indication that the Fmoc groups were less accessible to solvent than normal and was a warning that the next amino acid might be difficult to couple. In such a case the catalyst, HOBt, was added before the next amino acid. The size of the deprotection peak should be constant throughout each synthesis and a decrease in size is usually indicative of a problem. When a significant decrease in the area of the deprotection peak was observed the synthesis was continued but samples of resin were taken for subsequent hydrolysis and amino acid analysis. Reagents were also rechecked to try and establish the root of the problem.

(b) Ninhydrin (Kaiser) Test.

Small samples of resin (5mg) were removed at the end of the DMF wash following deprotection and 1 hour after the addition of each amino acid. Both samples were washed with 2ml each of DMF, t-amyl alcohol, acetic acid and t-amyl alcohol, followed by 4ml of diethyl ether. The washing was carried out in sintered glass filters under suction. A few beads of resin were transferred to eppendorff tubes and 2 drops each of (a) 10% ninhydrin in ethanol, (b) phenol - 80g in 20ml ethanol and (c) 20μM potassium cyanide in 98% pyridine were added. The tubes were placed in a heating block set at 120°C for 5 minutes. A strong blue colour was indicative that amino groups were present, whilst a brown/yellow colour showed complete coupling. When the N-terminal residue was aspartic acid, asparagine, glutamine or glycine a weaker blue or brown colour was observed (Kaiser *et al.*,1970). In the case of the secondary amino acid proline the isatin colour test gives a stronger reaction with the imino group and so was preferred.

(c) Isatin Test.

Small samples of resin were taken, washed and dried as described for the ninhydrin test and a few beads were placed in an eppendorff tube along with 2 drops of isatin solution (prepared by adding 2g of isatin to 60ml of benzyl alcohol, stirring for 2 hours, filtering and adding 2.5g of t-butoxycarbonyl phenylalanine to 50ml of the filtrate). The beads were heated for 5 minutes at 100°C and then allowed to cool before being washed x3 with acetone. A blue/ red colour indicated the presence of imino or amino groups, whilst colourless beads were indicative of complete coupling (Atherton and Sheppard,1989).

1.4 Cleavage of Peptides from Resin.

The final cleavage step is a critical factor in determining the overall yield and the quality of the peptide. The cleavage of the peptide from the solid support is combined with the removal of the side chain protecting groups. The 4-hydroxy-methyl-phenoxy-acetate linkage between the peptide and the resin is acid sensitive and is readily cleaved by 95% TFA in 1-2 hours. The cations generated in this reaction are relatively stable and have the potential to react with electron-rich amino acid side chains such as those of methionine, cysteine, tryptophan and tyrosine. This risk can be reduced by the inclusion of carbonium scavengers in the cleavage mixture. The cleavage of peptides including arginine in their sequence will be dealt with separately since the side chain of arginine is protected by the 4-Methoxy-2,3,6-trimethylbenzene sulphonyl- group (Mtr), which is removed at a much slower rate than the other side chain protecting groups.

(a) Cleavage of Non-arginine Containing Peptides.

0.5g of resin were placed in a clean, dry round bottom flask along with 19ml of TFA, 0.5ml of anisole and 0.5ml of ethanedithiol. The flask was attached to the base of a rotary evaporator in the vertical position (no vacuum) which was switched on at low speed to provide mixing. 100µl samples were taken from the flask after 1, 2 and 3 hours and dried down in eppendorff tubes under a stream of N_2. The residue was taken up in 100µl of ddH_2O and extracted with 2x 200µl of diethyl ether. The aqueous layer was filtered through a 0.2µ nylon filter before being analysed by reverse phase chromatography on an HPLC (see below). A typical trace showing the cleavage of the octapeptide MSGGAFSW can be seen in **fig.1.9a**.

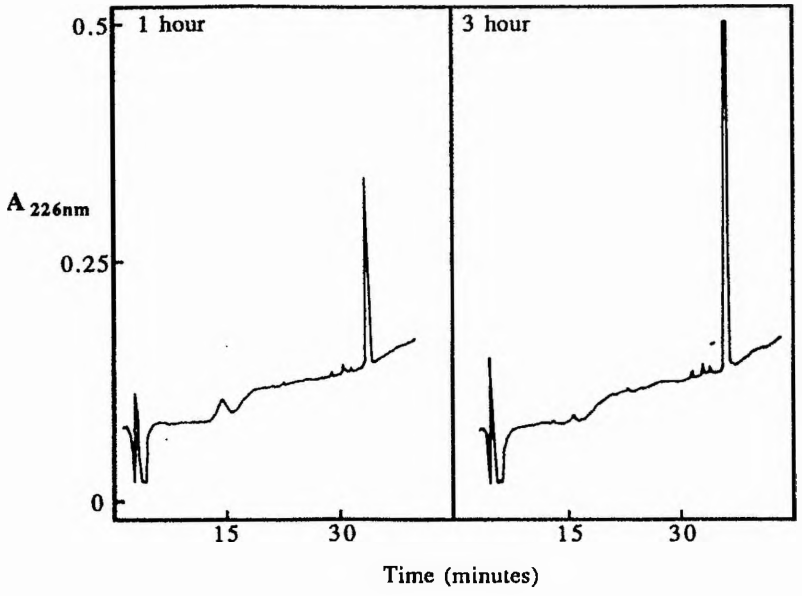


Fig. 1.9a Cleavage of MSGGAFSW
HPLC traces showing the cleavage of the peptide MSGGAFSW as an example of a non-arginine containing peptide. The full scale deflection was set at 0.5 and the solvent gradient was from 0 to 50% B in 20 minutes as described in method 1.5 below. The traces on the left and right show the amounts of peptide cleaved from the resin after 1 and 3 hours exposure to TFA respectively (see method 1.4 above).

After 3 hours, or when the cleavage reaction appeared to be complete, the resin was filtered through a sintered glass funnel and washed with 2x 10ml of TFA. The TFA was removed by rotary evaporation under vacuum at 40°C leaving an oily residue. 30ml of diethyl ether were added to the residue and after the suspension had settled the ether was decanted or pipetted off. This process was repeated 4 or 5 times until all traces of ethanedithiol and anisole had been removed. After the final wash the last traces of ether were removed by rotary evaporation and at this stage the peptides appeared as a dry, white coat on the inside of the flask. The peptides were taken up in 5ml of water and lyophilised.

(b) Cleavage of Arginine Containing Peptides.

0.5g of resin were placed in a flask along with 19ml TFA, 333mg of redistilled phenol, 333µl of ethanedithiol and 333µl of anisole. The flask was mixed and samples were taken each hour and analysed by reverse phase HPLC, as described for the non-arginine containing peptides. The removal of arginine can be readily monitored by HPLC as can be seen from the traces in **fig.1.9b** showing the cleavage of peptide, LAGGFRHR. The peptide with both arginines protected by Mtr groups is the most hydrophobic, followed by the peptide with a single Mtr group, with the fully deprotected peptide having the shortest retention time.

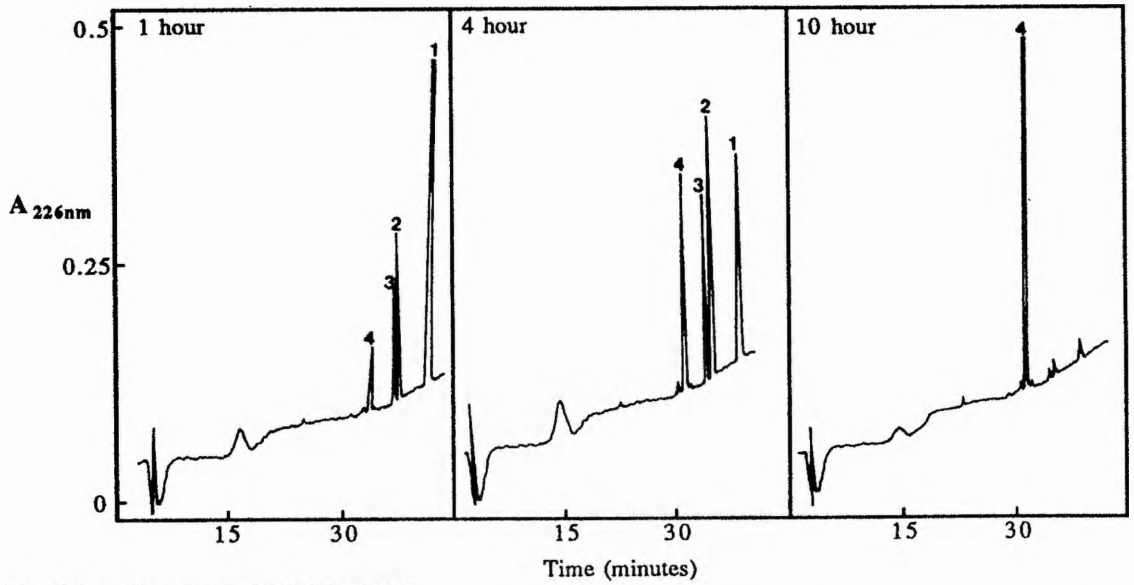


Fig. 1.9b Cleavage of LAGGFRHR
HPLC traces showing the cleavage of the peptide LAGGFRHR, as an example of a peptide that contains two arginines. The peak labelled 1 corresponds to the peptide with both arginines protected with Mtr groups, whilst peaks 2 and 3 are the peptides with a single Mtr group (one of the arginine residues protected). Peak 4 is the fully deprotected peptide. The details of the cleavage and HPLC are given in methods 1.4b and 1.5 respectively.

Generally it was found that peptides with a single arginine were fully deprotected after 6 hours, those with 2 arginines required 10 hours whilst those with 3 or 4 arginines required to be left overnight in TFA. After deprotection was complete the drying and ether extraction steps were identical to those described for non-arginine containing peptides. Fmoc pentafluorophenyl ester derivatives of arginine are now available with different side chain protection groups that are more susceptible to acid hydrolysis (Ramage and Green,1987). The use of such a derivative would remove some of the problems associated with arginine containing peptides.

1.5 Monitoring of Cleavage of Peptides by Reverse Phase HPLC.

A reverse phase column (18 μ bondpack-Anachem) was used to monitor the cleavages of the peptides synthesised. The solvents used were (A) 0.1% TFA in ddH_2O and (B) 0.1% TFA in acetonitrile. The peptides were separated on a solvent gradient from 0 to 50% B in 20 minutes and then from 50 to 100% B in 10 minutes. The flow rate was 1ml per minute and the flow through was monitored at 226 nm on a full scale deflection of 0.5. The peak areas were determined where necessary using a Shimadzu C-R1B integrator.

2 ANALYSIS AND PURIFICATION OF PEPTIDES.

2.1 Purification of Peptides By Reverse Phase FPLC.

Peptides were purified on a reversed phase column (PepRPC, Pharmacia) using the solvents (A) 0.1% TFA in ddH_2O and (B) 0.1% TFA in acetonitrile. The solvent gradients used varied depending on the retention times of the individual peptides, but all went from 0 to 100% B in 35 minutes. Typically 5mg of peptide was loaded onto the column per run. Most of the peptides were taken up at a concentration of 5 mg/ml in ddH_2O, but in the case of some of the more acidic peptides 0.1M ammonium bicarbonate was used to aid solubility. The full scale deflection was 2 and the flow through was monitored at 226nm. The flow rate was 0.7ml/ minute, 0.7ml fractions were collected and fractions corresponding to the peptide peaks were pooled and dried down by rotary evaporation. Peptides were washed with ether to remove residual TFA, dried, taken up in ddH_2O and lyophilised as before. The compositions of all the peptides were confirmed using amino acid analysis.

2.2 Amino Acid Analysis.

10μmoles of each peptide were hydrolysed in 300μl of 6M HCl for 5 hours at 108°C. Hydrolysates were dried in *vacuo*, and amino acid analysis carried out using reverse phase HPLC and pre-column derivitisation with o-phthaldialdehyde (OPA). The method used was based on that described by Roth (1971).

The OPA reagent was prepared by dissolving 10mg of OPA (Sigma) in 100μl β-mercaptoethanol, 5ml of 0.8M sodium borate pH12.5 and 4ml of H_2O and was stable for 3 days. The hydrolysed peptide was resuspended in 15μl of H_2O, filtered and added to 30μl of OPA reagent, mixed and left for 90 seconds before injection onto an HPLC reverse phase column (Ultrasphere-ODS, Altex). Derivatised amino acids were separated using the gradient:

Time (minutes)	%B
0	20
5.5	20
50	100
50.1	20
55	20

The buffers used were (A) 0.1M sodium acetate pH5.3 and (B) 90% methanol and the flow rate was 1ml per minute. The excitation and emission wavelengths were 340nm and 450nm respectively and the full scale deflection was 0.5. Peak areas were determined using a Shimadzu C-R1B integrator and retention times were compared with those of standard amino acid mixtures. It is noted that tryptophan is destroyed by the hydrolysis procedure so cannot be detected using this method.

3 CELLS AND ADENOVIRUS PREPARATION

3.1 Maintenance of HeLa Cells in Spinner and Monolayer Cultures.

HeLa cells were grown in suspension culture in S-MEM (GIBCO-BRL) with 5% newborn calf serum (Seralab) at 37°C to a density of 5 x 10^5 cells/ml. They were passaged every 3 to 4 days to give a cell concentration of 2 x 10^5 cells/ml.

Monolayer HeLa cultures were maintained in plastic 75cm^3 bottles (Sterilin) in 25ml modified G-MEM (GIBCO-BRL) supplemented with 10% newborn calf serum (Seralab) in an

atmosphere of 5% CO_2 at 37°C. Cells were passaged every 3 to 4 days on reaching confluence as follows. Spent medium was poured off and the cells were washed with 5ml of Trypsin/EDTA. 5ml of fresh Trypsin/EDTA were added and left on the cells for 2 minutes or until cells started to come off the plastic. Trypsin/EDTA was poured off and the cells were resuspended in 10ml of MEM/10% calf serum. 2ml of the cell suspension were transferred to a new 75 cm^3 bottle along with 25ml of MEM/10% calf serum and returned to 37°C.

Procedures for checking the viability of cells in tissue culture and storing them at -70°C are described by Freshney (1987).

3.2 Preparation of Adenovirus Type 2

6 litres of HeLa cells were grown in suspension culture, as above, to a density of $3x10^5$ cells per ml. The cells were centrifuged for 15 minutes at 2500rpm (1500g) in a Mistral 6L centrifuge and then resuspended in 250ml of S-MEM prior to infection with 1ml of Ad2 arcton extract (10^9 p.f.u./ml). Virus was allowed to absorb onto the cells in the absence of calf serum for 2 hours at 37°C after which the cells were made up to 6 litres with S-MEM and newborn calf serum was added to 2%. Cells were harvested at 72 hours post infection by centrifugation as before. The infected cell pellet was washed x1 with cold phosphate buffered saline pH6.2 (PBS) and then resuspended in 48ml of 10mM Tris/HCl pH7. 12ml of the cell suspension and 10ml of trichloro-trifluoro-ethane (arcton) were added to each of four 50ml sterile centrifuge tubes and shaken for 30 minutes at 4°C using a mechanical shaker. Extracts were centrifuged at 2500rpm (800g) for 15 minutes at 4°C using an IEC CENTRA-3R bench top centrifuge. The top (aqueous) layer was removed and clarified by centrifugation at 10000rpm (10000g) for 30 minutes at 4°C using a JA17 rotor in a Beckman J2-21 centrifuge before being stored at -70°C. This fraction contained infectious virus and was used either as the seed for subsequent virus preparations or subjected to further purification as described below. The titre of the virus was determined by carrying out plaque assays as described by Freshney (1987).

3.3 Purification of Adenovirus Type 2.

Purified virus was obtained by fractionating the arcton extracts, prepared in 3.2 on

caesium chloride density gradients as described previously by Russell and Blair (1977). Gradients were set up in 14cm^3 ultraclear centrifuge tubes (Beckman) by pipetting 3ml of 2M CsCl, 10mM Tris/HCl pH7.5, 1mM EDTA into each tube. This was gently underlayed with 3ml of 3M CsCl, 10mM Tris/HCl pH7.5, 1mM EDTA using a syringe to give a step gradient. The tubes were filled to about 3mm from the top by overlaying the caesium chloride with the arcton extract. Tubes were balanced and centrifuged in the sw40Ti rotor in a Beckman L8-55M ultracentrifuge for 2 hours at 30 000rpm (100000g) and 20°C. The virus band was removed using a syringe (**fig.1.10**). The virus was stored in 50% glycerol at -20°C. When

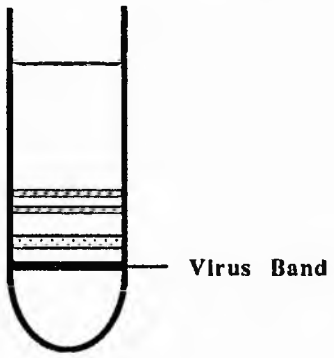


Fig. 1.10 Purification of Ad2
Diagram showing the position of the virus band after centrifugation of an arcton extract on a discontinuous CsCl gradient (method 3.3). The higher bands include empty virus particles and are referred to as top components.

highly purified virus was required the viral band was diluted 1:3 with 10mM Tris/HCl, 1mM EDTA, pH7.5 and purified on a second CsCl gradient. This time a continuous gradient was prepared, from 5ml 3M CsCl and 5ml 2M CsCl, in each tube using a gradient maker. The CsCl gradients were overlayed with the diluted virus band and centrifuged for 18 hours as before. The viral band was removed and stored as described above.

4 ASSESSMENT OF PEPTIDES AS SUBSTRATES AND INHIBITORS

4.1 Preparation of Disrupted Virus

Disrupted, purified Ad2 was used as the source of protease in all the experiments described in this chapter. The purified virus band, prepared as described in method 3.3, was dialysed for 15 hours at 4°C against 10mM Tris/maleate, 20% glycerol, pH6.8 and then stored in 100µl aliquots at -70°C.

4.2 Assessment of Peptides as Substrates of the Ad2 Protease

5µg of each peptide (1mg/ml in 25mM phosphate, 0.2M NaCl, pH7) were incubated with 5µl of disrupted purified virus for 0, 1, 2, 5 and 12 hours at 37°C. Reactions were

stopped by the addition of 0.1% TFA and the samples were filtered through 0.2μ nylon filters prior to analysis by HPLC. Peptides were separated on a reverse phase column (Ultrasphere-ODS, Altex) using a number of different gradients based on:

Time(minutes)	%B
0	0
5	0
10	20
25	35
30	100
30.1	0
40	0

with the buffers being (A) 0.1% TFA and (B) 0.1% TFA in acetonitrile. The flow rate was 1ml per minute and the full scale deflection was set at 0.5, for tryptophan containing peptides, or otherwise at 0.1. The flow through was monitored at 226nm.

In order to confirm that digestion had taken place and if so, to ascertain the point of cleavage, it was necessary to carry out amino acid analysis. 50μmoles of each peptide were incubated with 25μl of disrupted purified Ad2 at 37°C. Where digestion was evident, an incubation time giving about 50% digestion was chosen. Where preliminary analysis gave no indication of digestion, 12 hour incubation times were chosen. Samples were analysed by HPLC essentially as described above, but 1ml fractions were collected and the full scale deflection was set at 1-2. All the fractions were dried down and hydrolysed in preparation for amino acid analysis which was carried out as described in method 2.2.

4.3 Assessment of Peptides as Inhibitors of the Ad2 Protease

HeLa cells, infected with wild type Ad2, were labelled with ^{35}S methionine for 2 hours at 48 hours post infection and harvested as described in the methods section of chapter 2. 10μl of the resulting cell extract were incubated with 10μl of each of the peptides (1mg/ml) for 6 hours at 37°C. Reactions were stopped by boiling and the samples were analysed by SDS-PAGE and autoradiography as described in chapter 2.

5 COMPUTER MODELLING AND SEARCHES

The structures of the peptides were simulated on the CHEM-X software package (Chemical Designs Ltd.) by energy-minimisation of random structures. The Staden

programmes (Staden,1986) were used to translate the Ad2 genome in the 6 different reading frames and to search for proteins containing the sequences M(L,I)XGG and M(L,I)XGXG. Programmes in the Wisconsin computing package were used to carry out hydrophobicity and secondary structure predictions of the adenovirus proteins (Devereux *et al.*,1984).

RESULTS

1 Peptide Substrates for Ad2 Protease

The initial aim of this section is to establish whether or not synthetic peptides can be used as substrates for the adenovirus protease. Five peptides, based on the established cleavage sites in pVI and pVII, were synthesised and incubated for 1-12 hours with disrupted, purified Ad2 as the source of protease (**table 1.3**). The samples were analysed by reverse phase HPLC and fractions were collected for amino acid analysis.

The octapeptides MSGGAFSW (pVI) and MFGGAKKR (pVII) were both cleaved by the protease and amino acid analysis confirmed the GA bond to be the scissile one in each case.

Protein	Peptide	Cleaved
pVII	F G G - A K K	No
	M F G G - A K K	Yes
	M F G G - A K K R	Yes
pVI	S G G - A F S W	No
	M S G G - A F S W	Yes

Table 1.3 Minimum requirements for a synthetic peptide substrate
The above table shows 5 peptides that were synthesised based on the Ad2 protease cleavage sites in pVI and pVII. The dash indicates the position of the scissile bond in the natural substrate and the column on the right indicates whether or not the peptides were cleaved by the protease *in vitro*.

In contrast, the corresponding peptides lacking the P_4 methionine, SGGAFSW and FGGAKK, were not cleaved by the protease. These results show that a minimum requirement for peptide substrates of the Ad2 protease is that 4 amino acids must be present on the N-terminal side of the cleaved bond. The specific cleavage of MFGGAKK at its GA bond indicates that the maximum number of amino acids required to the C-terminus of the scissile bond is 3 (**table 1.3**).

The FPLC profiles of the 1 hour digestions of MSGGAFSW and MFGGAKKR are shown in **fig.1.11** alongside the CHEM-X models of the 2 peptides. The efficiency of the MFGGAKKR cleavage was slightly greater than that of MSGGAFSW, with about 10% more peptide being cleaved in the first hour. Comparing the sequences of the pVI and pVII cleavage

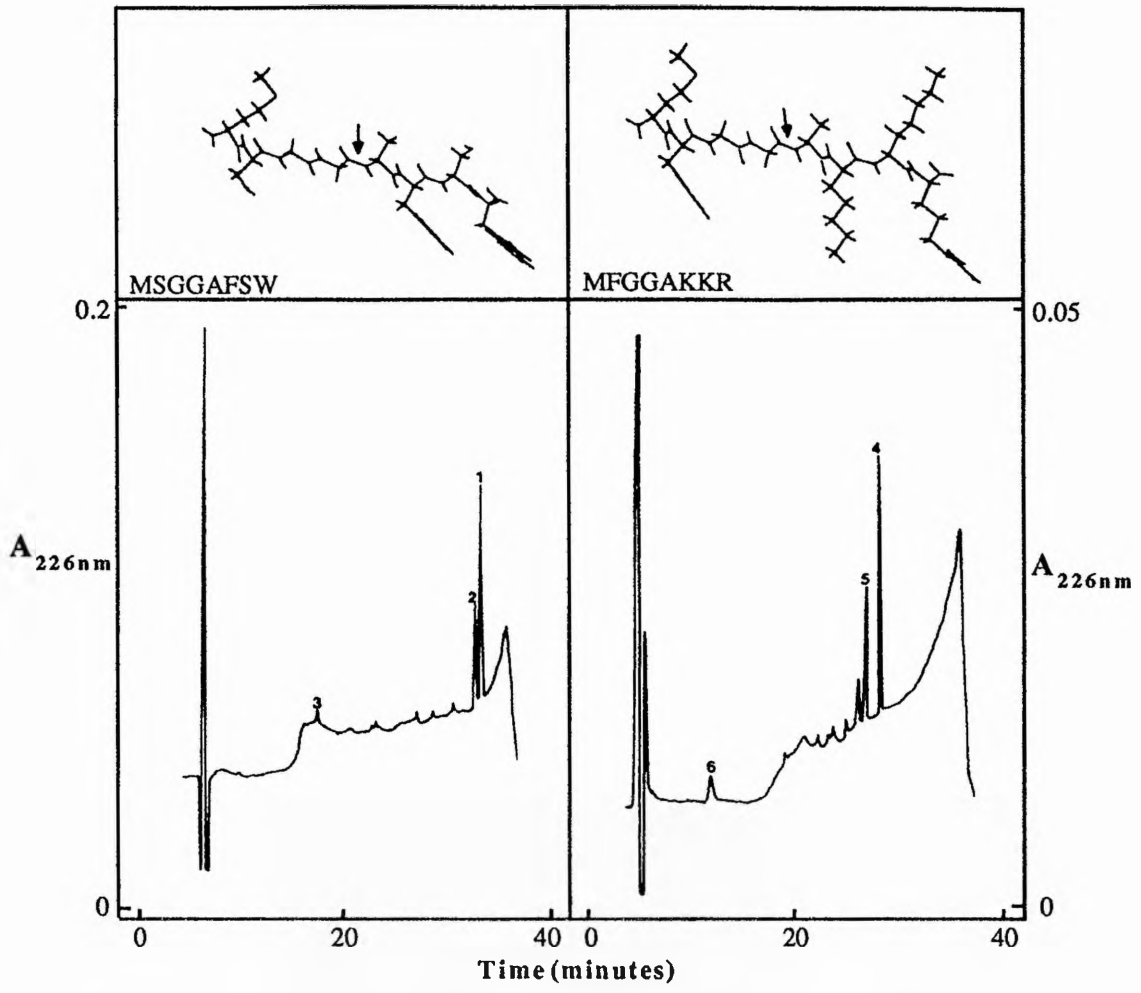


Fig. 1.11 Digestion of the peptides MSGGAFSW and MFGGAKKR
HPLC profiles of 1 hour incubations of the peptides MSGGAFSW and MFGGAKKR with disrupted purified Ad2 as a source of protease. Amino acid analysis identified the peaks as (1) MSGGAFSW, (2) AFSW, (3) MSGG, (4) MFGGAKKR, (5) MFGG and (6) AKKR. The gradients for the separations are described in method 4.2. Above each profile is the CHEM-X molecular model (see method 5) of the appropriate octapeptide and the arrow marks the position of the cleavage site.

sites, it is immediately apparent that the residues occupying the P_2' -P_4' positions, FSW in pVI and KKR in pVII, are completely different. Furthermore, it is noted that the P_3 position is occupied by a serine in the pVI peptide and a phenylalanine in the pVII peptide. These residues differ significantly in size and polarity, suggesting that P_3 does not contribute to the specificity of recognition. It seems probable, therefore, that the primary substrate specificity of the Ad2 protease is contained within positions P_4, P_2, P_1 and P_1'. Accordingly, a close examination of the requirements at each of these positions was undertaken.

2 Consensus Sequence for Ad2 Protease Substrates.

The pVI derived octapeptide, MSGGAFSW, was selected as the reference peptide for these studies. It was chosen in preference to MFGGAKKR because the presence of a tryptophan residue facilitates detection at 226nm.

The residues at P_4, P_2, P_1 and P_1' were varied in turn to produce 12 distinct octapeptides (**table.1.4**). Each was incubated with disrupted Ad2 and the resulting samples were analysed by HPLC and amino acid analysis as described in method 4.2. Five of the peptides, LSGGAFSW, ISGGAFSW, MSGGGFSW, MSGAGFSW and MSGSGFSW were cleaved efficiently by the protease at GA, GA, GG, AG and SG bonds, respectively, with tetrapeptides being generated in each case. The remaining 7 peptides VSGGAFSW, YSGGAFSW, QSGGAFSW, ESGGAFSW, TSGGAFSW, MSAGAFSW and MSGSAFSW were not cleaved by the protease, even after 12 hour incubations.

Postion varied	Peptide	Relative susceptibility
	M S G G - A F S W	1.0
P_4	**Y** S G G - A F S W	0.0
	Q S G G - A F S W	0.0
	E S G G - A F S W	0.0
	T S G G - A F S W	0.0
	V S G G - A F S W	0.0
	L S G G - A F S W	0.8
	I S G G - A F S W	0.9
P_2	M S **A** G - A F S W	0.0
P_1	M S G **S** - A F S W	0.0
P_1'	M S G G - **G** F S W	1.5
P_1 & P_1'	M S G **A** - **G** F S W	1.5
	M S G **S** - **G** F S W	1.2

Table 1.4 Cleavage of octapeptide variants of the pVI site.
The ability of a series of octapeptides to act as substrates for the Ad2 protease was examined. The right hand column shows the cleavage rate of each peptide relative to the reference peptide MSGGAFSW; with a value of 0.0 indicating that the peptide was not cleaved.

The P_4 Position

The substitution of 7 amino acids for the methionine at P_4 in the reference peptide

revealed that the Ad2 protease has very specific requirements at this position. Peptides with valine, tyrosine, glutamine, glutamate or threonine at P_4 were not hydrolysed by the protease. Indeed, of the amimo acids tested, only methionine, leucine and isoleucine were found to be permissible at P_4. Thus one of the requirements of the Ad2 protease specificity appears to be that a hydrophobic residue must be present at P_4. The observation that valine is not acceptable indicates that the size of the side chain is also important and strongly suggests that only methionine, leucine and isoleucine will be recognised by the adenovirus protease at P_4.

The P2 Position

The resistance of the peptide, MSAGAFSW to the protease reveals that even the conservative change of a glycine to an alanine at P_2 is sufficient to prevent cleavage. This result gives a strong indication that the Ad2 protease will only cleave peptides with a glycine at P_2. Thus the P_2 position is a primary determinant of the enzyme's specificity.

The P_1 Position

The requirements at the P_1 position are not as clear and are dependent, to some extent at least, on the identity of the residue at P_1'. This is demonstrated by the fact that the glycine to serine substitution at P_1 in MSGSAFSW prevented cleavage, but that an additional alanine to glycine substitution at P_1' resulted in the efficient cleavage of the peptide, MSGSGFSW. It appears, therefore, that a serine is only permissible at P_1 if a glycine is present at P_1'.

The P_1' Position

Rate studies reveal that the peptides with a glycine substituted for the alanine at P_1' were cleaved more efficiently than MSGGAFSW, suggesting that glycine is the preferred amino acid at this position. Indeed, MSGGGFSW proved to be the best substrate of all being cleaved at a rates 50% greater than MSGGAFSW (**table 1.4**).

Based on this systematic study M(L,I)XGX-X is suggested as a preliminary consensus sequence to describe all of the peptides cleaved by the Ad2 protease. It is noted, however, that

some peptides (e.g. MSGSAFSW) fit this consensus sequence, but are not hydrolysed by the protease, so clearly a more specific notation is required. At present, all the peptides cleaved by the protease are either of the form:

M(L,I) X G G - X or M(L,I) X G X - G

and so it is proposed that these sequences define the substrate specificity of the Ad2 protease. It is conceded, however, that some peptides conforming to these sequences will probably not be cleaved. This is because an exhaustive study of the amino acids, permitted at P_3, P_1 and P_1', has not been carried out and it is likely that there will be additional restrictions at some of these positions. For example, a charged residue such as arginine or aspartate may not be permitted at P_1' and the presence of a proline residue anywhere in the sequence might have a detrimental effect on the peptide's secondary structure. It is also possible that some peptides of the form M(L,I)XGX-X will be substrates of the Ad2 protease. Nevertheless, the primary determinants of the enzyme's substrate specificity have been shown to reside in the P_2 and P_4 positions and the validity of the above consensus sequences will be tested by an examination of the natural substrates of the Ad2 protease.

3 Cleavage Sites in Ad2 Proteins.

The Ad2 genome was translated in its 6 reading frames using the Staden programmes and a search was carried out to determine how frequently the sequences M(L,I)XGG-X and M(L,I)XGX-G occur in viral proteins. The search identified 25 such sites in Ad2 open reading frames, including 14 sites in pVI, pVII, pVIII, pTP, IIIa and the 11-kDa protein (**table 1.5**).

Peptides corresponding to the 4 predicted cleavage sites in the pTP and one of the sites in pVIII were synthesised and assessed as substrates for the Ad2 protease as described above. The peptides LAGG-FRHR (pVIII), MRGF-GVTR and MGGR-GRHL (both pTP) were cleaved at rates of 0.2, 0.5 and 0.6 relative to the reference peptide MSGG-AFSW (**fig.1.12**). Amino acid analysis confirmed that the hydrolysed bonds were G-F, F-G and R-G respectively. Thus at least 3 cleavage sites in Ad2 proteins appear to have been successfully predicted, 2 in the pTP and 1 in pVIII. In contrast, the peptides corresponding to the 2 remaining predicted sites in the pTP, SLGGGVPW and MTGGVFQ were not hydrolysed

Protein	Putative Cleavage Site	Turn Predicted?	Hydrophilic Index	Serotypes Sequenced	Conserved?
pVI	MSGG-AFSW	Yes	0.7	2	-
	IVGL-GVQS	No	-1.0	2,41	Yes
pVII	MFGG-AKKR	Yes	1.2	2,5	Yes
pVIII	LAGG-FRHR	Yes	1.2	2,3,41	Yes
	IRGR-GIQL	Yes	1.1	2,3,41	Yes
	IGGA-GRSS	Yes	0.8	2,3,41	Yes
pTP	MRGF-GVTR	Yes	0.8	2,5,7,12	Yes
	MGGR-GRHL	Yes	1.8	2,5,7,12	No
	LGGG-VPTQ	Yes	-0.6	2,5,7,12	No
	MTGG-VFQL	Yes	1.5	2,5,7,12	Yes
IIIa	LGGS-GNPF	Yes	0.7	2,5	Yes
11K	MAGH-GLTG	No	1.3	2	-
	LTGG-MRRA	Yes	1.0	2	-
	MRGG-ILPL	Yes	2.2	2	-
Hexon	LGGI-GVTD	Yes	-0.5	2,5,40,41	No
	LLGN-GRYV	Yes	0.4	2,5,40,41	Yes
Penton	LEGG-NIPA	Yes	1.5	2,5	Yes
	LPGC-GVDF	Yes	-2.4	2,5	Yes
L1 52K	LAGT-GSGD	Yes	0.2	2	-
Polymerase	IRGG-RCYP	Yes	1.5	2,5,7,12	Yes
E1B 21K	LSGG-YLLD	Yes	0.7	2,4,5,7,12 40,41	No
E3B 14.5K	LTGG-DD	Yes	1.1	2,3,5,7	Yes
E1B 52K	ISGN-GAEV	Yes	0.4	2,5,7,12, 40,41	Yes
D-172	ISGG-VLEG	Yes	-2.0	2	-
	LAGS-GSGS	Yes	0.3	2	-
E4 34K	IAGG-QVLA	Yes	-0.1	2	-

Table 1.5 Secondary Structure and Conservation of Predicted Cleavage Sites
Cleavage sites predicted by the concensus sequences are shown as octapeptides, with the exception of the E3B 14.5K sequence which is at the C-terminus. Turn regions, and the maximum hydrophilicity value between P_4 and P_1' were predicted using the Peptidestructure program in the Wisconsin package (Devereux *et al*, 1984).

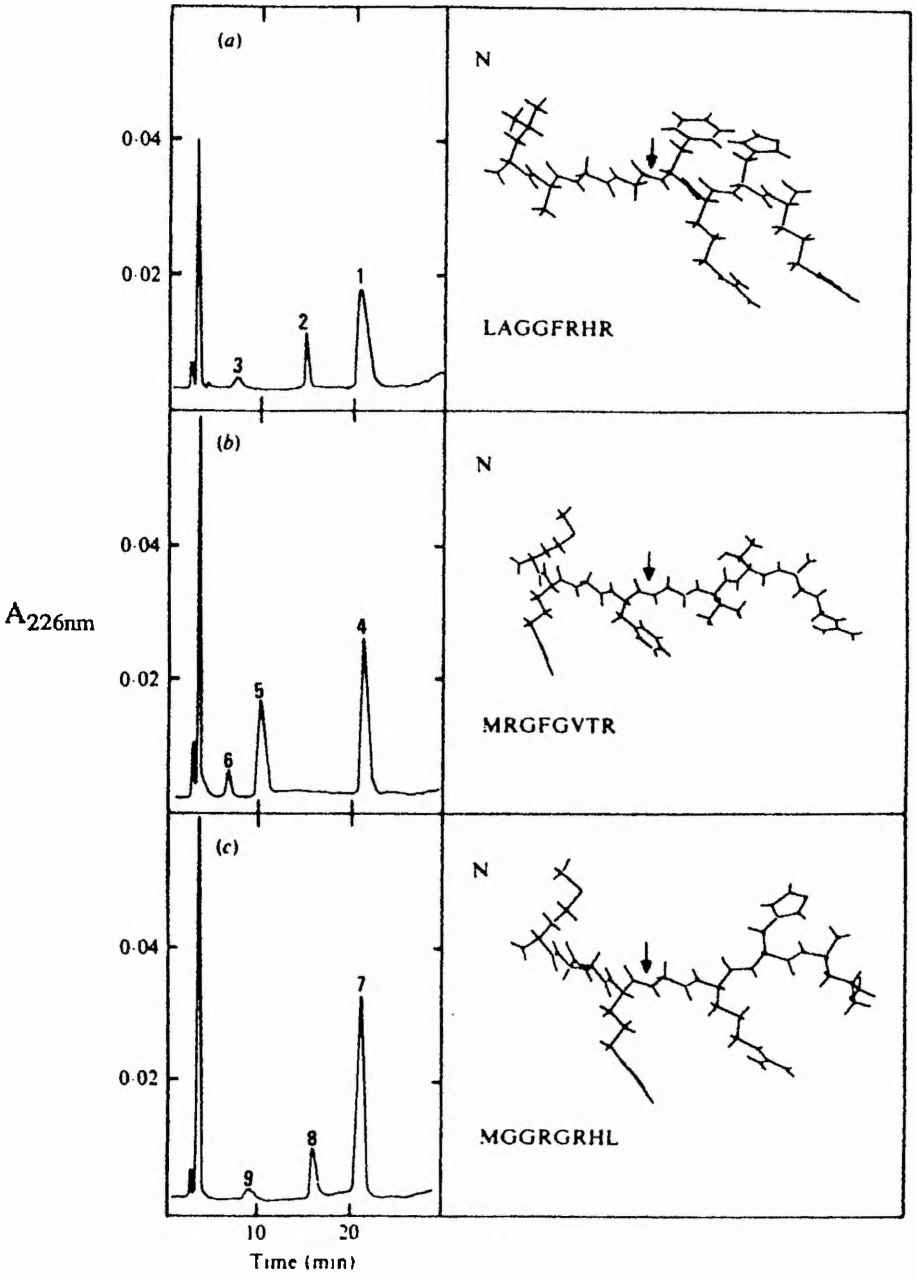


Fig. 1.12 Confirmation of putative cleavage sites in pVIII and pTP.
FPLC traces (PepRPC column) from protease digests of (a) LAGGFRHR from pVIII, (b) MRGFGVTR from the pTP and (c) MGGRGRHL from the pTP. Incubations of 100μg of peptide with 20μl of purified disrupted Ad2 at 37°C were for 12 hours (LAGGFRHR), 5 hours (MRGFGVTR) and 2 hours (MGGRGRHL). Amino acid analysis showed the numbered peaks to contain (1) LAGGFRHR, (2) FRHR, (3) LAGG, (4) MRGFGVTR, (5) GVTR, (6) MRGF, (7) MGGRGRHL, (8) GRHL and (9) MGGR. Peptides were eluted with a 5 minute isocratic step at 0.1% (v/v) TFA, followed by a 25 minute gradient to 0.1% (v/v) TFA in 20% (v/v) acetonitrile. CHEM-X models are shown in adjacent panels with the cleaved bond arrowed.

by the protease and further studies are required to confirm the identity of the scissile bond in the pTP giving rise to the TP (see chapter 4).

Two additional approaches were taken in an attempt to establish the significance of the predicted cleavage sites in the viral proteins. Firstly protein sequences were examined in other adenovirus serotypes in order to determine which of the sites are conserved (**table 1.5**). Secondly the programmes from the Wisconsin package were used to predict the probable secondary structures and hydrophilicity values of all the potential cleavages sites and the values obtained are included in **table 1.5**. Five of the sites, namely LGGG-V in the pTP, IVGL-G in pVI, LGGI-G in the hexon, LPGC-G in the penton and ISGG-V in D-172 were predicted to be in hydrophobic regions of the protein so might not be surface exposed. The predicted secondary structure at all the potential cleavage sites, except MAGH-G in the 11-kDa protein and IVGL-G in pVI, was a β-turn. This perhaps is not surprising given that all the potential cleavage sites contain at least 2 glycine residues.

In conclusion the consensus sequences derived to define the substrate specificity of the Ad2 protease, have survived this examination of the viral proteins. In the first instance they were used to select a relatively small number of potential cleavage sites, including sites in all the known substrates of the enzyme. Furthermore, peptides corresponding to 3 of the predicted sites were synthesised and found to be cleaved efficiently by the protease. The work also confirms that there are additional restrictions at P_1' in peptide substrates, since SLGGGVPW and MTGGVFQ, which contain valine residues at P_1', were not hydrolysed, despite the fact that a glycine is present at P_1 in each case. This serves to remind us that not all peptides of the form M(L,I)XGX-G or M(L,I)XGG-X are substrates of the Ad2 protease and suggests that some of the sites predicted in **table 1.5** will not be cleaved.

4 Molecular Modelling.

The CHEM-X molecular modelling package was used to predict the structures of a series of peptides, including all those synthesised. The models were constructed by energy minimisation of random structures and a selection of models is included to portray the key elements of substrate specificity of the adenovirus protease (**fig.1.11**, **1.12** and **1.13**).

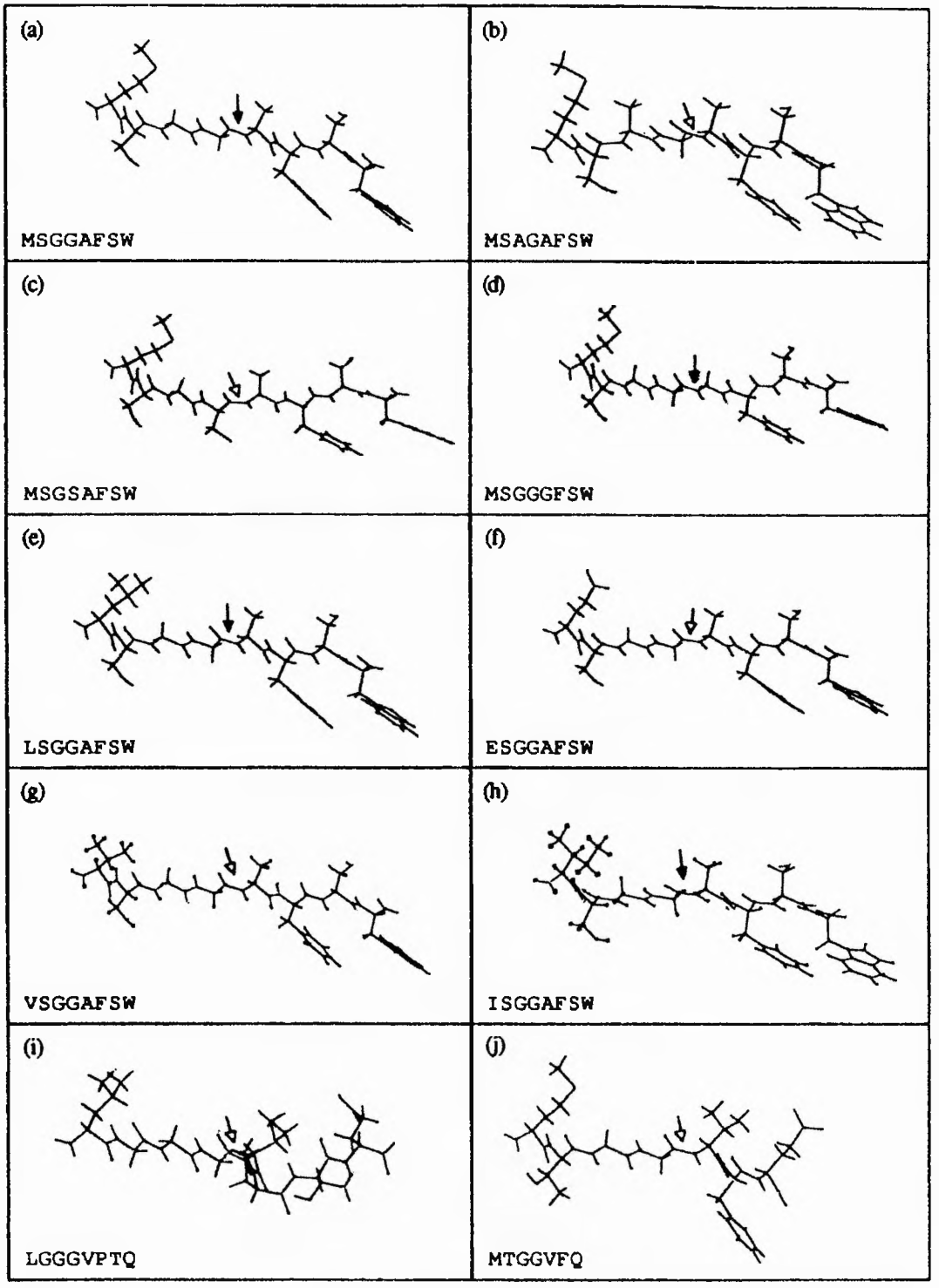


Fig. 1.13 CHEM-X Models of Selected Peptides

The models were created through the CHEM-X package using energy minimisation of random structures. The N-terminus is to the left; filled arrows indicate cleavage while open arrows indicate the analagous position in peptides which were not cleaved by the protease.

The P_2 and P_4 positions have already been shown to be the primary determinants of specificity and **fig.1.13b** demonstrates how even the small methyl side chain of alanine at P_2 protrudes and presumably is large enough to prevent access of the protease to the G-A bond. Comparisons of the models of MSGGAFSW (**fig.1.13a**) and ESGGAFSW (**fig.13f**) reveal only minor differences in the size and orientation of the amino acids at P_4. The fact that ESGGAFSW is not a substrate (**table 1.4**), suggests that there are specific requirements for a hydrophobic residue at P_4. Size, however, would still appear to be important since peptides containing the small branched hydrophobic amino acid, valine, at P_4 were not hydrolysed (**fig.1.13g**).

CHEM-X models (**fig.1.12b** and **c**) show that the side chain of P_1 is oriented in a direction away from those of the critical P_4 and P_2 residues. This would explain why peptides containing glycine, alanine, serine, phenylalanine and arginine at P_1 are all cleaved (**tables 1.4** and **1.5**). On the other hand, models of MSGGAFSW (cleaved) and MSGSAFSW (not cleaved) demonstrate that the residue at the P_1 position can affect the orientation of the P_1' residue and so in some cases does dictate whether or not a peptide will be cleaved (**fig.1.14a** and **c**).

The CHEM-X models also demonstrate that the side chain of the P_1' residue must be oriented away from the scissile bond. The branched side chain of the valine residues at P_1' in LGGGVPTQ (**fig.1.13i**) and MTGGVFQ (**fig.1.13j**) can be seen to protrude further across the P_1-P_1' bond than the alanine in MSGGAFSW (**fig.1.13a**). In contrast the bulky phenylalanine in the peptide, LAGGFRHR, which was cleaved, appears to be oriented well away from the scissile bond (**fig.1.12a**).

Molecular modelling has enabled us to visualise the features of secondary structure shared by all the peptide substrates of the Ad2 protease and these features are exemplified in the model of MSGGGFSW (**fig.1.13d**) which was found to be the best substrate.

5 Assessment of Peptides as Inhibitors of the Ad2 Protease.

Finally the possibility that some of the peptides synthesised, during the course of this project, might act as specific inhibitors of the Ad2 protease was considered. Particular attention was paid to the peptides that were not cleaved by the protease in these preliminary investigations.

The assay used relies on the use of ^{35}S methionine labelled Ad2 proteins as substrates, with the digestion of pVII to VII by the protease, being monitored by SDS-PAGE and autoradiography. A full description of the assay is given in chapter 2. The extracts used contained endogenous protease activity and were incubated for 6 hours at 37°C in the presence of a selection of peptides at concentrations of 0.5mg/ml. Only SGGAFSW and FGGAKK were found to inhibit the protease to any significant extent (**table.1.6**). Interestingly, these

Peptide	% Inhibition
MFGGAKKR	-
FGGAKK	55%
MSGGAFSW	-
SGGAFSW	60%
YSGGAFSW	-
QSGGAFSW	-
ESGGAFSW	-
TSGGAFSW	-
MSAGAFSW	-
MSGSAFSW	-
GGGGGG	-

Table 1.6 Assessment of peptides as inhibitors.
The effects of selected peptides on the cleavage of pVII to VII in the ^{35}S assay (see chapter 2). The peptides were incubated with the labelled Ad2 infected extracts for 6 hours at 37°C and the digestion of pVII to VII was monitored by SDS-PAGE and autoradiography. Quantitative data was obtained by scanning densitometry. In the control experiment without any added peptide 90% digestion of pVII to VII was observed.

peptides are based on the cleavage sites from pVI and pVII, respectively, but both lack the P_4 residue. Parallel incubations were carried out with $(Gly)_6$, but this peptide did not inhibit the protease. These results are initial observations that must be confirmed, in the future, in kinetically significant experiments. Neverthless, they may provide a starting point for the design of a specific inhibitor of the Ad2 protease.

DISCUSSION

Using peptide synthesis and computer modelling the substrate specificity of the Ad2 protease has been defined, with elements of both primary and secondary structure being important. The peptide, MSGGAFSW, based on the established pVI cleavage site, was selected as the reference peptide, since the presence of a tryptophan residue facilitated its detection. In terms of the primary structure of the substrates, the results show that the major requirements are at the P_2 and P_4 positions. P_2 must be a glycine and only methionine, leucine and isoleucine have been shown to be acceptable at P_4. The other major requirement appears to be that the side chain of P_1' must point away from the scissile bond. The secondary structure that enables this is best portrayed in **fig.1.13d** and is characterised by an area between the P_4 position and the cleaved bond devoid of intrusive side chains. At this stage it is important to note that proteolytic processing normally takes place at peptide bonds in relatively flexible interdomain segments or in surface loops of proteins (Neurath,1989). Thus the presence of a potential cleavage site in the primary structure of a protein does not necessarily mean that it will be accessible to the protease in the native molecule.

A search for the sequences M(L,I)XGGX and M(L,I)XGXG picked out 15 of the Ad2 proteins. These included all the known substrates of the protease, namely, pVI, pVII, pVIII, pTP, 11K and IIIa. The main purpose of this discussion will be to consider the significance of all the predicted cleavage sites in the Ad2 proteins. **Fig.1.14** shows the positions of the predicted sites in the known substrates of the protease.

pVI and pVII

In addition to the established cleavage sites at G_{33}-A_{34} and G_{24}-A_{25} in pVI and pVII, a second site was predicted towards the C-terminus of pVI at L_{240}-G_{241}. Evidence that this site is cleaved *in vivo* comes from work by Anderson (1990) who noted that the only cysteine in the Ad2 pVI is the penultimate residue at position 249 and went on to show that pVI, but not VI, was labelled with ^{35}S cysteine.

pVIII

pVIII, with an apparent MW of 26-kDa, is cleaved to generate a protein of apparent MW 13-kDa, as determined by SDS-PAGE (Weber, 1976). As yet, only G-A and A-G bonds have been considered as potential cleavage sites (Hannan *et al.*, 1983). Consequently, these authors proposed that either a single cleavage at amino acid 105 (TNSG-AQLA) with the retention of the amino terminus of pVIII, or a double cleavage at amino acids 22 (LAAG-AAQD) and 156 (QIGG-AGRS) takes place. None of these sequences, however, fits the consensus sequence, proposed in this project and so we would not expect any of them to be

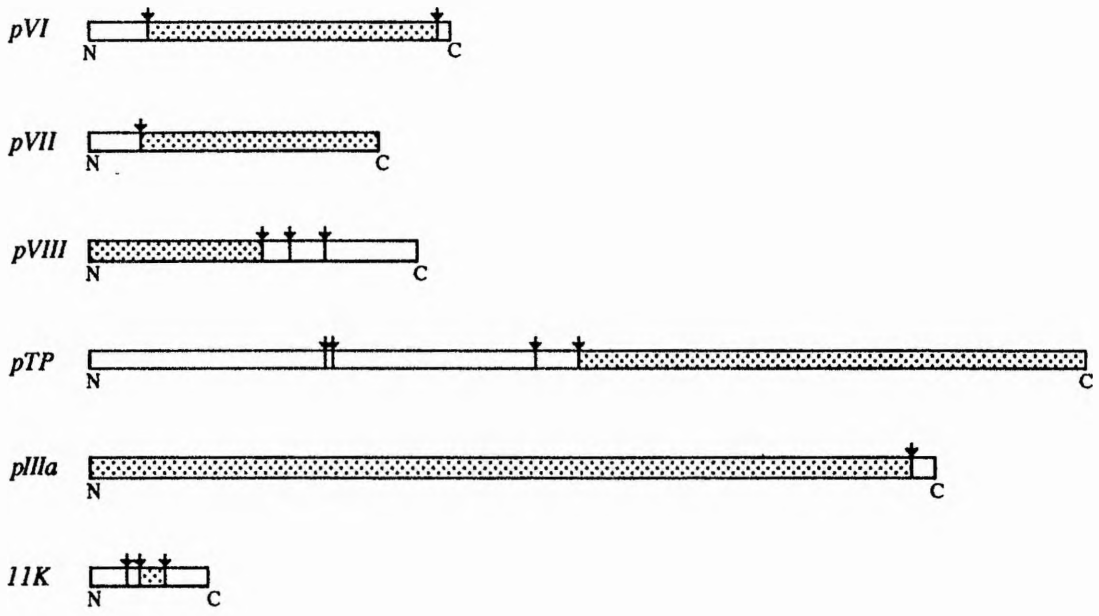


Fig. 1.14 Predicted Cleavage Sites in the Established Substrates of the Ad2 protease
The locations of the predicted cleavage sites in pVI, pVII, pVIII, pTP, pIIIa and the 11-kDa protein are indicated by arrows (see table 1.5). The shaded area represents the retained fragments of the proteins, namely VI, VII, VIII, TP, IIIa and mu respectively. The N- and C-termini are marked in each case.

cleaved by the Ad2 proteinase. Three sequences from pVIII, however, centred on amino acids 111, 131 and 157, respectively, do satisfy the conditions outlined above. These are LAGG-FRHR, IRGR-GIQL and IGGA-GRSS respectively. Cleavage at all of these sites would be consistent with the results of Hannan *et al.* (1983), whose comparative ^{125}I and ^{35}S labelling studies on pVIII and VIII suggest a cleavage site near the centre of the protein, with the retention of the N-terminus. The peptide based on the site centred at position 111, LAGG-FRHR, was synthesised and shown to be cleaved by the adenovirus protease, albeit at

a slower rate than any of the other peptide substrates. The fact that a phenylalanine was acceptable in the P_1' position was surprising, until a close examination was made of the computer model (**fig.1.12a**), which shows that, although this is a large and bulky residue, the ring structure is bent well away from the cleaved bond. It is proposed, therefore, that a cleavage at G_{111}-F_{112} releases the 13-kDa polypeptide VIII from the N-terminus and that the C-terminal portion of the precursor, which includes the 2 additional cleavage sites, is broken down further . The phenylalanine in P_1' is not only unexpected as far as the adenovirus protease is concerned, but is also unusual as far as substrates of viral proteases in general are concerned (Wellink and van Kammen, 1988). The only other viral proteases known to cleave a substrate with an aromatic group in the P_1' position are the alphavirus nsP2 proteases. In this case the nsP3-nsP4 cleavage involves a tyrosine at P_1' and processing takes place at a much slower rate than at the other cleavage sites in the polyprotein. Perhaps slow processing of the pVIII protein is required by the adenovirus and the presence of a phenylalanine at P_1' ensures this.

Preterminal Protein (pTP)

Smart and Stillman (1982) have shown that the pTP (87-kDa) is cleaved, via a 62-kDa intermediate (iTP), to the 55-kDa terminal protein (TP) when incubated with a source of Ad2 protease. Four putative cleavage sites were found in the pTP. Two of these were separated by only 8 amino acids; the sequences MRGF-GVTR and MGGR-GRHL, centred on residues 168 and 176 respectively. These peptides were synthesised and both were found to be cleaved by the protease. Cleavage at either or both of these sequences would give a protein with a MW of around 62-kDa. This would explain the presence of the intermediate when the pTP is cleaved. A split band at 62-kDa has been observed on some gel systems, indicating that it may contain 2 species (Leith *et al.*, 1989). The MRGF-GVTR, but not the MGGR-GRHL site is conserved in the pre-terminal proteins of the other adenovirus serotypes in which it has been sequenced suggesting that this is the main site giving rise to the iTP. The bond cleaved by the protease to give the 55-kDa terminal protein is slightly more elusive. Two sequences were predicted as potential cleavage sites; MTGG-VFQL and LGGG-VPTQ. The peptides MTGGVFQ and

SLGGGVPW incorporating the putative cleavage sites were synthesised, but neither was cleaved by the protease. The modelling studies on LGGGVPTQ and MTGGVFQL (**fig.1.13i** and **j**) show that the valine intrudes slightly more than the alanine at P_1' in MSGGAFSW. It seems probable that the cleavage site is at one of these 2 positions and that the tertiary structure, which cannot be mimicked by a 7 or 8 residue peptide, pulls the valine into an orientation which allows the protease access. The LGGG-V site is not conserved in the pre-terminal proteins from other serotypes that have been sequenced, suggesting that the MTGG-V site originally suggested by Smart and Stillman (1982) is the cleavage site giving rise to the mature 55-kDa terminal protein. Further support for this hypothesis comes from results which suggest that the TP is produced much more slowly than the iTP (see chapter 4). As with the processing of pVIII-VIII it could be advantageous to the adenovirus if the generation of the mature TP is controlled.

pIIIa

The 66-kDa capsid-associated viral protein, IIIa, has been shown to result from the cleavage of a slightly larger 67-kDa molecule (Boudin *et al.*,1980). The putative site, reported here, is situated only 15 residues from the C-terminus of the precursor polypeptide and cleavage at this point would account for the 1-kDa loss in molecular weight.

L2 11-kDa protein

The last of the viral proteins that has been shown to be cleaved by the protease is the L2 11-kDa protein. The computer search picked out 3 putative cleavage sites in the 11-kDa protein; MAGH-G, LTGG-M and MRGG-I. Cleavage at the LTGG-M site was previously demonstrated by Weber and Anderson (1988) and more recently Anderson *et al.*, (1989) located the second cleavage site in the 11-kDa protein that gives rise to the basic protein mu and they have used sequence analysis to show that this is at the G-I bond predicted in **table 1.5**. It seems probable, therefore that the 11-kDa protein is cleaved in 3 places with the G-M and G-I cleavages resulting in the production of the mature protein.

Other Adenovirus Proteins

Consider now the predicted cleavage sites in adenovirus proteins that have not been previously shown to be cleaved by the protease (**table 5, fig.1.15**). It is possible that some,

previously shown to be cleaved by the protease (**table 5, fig.1.15**). It is possible that some, if not all, of these sites are hidden in hydrophobic or highly structured areas of their proteins and so are not accessible to the protease. On the other hand, their processing may not have been noticed if the proteins are present at a low copy number, or if only a small percentage of the proteins are cleaved, perhaps as a result of a conformational change at a particular point in the infectious cycle. Surprisingly, both of the adenovirus coat proteins, the hexon and the penton, have putative cleavage sites near the centre of their amino acid sequences. The LLGN-G sequence at 564 to 569 of the hexon is located at the base of the tower region in a highly conserved segment of random structure (see **fig.2**, p8), whereas the LLGI-G sequence

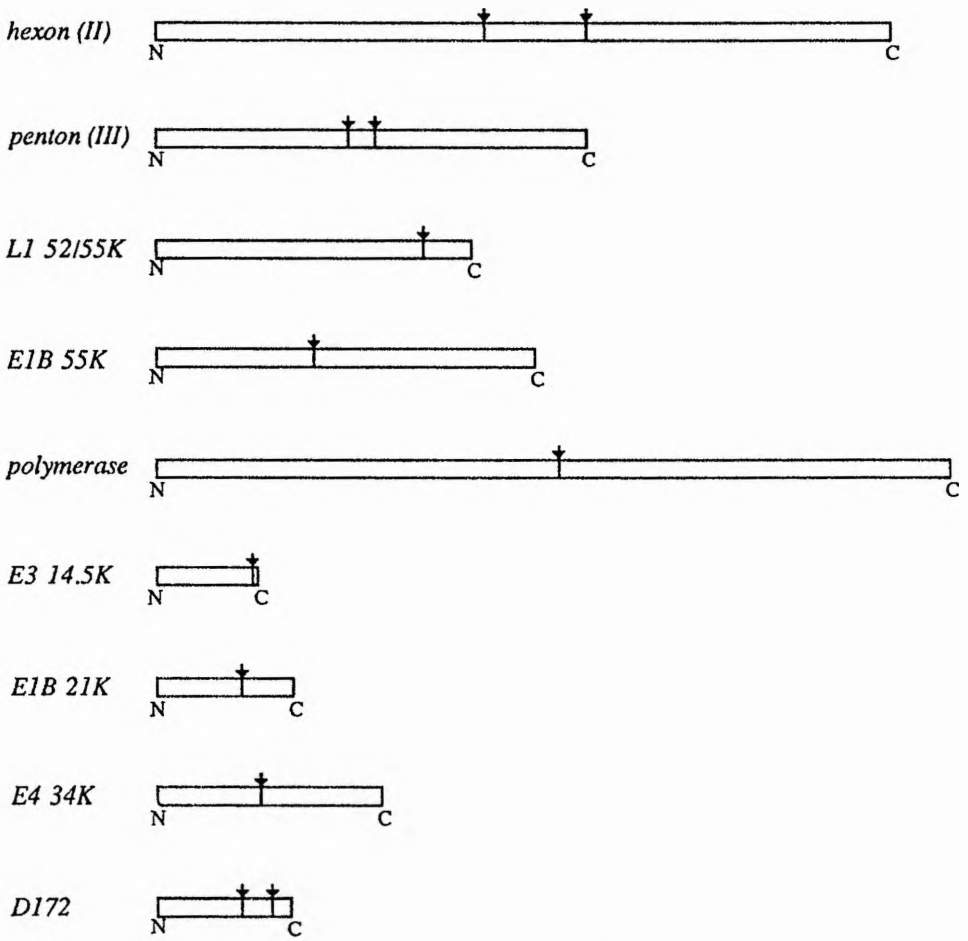


Fig. 1.15 Predicted Cleavage Sites in other Ad2 Proteins
The locations of the predicted cleavage sites in the hexon, penton, L1 52/55K, E1B 55K, E2B polymerase, E3 14.5K, E1B 21K, E4 34K and D172 proteins are indicated by the arrows (see table 1.5). 1cm represents 100 amino acids and the N- and C-termini of the proteins are marked.

(residues 428 -432) is in the tower region and is not conserved in the various serotypes (Roberts *et al.*, 1986; Toogood *et al.*, 1989). The position of these sequences within the

quarternary structure of the of the hexon, making them inaccessible from the inside of the virion, could explain why the protein does not appear to be cleaved. On the other hand, it has been suggested that viral proteases may play a part in uncoating (Krausslich and Wimmer, 1988). Although there is no direct evidence for this in the case of the adenovirus, the Ad2 ts1 mutant, which lacks protease activity, is defective in uncoating (Miles *et al.*, 1980; Hannan *et al.*, 1983). A report by Varga *et al.* (1990) reveals that the hexon undergoes a conformational change at low pH and that the N-terminus of the protein becomes exposed and susceptible to proteolysis. Since the major route of Ad2 internalisation is by receptor mediated endocytosis and the virions are subsequently located within endosomes, such a conformational change at low pH could have a significant part to play in uncoating. The penton base also undergoes a conformational change at low pH so the possibility remains that the putative cleavage sites in the coat proteins are hidden from the viral protease when the other maturation cleavages take place. On entering a mildly acidic environment, such as that of the endosomes, however, the coats of the virions undergo a conformational change, perhaps exposing the cleavage sites in the hexon and/or the penton to the protease present inside the virions. This could be the first step in uncoating and is an indication that the presence of putative cleavage sites in both of the major coat proteins of the adenovirus should not be dismissed without due consideration.

The potential cleavage site in the L1 52/55-kDa proteins is also of interest, since the latter are unusual among the viral late proteins, in that they are synthesised at early and late times during infection. This has led to proposals that they have a dual role in the lytic cycle. The 52/55-kDa proteins are phosphoproteins and are required for viral assembly, but they are not thought to be components of the mature virion (Hasson *et al.*,1989). It is noted that the predicted site, LAGT-GSGD has a glycine at P_1' and is not significantly different from some of the known substrates of the protease (**tables 1.4** and **1.5**).

Of the remaining predicted sites in adenovirus proteins only those in the E1B 55-kDa, E2B polymerase and E3 14.5-kDa proteins are conserved in other serotypes. In the polymerase and 14.5-kDa sites the charged residues arginine and aspartic acid are present at P_1' and it remains to be established that such amino acids are permissible at this position. Furthermore, the site in the E3 14.5-kDa protein is only 2 residues in from the C-terminus and

so processing would be difficult to detect. The site in the E1B 55-kDa protein is conserved in serotypes 2, 4, 5, 7, 12, 40 and 41, but further investigations would be required to establish whether or not it is processed.

Finally, it is noted that no potential cleavage sites were found, in either the L4 100-kDa protein or the L3 23-kDa protein. This suggests that, contrary to the proposals of Morin and Boulanger (1986) and Chatterjee and Flint (1987), the 100-kDa and 23-kDa proteins are not processed by the virus encoded protease. It cannot be ruled out, however, that a sequence of the form M(L,I)XGX-X could be cleaved, but to date there is no evidence to suggest this.

Comparisons with substrate specificities of other viral proteases.

Table 1.7 displays the preferred substrate specificities of a selection of viral proteases including that of the Ad2 protease outlined above. As can be seen there are a number of similarities. In particular, the alphavirus nsP2 protease and the Ad2 protease both require a

	P_4	P_3	P_2	P_1	P_1'	P_2'
Rhino (2A)					G	
Polio (3C)	A			Q	G	P
CMPV (24K)	A		AP	Q	G	
TEV (NIa)		V	RK	Q	G	
TVMV (NIa)		Y		Q	G	
Sindbis (nsP2)			G	GA	GA	
YF (NS3)			RK	RK		
Ad2 (23K)	MLI		G		G	
Vaccinia (?)			A	G		
ASF (?)			G	G		

Table 1.7 Preferred substrate specificities of viral proteases
The table shows the preferred substrate specificities of a number of viral proteases in the P_4 to P_2' positions. In most cases the data is derived from comparisons of the natural substrates of the enzymes; but in the cases of the rhinovirus 2A, poliovirus 3C, Sindbis virus nsP2 and the adenovirus 23K proteases the results are derived from mutagenesis and/or peptide studies. The cleavage sites of the vaccinia and African swine fever viral proteins are included, although the identity of the proteases remains to be confirmed. The specificities of the HIV proteases are given in fig.1.5 and table 1.2 so are not included in this table. For references see introduction to this chapter.

glycine at P_2 and appear to have a preference for enzymes with a small side chain in the P_1' position. Also of interest are the cleavage sites in the other DNA virus proteins, where the P_1 and P_2 residues have been shown to be AG and GG in vaccinia and African swine fever virus proteins respectively. Furthermore a number of the viral proteases appear to have specific requirements at the P_2 and P_4 positions including the 3C proteases of the picorna- and related viruses. There has also been a suggestion that a number of proteases including the picornavirus 3C proteases cleave at β-turns (Hellen *et al.*,1989). In this regard, it may be of significance that all the predicted cleavage sites in the Ad2 proteins, except IVGL-G in pVI and MAGH-G in the 11-kDa protein, are predicted to be at β-turns (**table 1.5**). The question as to whether similar substrate specificities means that some of these enzymes from different viruses are structurally and mechanistically related will be considered in chapter 2.

Applications

The defined substrate specificities of selective proteases have been used to engineer appropriate cleavage sites into the interdomain regions of a fusion proteins (Sassenfeld,1990). In the future it is possible that the coding sequence for MSGGG could be included in such a system and assessed for cleavage by the Ad2 protease. This in itself would not be of great use unless a pure, abundant source of active enzyme was available. Attempts to clone and purify the Ad2 protease will be made in chapter 3.

Having defined a substrate specificity for the Ad2 protease we are now in a position to consider how best to design a specific inhibitor for the enzyme. As a preliminary step to doing this some of the peptides synthesised were tested as inhibitors. Particular attention was paid to the peptides that were not cleaved by the Ad2 protease. The results (**table 1.6**) showed that the only peptides that inhibited the enzyme to any significant extent were FGGAKK and SGGAFSW. This was surprising in view of the fact that the P_4 residue is missing in these peptides. It appears then that binding can take place in the absence of a hydrophobic residue at P_4, but that the latter may be required to orientate the peptide correctly, in the active site of the protease, to allow cleavage. Further evidence for this comes from the observation of a possible inhibitory effect by the acetylated peptide, Ac-GGA-ONH_2, on the Ad2 protease (Tremblay *et*

al., 1983).

If it is possible to design a peptide inhibitor for the Ad2 protease, the question arises as to whether it will be effective against the enzymes from other serotypes. A recent report shows that the Ad2 pVII precursor was cleaved by nuclear extracts from cells infected with serotypes 1,2,3,4,5,6,7 and 9, suggesting that the proteases from different serotypes do, in fact, have very similar, if not identical, substrate specificities (Houde and Weber,1990b). Further evidence for this comes from the observation that the 23-kDa putative protease is highly conserved in all the serotypes so far sequenced (Houde and Weber,1990a).

Finally in this regard, the similarities in substrate specificities of a number of viral proteases raises the interesting possibility that an inhibitor effective against a range of viral proteases could be designed. Before embarking on such studies a greater understanding of the mechanisms involved in these proteases and pure sources of active enzymes would be desirable. The work carried out in the remaining chapters endeavours to achieve these aims for the adenovirus protease, to some extent at least.

CHAPTER 2

Characterisation and Classification of the Ad2 Protease

INTRODUCTION

The aim of this chapter is to find out more about the mechanism and classification of the Ad2 protease. After describing the work that has already been done to this end, a brief definition of the four classes of proteases will be given. This will be followed by a summary of recent studies undertaken to classify other viral proteases.

CHARACTERISATION OF THE ADENOVIRUS PROTEASE

Bhatti and Weber (1978) developed an assay for the adenovirus protease using ^{35}S methionine labelled, Ad2ts1 infected cell extracts as substrates. In their assay, labelled precursor proteins act as substrates and the cleavage of pVII-VII can be readily monitored by SDS-PAGE and autoradiography. All attempts to characterise the enzyme to date have made use of this assay and as a result, some of its properties have been defined. The Ad2 protease has been shown to operate best at neutral pH, with no activity being detected below pH5 or above pH9.5. It is reported that the enzyme is nuclear localised and that it is packaged within the virion at a low copy number estimated at 5-10 copies per virion (Bhatti and Weber,1979a,b; Anderson,1990). Whilst the protease activity from purified virus was activated by 5mM dithiothreitol (DTT), the thiol reagent had no effect on the nuclear localised enzyme and indeed the latter was found to be inhibited by β-mercaptoethanol (Tremblay *et al.*,1983). This result serves to emphasise the need for a cleaner assay system, so that any observations made can be attributed to the enzyme and not to effects on complex substrates or additional components present.

Limited inhibitor studies revealed that phenylmethylsulphonylfluoride (PMSF), tosyl-lysine chloromethyl ketone (TLCK) and tosyl-phenylalanine chloromethyl ketone (TPCK) gave 64, 90 and 100% inhibition of pVII-VII processing when each was present at a

concentration of 1mM (Bhatti and Weber,1979a). Diisopropylfluorophosphate (DFP) was also reported to inhibit the enzyme, although the concentrations and degree of inhibition were not given (Tremblay *et al.*,1983). Based solely on these observations it has generally become accepted that the adenovirus codes for a chymotrypsin-like serine protease (Krausslich and Wimmer,1988; Hellen *et al.*,1989; Kay and Dunn,1990). Similarly, the mapping of the mutation in Ad2ts1 to the L3 open reading frame, coding for a 23-kDa protein, has led to the assumption that the latter is the protease (Yeh-Kai *et al.*,1983). Although this is likely to be the case, it cannot be ruled out that the 23-kDa protein plays an essential regulatory role and is not, itself, the protease.

CLASSIFICATION OF PROTEASES

Proteases have been divided up into 4 mechanistic classes according to the amino acid or ion usually associated with the catalytic process. These are termed *serine*, *cysteine*, *aspartic* and *metallo* proteases (Barrett,1977). In terms of the cleavage of the peptide bond the overall process is identical for all the proteases; but the various groups required to carry out the cleavages are different. In simple terms a nucleophile in the form of an oxygen or sulphur atom attacks the slightly electrophilic carbonyl carbon atom in the peptide bond. General base catalysis is required to remove the proton from the attacking nucleophile and an electrophilic atom is required close to the carbonyl oxygen to increase the polarisation of the C-O bond. After the formation of the intermediate, general acid catalysis takes place to facilitate the departure of the amine group (Dunn,1989).

The serine proteases include the well characterised trypsin, chymotrypsin and subtilisin and each has a catalytic triad (His, Asp, Ser), the residues of which make up the charge relay mechanism of the enzyme. The residues adjacent to the active centre amino acids are well conserved among serine proteases, as is the secondary structure of these evolutionarily related enzymes (Polgar,1987). In papain, the best characterised of the cysteine proteases, a thiolate-imidazolium ion pair is formed between Cys_{25} and His_{159} with the sulphur atom of the cysteine side chain acting as the nucleophile (Brocklehurst *et al.*,1987). In both serine and cysteine proteases an ester is formed between the nucleophilic oxygen or sulphur atom and the acyl portion of the peptide linkage. This acyl enzyme intermediate is stabilised in each case by

hydrogen bonding to 2 backbone -NH- groups.

The oxygen atom of a water molecule serves as a nucleophile in the aspartic and metallo proteases in which no covalent acyl enzyme intermediates are formed. In aspartic proteases the H_2O molecule is hydrogen bonded to the carboxyl groups of 2 highly conserved aspartic acid residues. In the metallo proteases the H_2O is coordinated to a divalent metal ion, usually Zn^{+2}, and the carboxyl group of a glutamic or aspartic acid residue. The latter serves as a general base to remove a proton and to assist the attack of the water molecule on the peptide carbonyl carbon atom (Auld and Vallee,1987).

The assignment of a protease to one of the mechanistic classes is not a trivial process and usually requires an examination of the enzyme's physical properties, a study of its susceptibility to a range of inhibitors, comparisons of its amino acid sequence with those of established proteases and site directed mutagenesis. For absolute proof of its mechanism active site labelling and X-ray crystallography are required. **Tables 2.1, 2.2** and **2.3** list some of the properties that may be used to distinguish the 4 classes. Kinetically significant pH optimum values of an enzyme can be used to predict the catalytically important residues (**table 2.1**). **Table 2.2** lists a set of diagnostic protease inhibitors and their effects against the various classes, whilst **table 2.3** includes a set of sequence motifs commonly associated with proteolytic enzymes.

Amino Acid	pKa
Aspartic Acid	3.9
Glutamic Acid	4.3
Histidine	6.0
Cysteine	8.3
Tyrosine	10.1
Lysine	10.8
Arginine	12.5
Serine	13.0
Threonine	13.0

Table 2.1 pK Values of Ionisable Amino Acid Side Chains
The values given are for the individual amino acids and will vary under the influence of their immediate environment within the protein structure.

Inhibitor	Protease Class			
	Serine	Cysteine	Aspartic	Metallo
DFP	Inhibited	-	-	-
PMSF	Inhibited	(Inhibited)	-	-
DCI	Inhibited	-	-	-
SBTI	Inhibited	-	-	-
PCMB	-	Inhibited	-	-
Cystatin	-	Inhibited	-	-
E-64	-	Inhibited	-	-
DTT	-	Activated	-	Inhibited
Pepstatin	-	-	Inhibited	-
EDTA	-	-	-	Inhibited
o-Phenanthroline	-	-	-	Inhibited

Table 2.2 Diagnostic Protease Inhibitors
A list of the more common inhibitors which can be used to aid the classification of proteases. The inhibition of cysteine proteases by PMSF is reversible by DTT. A dash (-) indicates no effect.

Protease	Motif	Predictive Success	False Positives
Trypsin-like Serine Proteases	G D S G G	87%	4
	L S V S A T H C I A M G	92%	2
Subtilisin-like Serine Proteases	T G T S X S X P X X S A A A V	100%	2
Cysteine Proteases	S Q X X X G X C W X X T E A G	100%	1
Eukaryotic Aspartic Proteases	L S S I D T G T T V A A F N A	100%	21

Table 2.3 Active Site Sequence Motifs
The table shows some of the characteristic motifs encompassing active site residues of proteases, taken from the Wisconsin package. The active site residue is underlined in each case, X denotes no preference and the vertical columns indicate alternatives at that position. Also shown are the % of known proteases of the appropriate class detected by the motif, and the number of false positives from approximately 30 000 sequences in the database.

The use of these techniques in the characterisation of viral proteases has proved successful in a number of cases, but also serves to demonstrate the problems involved with classification.

VIRAL PROTEASES

So far the structures and mechanisms of 3 viral proteases, those from HIV-1, RSV and Sindbis virus, have been elucidated by X-ray crystallography. The classification of other viral proteases has been achieved mainly by sequence alignments and site directed mutagenesis. Such studies have shown that the vast majority of viral proteases appear to be either trypsin-like or aspartic proteases. The only major exceptions at present appear to be the nsP2 proteases of the alphaviruses, which have recently been proposed to be papain-like cysteine proteases (Gorbalenya *et al.*,1991).

Trypsin-like Proteases

A list of viral proteases that are currently thought to be related to trypsin is given in **table 2.4.** As can be seen these include the 2A and 3C proteases of the picorna- and related viruses, the NS3 protease of the flaviviruses and the capsid protein of the sindbis virus. A surprising feature of the list is that many of the enzymes are categorised as trypsin-like *cysteine* proteases. To the classical enzymologist this is perhaps a contradiction in terms, but an increasing body of evidence supports the existence of such a subclass.

It is suggested that this group of cysteine proteases is structurally and mechanistically related to the cellular serine proteases and not to papain. In other words they contain a His, Asp, Cys catalytic triad in place of the cysteine/histidine thiolate imidazolium ion pairing of papain. Computer based alignments of primary and secondary structures have revealed that the 2A and 3C proteases are homologous to the small and large structural subclasses of trypsin-like enzymes respectively (Gorbalenya *et al.*,1986; Brenner,1988; Bazan and Fletterick,1988). Subsequent alignments have predicted that the flavivirus NS3 and pestivirus p80 non-structural proteins are also trypsin-like proteases. Unlike the picornaviral enzymes, however, it is proposed that a serine is the active site residue (Bazan and Fletterick,1989; Gorbalenya *et al.*,1989). Significantly, recent work has shown that both NS3 and p80 are proteases and

mutagenesis studies support the proposed active site residues (Chambers *et al.*,1991; Wiskerchen and Collet,1991). **Table 2.4** includes the predicted active site histidine, cysteine and serine residues of a number of viral proteases aligned with the corresponding sequences from cellular enzymes. A number of site directed mutagenesis studies have been carried out in

		*	*	*
Cellular	Trypsin	AAHCG	NDI	GDSGG
Cellular	Chymotrypsin	AAHCY	NDI	GDSGG
Entero-	Polio (type 1) 3C	PTHAS	RDI	GQCGG
Rhino-	Rhino (type 14) 3C	PTHAQ	RDI	GQCGG
Cardio-	EMCV 3C	NRHMA	RDI	GWCGS
Aphtho-	FMDV 3C	PRHLF	RDI	GYCGG
Como-	CPMV 24K	CKHFF	WDL	EDCGS
Poty-	TEV NIa	NKHLF	RDM	GQCGS
	TVMV NIa	NQHLF	RDI	GQCGS
Flavi-	YFV NS3	MWHVT	EDL	GTSGS
Pesti-	BVDV p80	VDHVT	TDS	GWSGL
Cellular	α-lytic protease	AGHCG	NDR	GDSGG
Alpha-	Sindbis capsid	PLHVK	YDM	GDSGR
Entero-	Polio (type 1) 2A	NYHLA	RDL	GDCGG
Rhino-	Rhino (type 1) 2A	MYHLM	RDL	GDCGG

Table 2.4 Proposed active site residues of trypsin-like viral proteases
A subset of the viral proteases that have been classified as being trypsin-like serine or cysteine proteases taken from Bazan and Fletterick (1989). The virus family is indicated in bold on the left and the identity of the virus or cellular protease is given in column 2. The amino acids surrounding the predicted active site histidine, aspartate and serine/cysteine residues are listed for each protease. In most cases the active site residues have been predicted by sequence alignments with trypsin/ chymotrypsin or the smaller bacterial α-lytic proteases (Bazan and Fletterick,1989; Gorbalenya *et al.*, 1989). The active site residues in the capsid protein of the alphavirus, Sinbis, however, have been confirmed by X-ray crystallography. For references see text.

an attempt to confirm the identity of these active site residues. Although most are able to do so, few are extensive enough to be absolutely conclusive (Sommergrubber *et al.*,1989; Cheah *et al.*,1990; Kean *et al.*,1991; Yu and Lloyd,1991). One interesting observation is that the replacement of the predicted active site cysteine of the polio 3C protease with a serine did not completely abolish protease activity. This taken in conjunction with the observation that a serine to cysteine mutation in trypsin does not remove all its protease activity suggests that serine and cysteine may be interchangeable, to some extent at least, in the trypsin-like proteases (Lawson and Semler,1991).

Extensive inhibitor studies have been carried out on the rhinovirus 3C protease by Orr *et al.* (1989) revealing that its inhibitor profile fits better into the cysteine than the serine class. It is not, however, inhibited by the peptide based inhibitors, E64 and cystatin, regarded as

diagnostic of cysteine proteases. Again these observations are in accord with the existence of a subgroup of proteases that link the cysteine and serine classes.

Finally consider the best studied of the viral proteases belonging to this group, the p26 capsid protein of the Sindbis virus, which has been crystallised. p26 is responsible for a single autocatalytic cleavage that removes it from the other structural proteins and is rather more conventional than some of its viral counterparts in that it is a trypsin-like protease with an active site serine. Sequence alignments and mutagenesis studies suggested that His_{141}, Asp_{163} and Ser_{215} make up the catalytic triad and this was confirmed by the elucidation of its crystal structure (Hahn *et al.*,1985; Hahn and Strauss, 1990; Choi *et al.*,1991). **Fig.2.1** shows the structure of p26 and the positions of the active site residues. Also note that after autocatalytic processing, the C-terminus of the protein remains in the active site cleft where it is proposed to inhibit the enzyme (Choi *et al.*,1991).

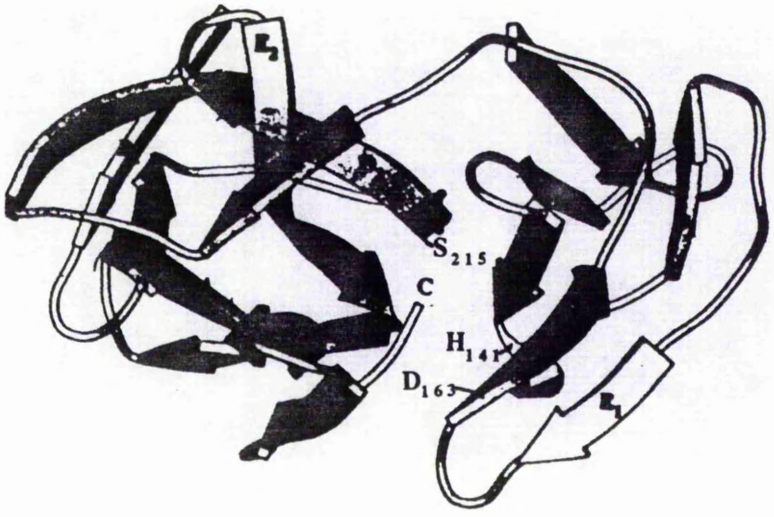


Fig. 2.1 Structure of the Sindbis Virus p26 Capsid Protein
The alpha carbon trace of the Sinbis virus p26 protease is shown above. The location of the active site residues are marked along with the C-terminus of the protein. Note the position of the C-terminus in the active site cleft (from Choi *et al.*, 1991).

Viral Aspartic Proteases

At present this group of enzymes is made up entirely of the retroviral proteases, although putative aspartic protease sequences have been found in vaccinia, hepadna- and

parapoxvirus proteins (Slabaugh and Roseman,1989; Nassal *et al.*,1989; Mercer *et al.*,1989). As yet, however, none of these proteins have been shown to have protease activity so this discussion will concentrate on the classification of the retroviral proteases.

Initial reports showed that the AMV, p15 protease was inhibited by p-chloromercuri-benzoate (PCMB) and N-ethylmaleimide (NEM), suggesting that it was a cysteine protease, whilst the murine retrovirus proteases were inhibited by PMSF which is specific for serine proteases. When the amino acid sequences of a number of retroviral proteases had been determined, however, it was realised that the preliminary classifications were incorrect (reviewed in Krausslich and Wimmer,1988). All the enzymes contained a highly conserved sequence similar to the sequence around the active site aspartic acid residue of cellular aspartic proteases (**table 2.3**). They were also inhibited by pepstatin A, although at higher concentrations than their cellular counterparts (Katoh *et al.*,1987). The fact that the retroviral enzymes contain only one conserved active site aspartate and are about half the size of other aspartic proteases has led to the suggestion that they are active as dimers (Pearl and Taylor,1987). Elucidation of the crystal structures of RSV and HIV-1 confirmed this revealing that 2 polypeptides combine in an asymmetric manner, each providing one of the necessary active site aspartic acid residues as is shown in **fig.2.2** (Leis *et al.*, 1989; Navia *et al.*,1989). It has been proposed that the cellular aspartic proteases, which contain 2 homologous domains, may have evolved through gene duplication and fusion from a

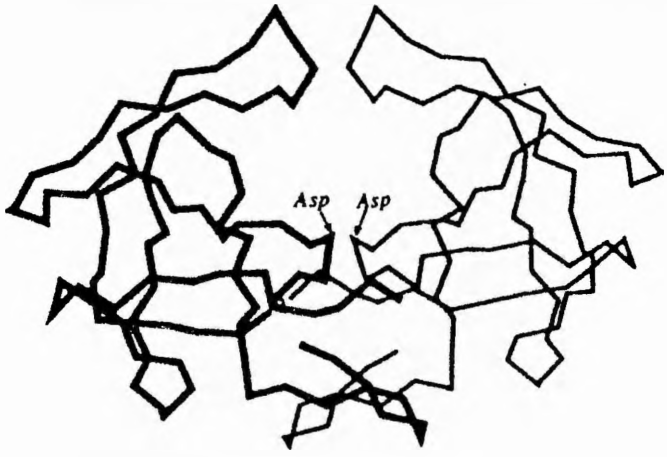


Fig. 2.2 Structure of the HIV-1 Protease
Alpha carbon trace of the HIV-1 protease taken from Navia *et al.* (1989). The trace of one of the monomers is shown in bold and the positions of the active site aspartate residues are marked, with one being donated by each subunit. Note also the position of the flaps over the active site.

primordial enzyme to which the retroviral proteases are equivalent. Krausslich (1991) constructed a plasmid designed to produce fused dimers of the HIV-1 protease. He went on to show that when substituted for the wild type monomer in *gag-pol,* the dimeric protease was capable of efficiently cleaving the polyprotein. The presence of the dimer, however, was found to give premature processing of the polyproteins in cells which prevented assembly of the virus (Krausslich,1991). This emphasises further the crucial role that dimerisation plays in the control of retrovirus protease activity and infers that the viral enzymes are more than simply backward ancestors of their cellular counterparts (Burstein *et al.*,1991).

A more detailed understanding of the mechanism of the HIV-1 protease has been obtained by the complete mutagenesis of the enzyme and also by the determination of kinetically significant pK_a values (Loeb *et al.*,1989; Hyland *et al.*,1991).

Other Viral Proteases

At present the nsP2 proteins of alphaviruses are among the few well characterised viral proteases that have not been assigned to the trypsin-like or aspartic groups. The protease activity is located at the C-terminus of nsP2, but the sequence contains no obvious protease motifs. Site directed mutagenesis studies show that Cys_{481} and His_{558} of the Sindbis virus nsP2 protease are required for activity and the current proposal is that the enzyme is a member of the papain-like cysteine protease class (Strauss and Strauss,1990). This proposal has been supported by Gorbalenya *et al.* (1991) who have carried out sequence alignments of the alphavirus nsP2 proteins with the papain-like proteases. The sequences, however, differ quite considerably from papain and more studies are required to establish with certainty the mechanism by which nsP2 operates.

At present, like nsP2, the Ad2 protease must be included in the fifth group of proteases of unknown mechanism since its classification as a chymotrypsin-like cellular protease is based on limited studies. The first aim of this section is to design a cleaner assay system and to use it to carry out more extensive characterisation and inhibitor studies. Given the lack of homology between the Ad2 L3 23-kDa protein and other cellular/viral proteases it is also important to establish with certainty that the 23-kDa protein is the protease (Krausslich and Wimmer,1988). This will be achieved by purifying active protease from the virus.

METHODS

1 ASSAYS FOR THE ADENOVIRUS PROTEASE.

1.1 ^{35}S- Methionine Labelled Ad2ts1 Assay.

This method was developed by Bhatti and Weber (1978) and is essentially as described by Anderson (1990). HeLa cells were grown to 90% confluence in 75cm^3 plastic petri dishes (Celcult, Sterilin) and infected with 200µl of a 10^9 p.f.u./ml stock of the temperature sensitive mutant, Ad2ts1, in 2ml G-MEM. Plates were rocked for 1 hour at 37°C before the virus was removed, 15ml of G-MEM/2% calf serum were added and the infection was allowed to proceed for 24 hours at 39°C. The medium was removed, cells were washed with 5ml of methionine free MEM (Gibco-BRL) and labelled for 4 hours at 39°C with 200µCi of ^{35}S methionine (Amersham) in 3ml of methionine free medium/ 2% calf serum. Labelled cultures were harvested, washed twice in PBS, resuspended in 1ml of 10mM sodium phosphate buffer, pH7.4 and sonicated on ice for 4x15 seconds. The suspension was then centrifuged in a microcentrifuge at high speed (10000g) for 10 minutes and the pellet was resuspended in 1ml of phosphate buffer before being aliquoted and stored at -70°C.

When required, an aliquot of the labelled extract was thawed at 37°C, centrifuged at 10000g and the pellet resuspended in 0.5 volumes of phosphate buffer. Assays were performed by incubating 10µl of ^{35}S labelled substrate with 5µl of protease source and 5µl of the desired buffer for 1 to 18 hours at 37°C. The protease activity was analysed by assessing the cleavage of pVII to VII using SDS-PAGE and autoradiography (see method 2). Where quantitative information was required gels were scanned on a Vitatron TLD densitometer and peak areas were determined.

1.2 ^{35}S Methionine Labelled Ad2 Assay

A slightly modified version of the previous assay, involving the labelling of wild type Ad2 for 2 hours late in infection, was used in selected experiments. HeLa cells were infected with wild type Ad2 as described for Ad2ts1 in method 1.1, with the exception that the infection was for 48 hours at 37°C. Infected cells were labelled for 2 hours and the extracts were

prepared as in 1.1. For some assays it was necessary to remove the endogenous protease activity and this was done by heating the samples to 60°C for 2 minutes prior to freezing them at -70°C. When heat treated extracts were used, assays were carried out as described for the Ad2ts1 substrate. For other assays, the endogenous protease was utilised and in this case no heat treatment was included and 10µl of ^{35}S substrate and the desired buffer were mixed and incubated at 37°C.

1.3 Synthetic Peptide Assay for the Adenovirus Protease.

2.5µg of MSGGAFSW, in 5µl of H_2O, were incubated with 40µl of buffer (50mM Tris/HCl, 1M NaCl, 5mM EDTA, 2mM DTT unless stated) and 5µl of protease source for 30 minutes to 12 hours at 37°C. The reaction was stopped by the addition of 100µl of 0.1% TFA. Samples were filtered through a 0.2µ filter (Anachem) and 75µl of each was analysed by reverse phase chromatography as described in chapter 1, method 4.2, using the gradient:

time (minutes)	%B
0	0
5	0
10	20
25	30
30	40
30.1	0
40	0

The full scale deflection was set at 0.2 absorbance units. Quantitative results were obtained by measuring the peak areas using a Shimadzu C-R1B integrator.

Thin layer electrophoresis (TLE) was used to assess the cleavage of MSGGAFSW by the Ad2 protease in some experiments and was carried out as described by Kotler *et al.* (1988) with the exception that the buffer used was 1.5% (v/v) formic acid pH1.9. Plates were developed with fluorescamine (Kotler *et. al.*,1988) and photographed under u.v. light.

1.4 Inhibitor Studies.

Leupeptin, pepstatin, dithiodipyridine (DTDP), soya bean trypsin inhibitor (SBTI), 3,4-dichloroisocoumarin (DCI), 1,10 phenanthroline, cystatin, E64, TPCK, TLCK, PMSF, PCMB and DTT were obtained from Sigma; iodoacetic acid (IAA) and N-ethylmaleimide (NEM) were from BDH. PCMB was dissolved at 20mM in 0.01M NaOH and diluted to 2 or 0.2mM with 50mM bicarbonate buffer pH 9.5. Iodoacetate was freshly prepared in 0.05M

bicarbonate buffer pH 9.5. NEM, dithiodipyridine, E64, cystatin, SBTI and TLCK were prepared in 25mM phosphate pH 7.0. Dithiodipyridine was also prepared in 10mM Tris/maleate pH 5. Leupeptin was dissolved in 0.05M ammonium bicarbonate pH7.5 and TPCK was dissolved in methanol at a concentration of 10mM. Stocks of 10mM DCI (in DMF), 1mM pepstatin and 200mM 1,10 phenanthroline (both in methanol) were prepared. Inhibitors were incubated with 5µl of disrupted purified Ad2 for 10 minutes at 37°C. The final concentration of inhibitor in the preincubation was 1mM with the exception of PCMB which was also tested at 0.1mM, PMSF which was added as a solid to a final concentration of 5mM, 1,10 phenanthroline (10mM), DCI (0.1mM), E64 (10µM), leupeptin (2µM), cystatin (2µM), pepstatin (1µM) and SBTI (0.1mg/ml). Control incubations with the corresponding amount of the appropriate buffer, methanol or DMF were also prepared. Following the preincubation with the inhibitor, the protease was added to 10µl of heat treated ^{35}S labelled Ad2 extract or 2.5µg of MSGGAFSW. The ^{35}S and peptide assays were carried out as described previously with incubation times of 6 and 1 hours respectively. No DTT or NaCl was included in the peptide assay buffer for these studies.

2 SDS POLYACRYLAMIDE GEL ELECTROPHORESIS.

2.1 Preparation of Gels.

The protocol of See and Jackowski (1989) was employed to analyse the proteins present in various samples. Bio-Rad Mini Protean System gel rigs were assembled and used according to the manufacturers instructions. 7.5-15% running gels were prepared by mixing 30% acrylamide, 2% bis-acrylamide, 1M Tris/HCl pH8.7, 10% SDS and H_2O according to **table 2.5** and degassing for 15 minutes. 10µl of TEMED and 50µl of freshly prepared 10% ammonium persulphate were added to the running gel mixture, which was immediately poured into the gel cassettes leaving a 2cm gap at the top for the stacking gel. The running gel was overlayed with water saturated butanol, which was rinsed off with H_2O and 70% ethanol after the gel had set. The stacking gel was prepared by mixing 6.35ml H_2O, 1.7ml 30% acrylamide, 0.7ml 2% bis-acrylamide, 1.25ml 1M Tris/HCl pH 6.9, 100µl 10% SDS, 50µl 10% ammonium persulphate and 20µl TEMED and pouring on top of the running gel after

insertion of the comb.

	% Acrylamide						
	5	7.5	10	12.5	15	17.5	20
30% acrylamide (ml)	2.5	3.75	5	6.25	7.5	8.75	10
2% bis-acrylamide (ml)	1.95	1.45	1	0.75	0.65	0.55	0.5
1M Tris/HCl pH8.7 (ml)	5.6	5.6	5.6	5.6	5.6	5.6	5.6
10% SDS (µl)	75	75	75	75	75	75	75
H_2O (ml)	4.75	4.12	3.32	2.32	1.2	-	-

Table 2.5 Recipe for SDS-PAGE Running Gels

2.2 Sample Preparation and Running of Gels

Protein samples, typically 5-20µl (1µg protein) were mixed with 1/4 volume of reducing buffer (0.1% bromophenol blue, 6.25% SDS, 0.15 M Tris/HCl pH6.8, 25% glycerol, 12.5% mercaptoethanol) and heated to 80°C for 5 minutes. Reduced samples were loaded into the wells using a Hamilton syringe and proteins were separated by applying a constant current of 20mA to each gel for 1 hour (maximum voltage 180V) using 0.19M Glycine, 0.025M Tris, 0.1% SDS as the electrophoresis buffer.

2.3 Coomassie Blue Staining of Gels

SDS polyacrylamide gels were stained for 30 minutes in 0.2% Coomassie brilliant blue R250 (BDH), 50% methanol, 10% acetic acid and destained in 25% methanol, 10% acetic acid. The gels were agitated gently on an orbital shaker during the staining and destaining steps, after which they were dried onto 3MM Whatman filter paper using a Biorad 583 vacuum gel drier, heated to 80°C, or stored in sealed pastic bags with the air removed.

2.4 Silver Staining of Proteins

This was carried out essentially according to Rabilloud *et al.* (1988) whose modified method reduces the background of gels by the inclusion of sodium dithionite. Gels were fixed overnight in 25% methanol, 10% acetic acid or Coomassie blue stained/destained as in 2.3. They were then washed for 2x15 minutes in ddH_2O before being sensitised for 1 minute in 0.025% sodium dithionite and being given 2 further 1 minute washes with ddH_2O. Staining was for 30 minutes in 0.2% silver nitrate, 1mM formaldehyde and was followed by a 1 minute

H_2O wash. Colour development was carried out in 6% sodium carbonate, 6mM formaldehyde, 20μM sodium thiosulphate for 2x 5 minutes or until bands appeared, after which the reaction was stopped by the addition of acetic acid to 3.5% and leaving for 10 minutes. This was followed by 4x 10 minute washes with water. The gels were stored in 20% ethanol at 4°C prior to being dried or photographed.

2.5 Autoradiography

Gels on which radio-labelled proteins had been separated were stained and dried as described in method 2.3. In the case of ^{35}S labelled proteins, gels were washed for 2x 15 minutes in Amplify (Amersham) prior to drying. Dried gels were exposed to Fuji film or Amersham Hypersensitive film, adjacent to an intensifying screen. After various exposure times the films were developed using an automatic developer (Kodak).

2.6 Western Blotting

Protein samples were separated using the Bio-Rad Mini-Protean Gel system according to methods 2.1 and 2.2 above. When the dye front was 0.5cm from the bottom of the gel the power supply was switched off and it was removed from its cassette for blotting in a semi-dry blotter (LKB). Six pieces of 3MM Whatman filter paper were cut to the same size as the gel, soaked in anode buffer 1 (0.3M Tris, 20% methanol pH10.4) and placed on the bottom graphite plate. 3 further pieces of filter paper were cut and soaked in anode buffer 2 (0.025M Tris, 20% methanol, pH10.4) and placed on top of the first 6. A piece of nitrocellulose paper, also cut to the same size as the gel, was soaked in ddH_2O and placed on the pile of filter paper. Air bubbles were removed by rolling with a glass pipette. The gel was rinsed with ddH_2O and placed on top of the nitrocellulose. The sandwich was completed with 3 pieces of filter paper soaked in cathode buffer (0.025M Tris, 40mM 6-aminohexanoic acid, 20% methanol pH9.4). Both graphite plates were given a liberal sprinkling of ddH_2O before the upper plate was put in place and the electrodes attached. The gel was blotted for 45 minutes at a constant current of 50mA.

Blots were developed using a Bio-Rad Kit (alkaline phosphatase conjugated anti-rabbit Ig) according to the manufacturers instructions. The nitrocellulose was blocked for 1 hour in

3% gelatin in 20mM Tris/HCl, 0.5M NaCl, pH7.5 (TBS) and agitated on an orbital shaker at room temperature. It was then washed for 5 minutes in TBS, 0.05% Tween 20 (TTBS) before being incubated with the first antiserum for 1 hour on the shaker. The first antisera were made up in 1% gelatin, TTBS at dilutions ranging from 1:500 to 1:2000 depending on their titre. The filter was then washed for 2 x 5 minutes in TTBS prior to a 1 hour incubation with a 1:3000 dilution of the alkaline phosphatase conjugated anti-rabbit Ig in 1%gelatin, TTBS. The filter was given 3 final washes: 2 x 5 minutes in TTBS and 1 x 5 minutes in TBS before being placed in 10ml of 20mM Tris/HCl, pH9.6 containing 100µl each of colour development reagents A and B. When bands appeared (10-15 minutes) the nitrocellulose was washed for 2 x 5 minutes in ddH_2O and dried.

3 PREPARATION OF ANTI-PEPTIDE SERA

The following four peptides derived from the Ad2 L3 23-kDa protein sequence were synthesised as described in chapter 1, method 1:

(1) **M G S S E Q E L K A I V K V L G C** (N-terminal peptide)

(2) **C R E T G G V H W M A F A W N**

(3) **C Q V Q P T L R R N Q E Q L Y**

(4) **C Q I R S A T S F C H L K N M** (C-terminal peptide)

3.1 Coupling of Peptides to Carrier Using SPDP.

The peptides were coupled to human serum albumin (HSA) using N-succinimidyl-3-(2-pyridyl dithio)-propionate (SPDP) which couples the amino groups of the carrier to the cysteinyl residues of the peptides. The method used is essentially as described by Scheidtmann (1989). 10mg of HSA were dissolved in 1ml of 0.1M sodium phosphate, 0.1M NaCl, pH7.5 and then 225µl of 20mM SPDP in ethanol were added dropwise over 5 minutes. The solution was stirred for a further 30 minutes at room temperature before the free SPDP was removed by FPLC using the Fast Desalting HR10/10 column. The column was equilibrated with the above phosphate buffer, and 500µl of sample were loaded per run. The flow through was monitored at a wavelength of 280nm, the full scale deflection was 2 absorbance units, the flow rate 1ml/minute and the fraction size, 1.25ml. Fractions containing activated HSA were pooled and mixed with 4mg of peptide (in 1ml phosphate buffer) for 2 hours. The conjugated peptide was

then dialysed against PBS at 4°C overnight, aliquoted and stored at -20°C prior to immunisation of rabbits.

3.2 Immunisation Protocol.

The peptide conjugates were mixed with an equal volume of Freund's complete adjuvant to give a concentration of 0.5mg/ml and 1ml was injected subcutaneously into Dutch rabbits (2 rabbits were injected for each of the 4 conjugates). The rabbits were then given 1ml booster injections, at 2 week intervals, of peptide conjugates mixed with an equal volume of Freund's incomplete adjuvant. Again the injections were subcutaneous. 5ml of blood were withdrawn from the rabbits' ears before the initial immunisation (pre-immune sera) and 7 days after each subsequent immunisation. Blood was collected into plastic tubes and allowed to clot by incubating it for 1 hour at 37°C and then storing at 4°C overnight. The sera were collected by centrifugation for 10 minutes at 2000g and then stored in 0.5ml fractions at -70°C until required.

3.3 Assessment of Antisera.

In the first instance the titres of the antisera were determined using an ELISA procedure and then qualitative assessment was made by Western Blotting (method 2.6).

ELISA for Antipeptide Sera.

100µl of 1% glutaraldehyde in PBS were pipetted into each well of a 96 well plate (Immulon-2, Dynatech) and left overnight at 4°C. Wells were washed with 200µl of PBS before 50µl of a 10µg/ml solution of peptide (or carrier) in PBS were added to the appropriate wells and left overnight at 4°C. The wells were emptied and the plates dried for 30 minutes at 37°C. Non-specific binding was blocked by adding 150µl of 10% dried milk (Marvel) in PBS to each well and incubating for 3 hours at 37°C before washing with 2x200µl of 1% Marvel/PBS. A doubling dilution series of each antiserum in 1% Marvel/PBS was prepared in a separate plate and then 50µl from each well were transferred to the peptide plate which was incubated at room temperature for 1 hour. Wells were washed with 3x150µl of 1% Marvel/PBS before 50µl of a 1:500 dilution of horseradish peroxidase linked anti-rabbit Ig (Scottish Antibody Production Unit) in 1% Marvel /PBS were added to each well and incubated for a further hour at room temperature. Wells were washed with 2x200µl of 1%

Marvel/PBS followed by 2x200µl of 0.025M citric acid, 0.05M disodium hydrogen phosphate (phosphate/citrate buffer). 100µl of freshly prepared peroxidase substrate (40mg o-phenylene diamine in 50ml of phosphate/citrate buffer with 50µl of hydrogen peroxide) were added to each well and the colour allowed to develop (10-30 minutes). Plates were read in a multi-well plate reader with a 492nm filter.

4 PURIFICATION OF ADENOVIRUS PROTEASE

4.1 Extraction of Protease Activity From Ad2

Three 200µl samples of disrupted, purified Ad2, prepared as described in chapter 1 methods 3 and 4, were centrifuged at 60000rpm (100000g) in a Beckman L8-55M ultracentrifuge for 15 minutes at 4°C. The pellets were resuspended in 200µl of (A) 50mM Tris/HCl, 5mM EDTA, pH7.5, (B) buffer A made 2M with respect to NaCl and (C) buffer A made 2M with respect to urea. The samples were vortexed and extracted on ice for 30 minutes. 50µl of each sample were stored on ice for analysis, whilst the remainder of the samples were centrifuged at 100000g as before. Supernatants were removed for analysis whilst the pellet from the urea extraction (Buffer C) was resuspended in 150µl of buffer A made 2M with respect to NaCl and urea. The sample was vortexed and left on ice for 15 minutes and then centrifuged at 100000g. The resulting NaCl/urea supernatant was assessed for protease activity and analysed by SDS-PAGE along with the other supernatants.

4.2 Purification of Protease from Ad2

Ad2 was purified from 6 litres of infected HeLa cells and disrupted by dialysis against 10mM Tris/maleate pH6.4 as described in chapter 1, methods 3 and 4. The virus was mixed with an equal volume of 50mM Tris/HCl, 5mM EDTA, 4M urea, pH7.5, extracted on ice for 10 minutes and centrifuged for 5 minutes at high speed in a microcentrifuge. The resulting pellet was resuspended in 500µl of 50mM Tris/HCl, 5mM EDTA, 2M urea, 2M NaCl, pH7.5, vortexed, extracted on ice for 30 minutes and then centrifuged at 100000g as described in 4.1. The supernatant was fractionated by gel filtration using an FPLC Superdex 75 HR10/30 column equilibrated with 50mM Tris/HCl, 5mM EDTA, 1M NaCl, pH8. The flow rate was 0.5ml/minute, 1ml fractions were collected, the full scale deflection was 0.2 absorbance units and the detection wavelength 280nm.

4.3 Desalting Purified Ad2 Protease

Disrupted purified virus from 2 litres of Ad2 infected cells was mixed with an equal volume of 50mM Tris/HCl, 5mM EDTA, 4M urea, pH7.5 and centrifuged for 5 minutes at high speed in a microcentrifuge. The pellet was resuspended in 500µl of 50mM Tris/HCl, 5mM EDTA, 2M NaCl, 2M urea, pH7.5 and extracted as described in 4.2. Half of the resulting supernatant was applied to an FPLC desalting column HR10/10 equilibrated with 50mM Tris/HCl, 5mM EDTA, pH8. The remainder of the sample was loaded onto the desalting column equilibrated with the same buffer made 1M with respect to NaCl. The flow rate in each case was 1ml/minute, 1.25ml fractions were collected, the full scale deflection was 0.5 absorbance units and the detection wavelength was 280nm.

Four 10µl samples of the main protein fraction from each separation were tested for protease activity using the peptide assay (method 1.3). Two assays were set up using the standard assay buffer and 2 with the same buffer without the 1M NaCl. The salt buffer was added to the samples immediately after their elution from the column. Incubation times were for 1 hour at 37°C.

RESULTS

1 Assay Systems

Two assay systems were developed for the Ad2 protease. The first was based on the procedure described by Bhatti and Weber (1978) and utilised either wild type Ad2 or Ad2ts1 proteins labelled with ^{35}S methionine. Wild type Ad2 was labelled for 2 hours late in infection and 10µl of the resulting cell extract were incubated with an equal volume of 50mM Tris/HCl pH8 for 3 hours at 37°C. Analysis by SDS-PAGE and autoradiography revealed that there was more than 50% digestion of pVII to VII in the incubated sample (**fig.2.3a**, lane 2). When the extract was heated to 60°C prior to the incubation a strong band was present corresponding to pVII, but no VII was visible revealing that endogenous protease activity had been destroyed (**fig.2.3a**, lane 3). Incubation of 10µl of the heat treated substrate with 5µl of purified disrupted Ad2 and 5µl of 50mM Tris/HCl pH8 gave complete digestion of pVII to VII (**fig.2.3a**, lane 4). Thus a substrate is available which can either make use of endogenous protease or be heated to 60°C and used to assess the protease activity from external sources. Throughout this project 3 variations of the ^{35}S methionine assay were used with the substrate being (a) ^{35}S Ad2, (b) heat treated ^{35}S Ad2, or (c) ^{35}S Ad2ts1.

The second assay system developed made use of the octapeptide MSGGAFSW, derived from the cleavage site of pVI (see chapter 1). Analysis by HPLC revealed that MSGGAFSW elutes as a single peak at 29% acetonitrile using the gradient described. Incubation of the peptide with disrupted virus for 1 hour at 37°C prior to HPLC produced 2 additional peaks (**fig.2.3b**). The peaks, labelled 2 and 3, were collected and subjected to amino acid analysis. Peak 3 was shown to contain methionine, serine and glycine in the ratio of 1:1:2, while analysis of peak 2 revealed equimolar amounts of alanine, phenylalanine and serine. Under the hydrolysis conditions used, tryptophan was destroyed, but its presence in peak 2 was evident from the higher absorbance of this peak at 226nm. Thus the amino acid analysis confirmed that the adenovirus protease cleaves MSGGAFSW at the G-A bond giving MSGG (peak 3) and AFSW (peak 2). The use of the peptide substrate, therefore, provides a simple assay system from which quantitative results can be readily obtained either by

comparing the areas under the MSGGAFSW and AFSW peaks or by quantitative amino acid analysis.

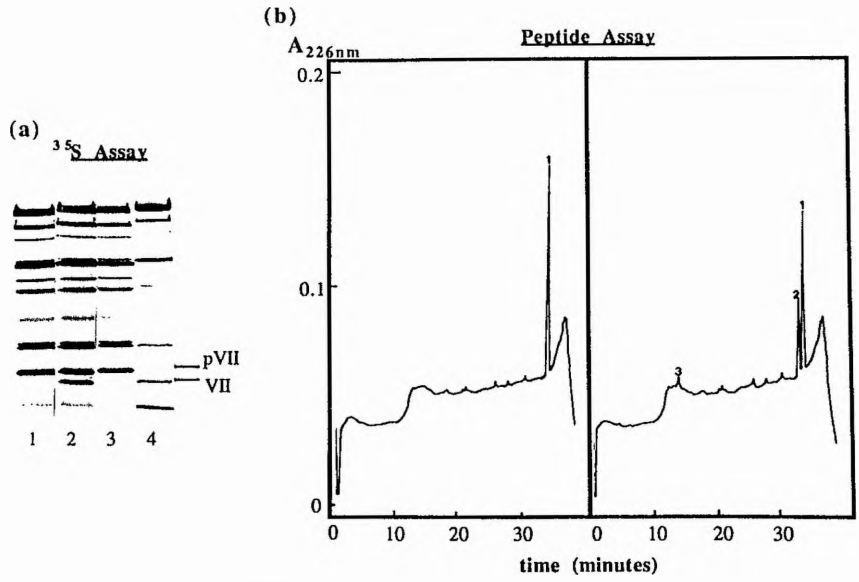


Fig. 2.3 Assays for the Ad2 Protease
(a) *^{35}S methionine assay.* Autoradiograms of reduced samples separated by SDS-PAGE. 10µl of ^{35}S methionine labelled Ad2 extracts were loaded onto the gel as follows: *lane 1*, unincubated; *lane 2*, incubated; *lane 3*, heat-treated and incubated; lane 4, heat-treated and incubated with 5µl of disrupted, purified Ad2 (for details see method 1.2). Incubations were for 3 hours at 37°C.
(b) *Peptide assay.* HPLC separations of 1.25µg of MSGGAFSW. *Left profile*, peptide alone; *right profile*, peptide after incubation for 1 hour at 37°C with 5µl of disrupted, purified Ad2 (for details see method 1.3). Amino acid analysis was used to identify peaks 1, 2 and 3 as MSGGAFSW, AFSW and MSGG respectively.

Digestion of the peptide MSGGAFSW was also monitored by TLE in an attempt to develop a rapid screening method (**fig.2.4**). The peptide was incubated at 37°C for 1, 2, 5, and 6 hours with disrupted, purified Ad2 and the resulting samples were analysed by TLE. Two spots were clearly seen to increase in intensity as the digestion proceeded and these presumably correspond to the tetrapeptides identified following HPLC separation of similar peptide digests (**fig.2.3b**).

The ^{35}S and peptide assays were used in parallel as far as possible to investigate further the physical characteristics of the Ad2 protease and to test the effects of a range of inhibitors on its activity. The experiments were carried out using disrupted, purified Ad2 as a source of protease unless otherwise stated. The absence of a pure source of protease in these studies

means that no attempts have been made to determine kinetic constants.

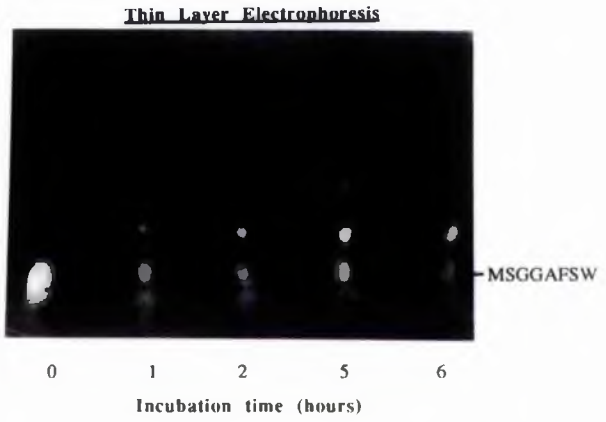


Fig. 2.4 TLE Analysis of a Timed Digest of MSGGAFSW
20µl of a 1mg/ml sample of MSGGAFSW were incubated with 5µl of purified, disrupted Ad2 at 37°C. Samples (2µl) were analysed by TLE after 0, 1, 2, 5 and 6 hours as indicated.

2 Effects of varying pH and ionic strength on Ad2 activity

The ^{35}S and peptide assays were used to determine the pH optimum of the Ad2 protease (method 1). Heat treated ^{35}S Ad2 and the peptide MSGGAFSW were incubated at 37°C, for 3 and 1 hours respectively, with disrupted Ad2 at pH values ranging from 4.5 to 8.5. The buffers used for both assays were 0.1M Tris/acetate (pH4.5-5.5), 0.1M Tris/maleate (pH6-7) and 0.1M Tris/HCl (pH7.5-9). The samples were analysed by SDS-PAGE and HPLC and the amount of digestion was determined by scanning densitometry of the autoradiograms and integration of the peptide peaks on the HPLC. The resulting pH curves show that in both systems the protease operates best at neutral to alkaline pH values with a decrease in activity below pH6 (**fig.2.5a**). Protease activity could be detected over a broad pH range in the peptide assay and even at pH4.5 20% of the maximum activity was evident. Examination of the pH profile obtained using pVII as a substrate reveals a narrower pH optimum (pH6.5-7.5) with no cleavage of pVII to VII taking place below pH5. These results which were obtained using heat treated ^{35}S Ad2 as a substrate are in good agreement with those described by Bhatti and Weber (1979b) who used ^{35}S Ad2ts1 as a substrate.

The effect of ionic strength on Ad2 protease activity was also determined using the ^{35}S

and peptide assay systems with incubation times of 3 and 1 hours respectively (method 1). Activity was tested at NaCl concentrations ranging from 0 to 2M and quantitative assays were carried out as described above. A significant increase in activity was observed between 0 and 1M NaCl in the peptide assay. In contrast no cleavage of the heat treated, ^{35}S labelled pVII was detected at salt concentrations above 0.5M (**fig.2.5b**).

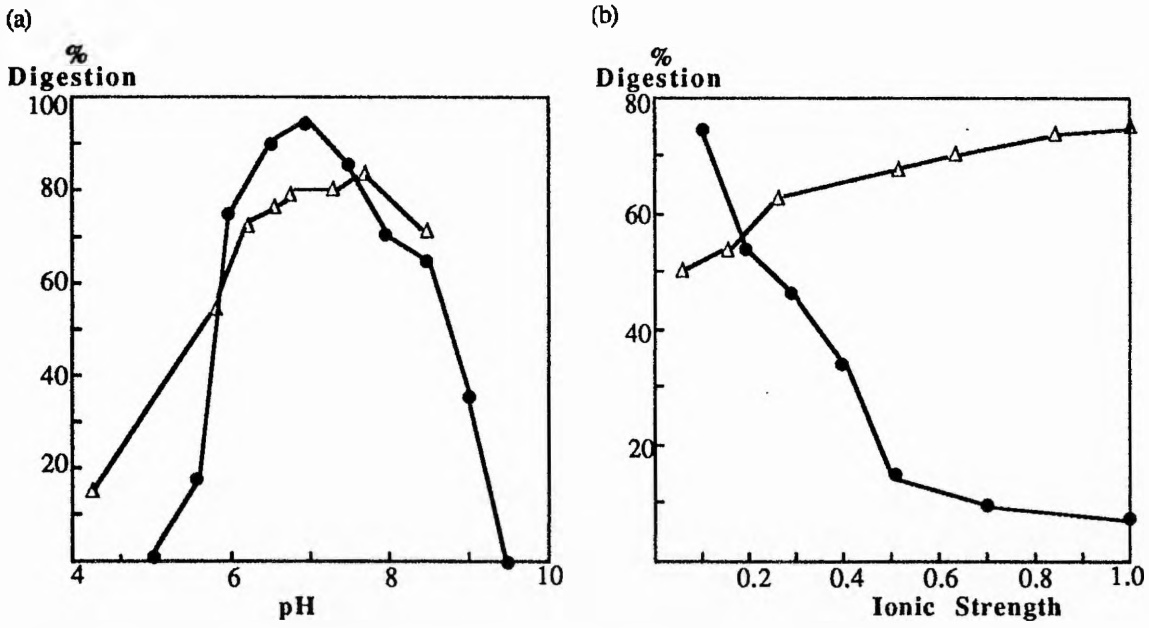


Fig 2.5 Effect of pH and Ionic Strength on Ad2 Protease Activity

The effect of (a) pH and (b) ionic strength on protease activity using the two assay systems. Activity is expressed as % digestion of pVII to VII using ^{35}S methionine labelled Ad2 as the substrate (●) following a 6 hour incubation at 37°C and as % digestion of the peptide MSGGAFSW (Δ) following a 1 hour incubation at 37° C. Assays were carried out according to method 1, with 2mM DTT being included in the peptide assay buffer Assays were carried out in triplicate. Standard deviations were <5%.

Thus the peptide assay was able to detect protease activity at more extreme values of pH and ionic strength than the ^{35}S assay. This is presumably because there are fewer conformational constraints on the octapeptide MSGGAFSW than on the protein substrates.

3 Effects of Divalent Ions on Ad2 Protease Activity

The peptide assay was used to assess the effects of Ca^{+2}, Mg^{+2}, Zn^{+2} and Cu^{+2} on Ad2 protease activity at concentrations ranging from 0.01 to 5mM. The counterion in each case was chloride and method 1.3 was followed, with the exception that no DTT was included in the assay buffer. Calcium and magnesium were found to have no effect on activity, but zinc and copper completely inhibited the enzyme when present at a concentration of 50μM

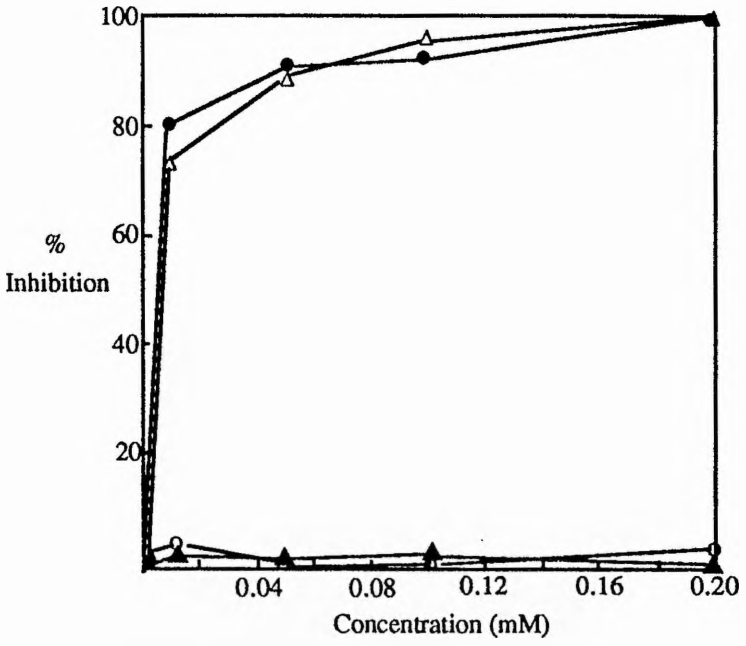


Fig. 2.6 Effect of Divalent Ions

The effect of divalent ions on activity of the adenovirus protease using the peptide assay. Digestions were carried out for 1 hour at 37°C in the presence of Cu^{+2} (●), Zn^{+2} (Δ), Ca^{+2} (○) and Mg^{+2} (▲). Assays were carried out in triplicate, with the standard deviations being <5% in each case.

(**fig.2.6**). In other experiments Ni^{+2} and Fe^{+2} were also found to inhibit protease activity at concentrations of 50μM, whilst Mn^{+2} had no effect on activity even at concentrations of 5mM (results not shown). The inhibitory effects of the divalent ions were completely reversed by the subsequent addition of 25mM EDTA.

4 Inhibition Studies

The neutral pH optimum and insensitivity to chelating agents of the Ad2 protease would normally exclude it from the aspartic and metallo protease classes respectively. It was decided, therefore, to concentrate mainly on serine and cysteine protease inhibitors in this section.

The susceptibility of the Ad2 protease to a range of inhibitors was investigated using both the ^{35}S and peptide assay systems as described in method 1.4. Samples were analysed by SDS-PAGE and autoradiography and the degree of digestion from pVII to VII in the presence of each inhibitor was compared with the control lanes (**fig.2.7**). The results from the peptide and ^{35}S assays were identical and are summarised in **table 2.6.**

The cysteine protease inhibitors iodoacetate, PCMB, dithiodipyridine (pH5 and pH7.2) and NEM gave complete inhibition in both assay systems (**fig.2.7**, lanes 2,3,5,9 and 10;

Inhibitor	Peptide Assay	^{35}S Assay
PMSF (5mM)	*20% inhibited	-
DCI (0.1mM)	-	-
TLCK (1mM)	NT	40% inhibited
TPCK (1mM)	NT	60% inhibited
Chymostatin (10μM)	-	-
SBTI (0.1mg/ml)	-	-
Iodoacetate (1mM)	inhibited	inhibited
PCMB (0.1mM)	inhibited	inhibited
DTDP (1mM)	inhibited	inhibited
Cystatin (2μM)	-	-
E64 (10μM)	-	-
Leupeptin (2μM)	-	-
o-Phenanthroline (10mM)	-	-
EDTA (10mM)	-	-
Pepstatin (1μM)	-	-

* inhibition by PMSF was reversed by DTT (see fig.2.8)

Table 2.6 Effects of Inhibitors on the Ad2 Protease
The above table summarises the effects of a range of inhibitors on Ad2 protease activity in the peptide and ^{35}S assays (see method 1.4). A dash (-) indicates that the inhibitor had no effect and NT is included where the inhibitor was not tested. Compare results with the list of diagnostic inhibitors in table 2.2.

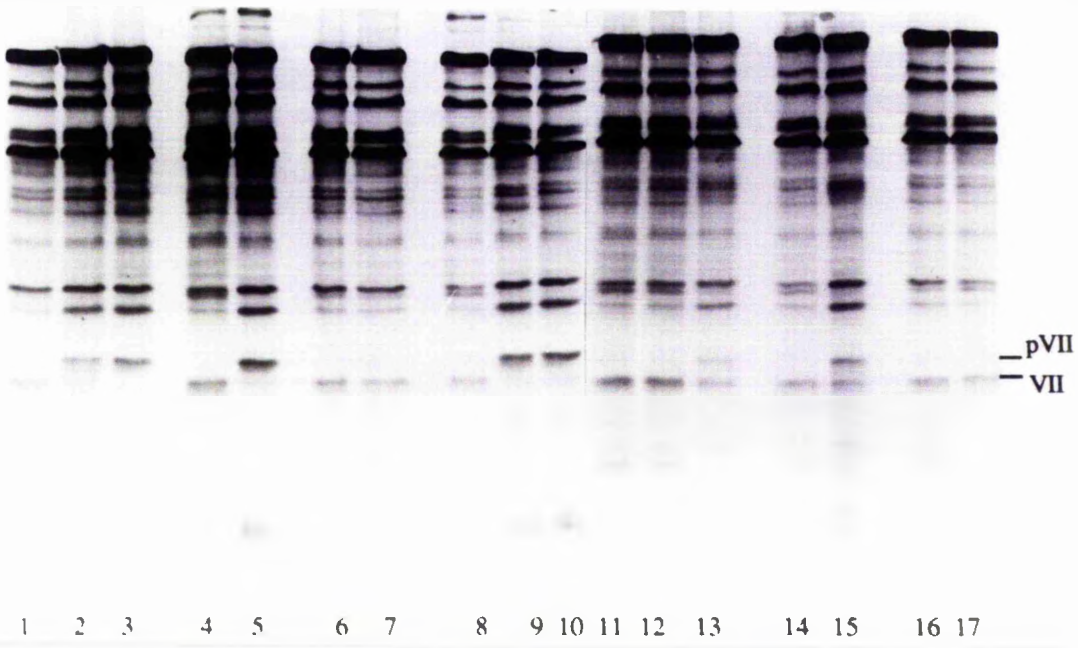


Fig. 2.7 Effects of Inhibitors on Ad2 Protease Activity using the ^{35}S Assay
Autoradiogram of reduced samples separated on a 15% SDS polyacrylamide gel. ^{35}S assays were carried out for each inhibitor or control buffer/solvent as described in method 1.4. *Lanes* 1, 4, 6, 8, 11, 14 and 16 are the control incubations for lanes 2-3, 5, 7, 9-10, 12-13, 15 and 17 respectively. Disrupted, purified Ad2 was pre-incubated with *lane 2*, iodoacetate; *lane 3*, PCMB; *lane 5*, DTDP at pH5; *lane 7*, leupeptin; *lane 9*, DTDP at pH7.2; *lane 10*, NEM; *lane 12*; PMSF; *lane 13*, TLCK; *lane 15*, TPCK and *lane 17*, SBTI. Incubations of the disrupted purified Ad2 with the 35S labelled Ad2 substrate were for 6 hours at 37°C.

table 2.6). Set against this, however, the peptide based inhibitors E64 and cystatin which are diagnostic for papain-like cysteine proteases did not inhibit the enzyme (**table2.6**). In agreement with Bhatti and Weber (1979b) the Ad2 protease was found to be partially inhibited by TLCK and TPCK in the ^{35}S methionine assay (**fig.2.7**, lanes 13 and 15). It was not possible to assess the effects of these 2 inhibitors in the peptide assay because their high u.v. absorbance interfered with detection. The usual solvents of PMSF (propanol, dimethylformamide and acetonitrile) were all found to have significant inhibitory effects on the proteinase on their own, even at concentrations of 5% (results not shown). As a consequence it was necessary to add solid PMSF to the incubation mixtures. Since this inhibitor is sparingly soluble in aqueous solution, control experiments were carried out using trypsin. Rapid and complete inhibition of trypsin was achieved in the parallel control experiments. In the ^{35}S methionine assay complete digestion of pVII to VII had occurred after 6 hours in the presence of PMSF (**fig.2.7**, lane 12). 5mM PMSF showed partial inhibition (20%) in the peptide assay at short digestion times, but this inhibition was completely reversed by the addition of 2mM DTT (**fig.2.8b**). The effects of DFP on the protease could not be determined due to the strong inhibitory effect of its solvent, propanol, in both assay systems; but the serine protease inhibitors 3,4-DCI and chymostatin failed to inhibit it. SBTI and the peptide based inhibitor leupeptin had no effect on protease activity (**fig.2.7**, lanes 17 and 7; **table 2.6**).

The conclusion from these studies is that the inhibitor profile of the Ad2 protease fits better into the cysteine class than the serine class; but the enzyme is not inhibited by cystatin and E64 (**table 2.6**), so does not appear to be a papain-like protease.

5 Effect of DTT and Cysteine on Ad2 Protease Activity

To investigate further the possibility that the Ad2 protease is a cysteine protease the effects of the thiol reagents DTT and cysteine on the enzyme's activity were examined using the peptide assay (method 1.3). Optimum activity was obtained at 0.5-2.5mM DTT or cysteine. The degree of activation decreased as the concentrations of the thiol reagents were raised to 10mM (**fig.2.8a**). The activation of the protease by thiol reagents would be consistent with it being a cysteine protease. Furthermore the presence of 2mM DTT completely reversed the

inhibitory effects of zinc, copper, nickel and iron (results not shown).

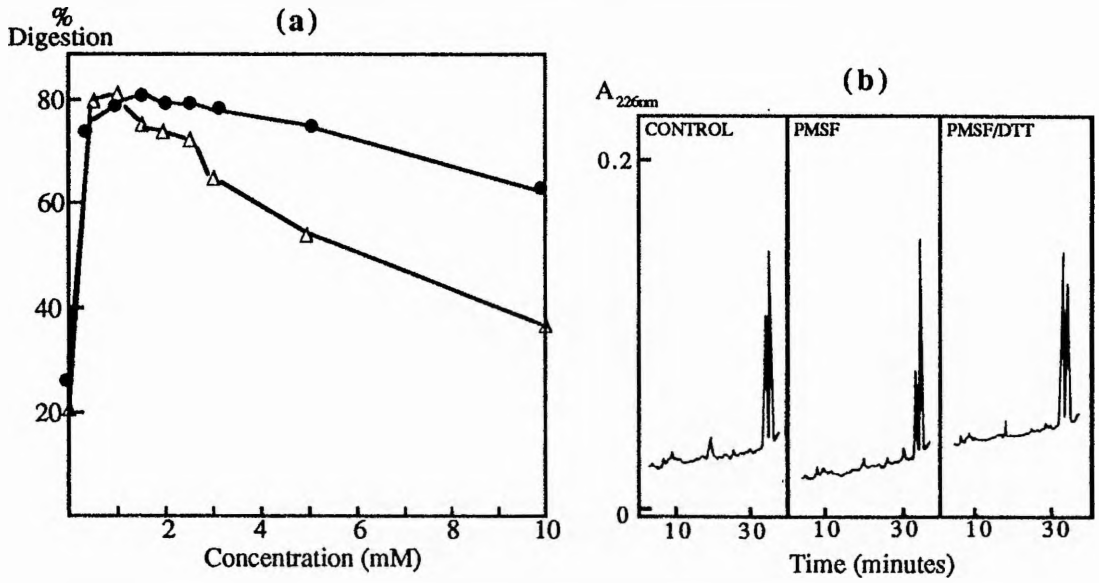


Fig. 2.8 Effect of DTT and Cysteine on Protease Activity
The effect of DTT and cysteine on protease activity using the peptide assay. The graph in (a) shows the % digestion of the peptide MSGGAFSW in 1 hour incubations at 37°C at a range of concentrations of DTT (open triangles) and cysteine (filled circles). Assays were carried out in triplicate and standard deviations were less than 5%. The HPLC profiles in (b) show the reversal of PMSF inhibition by DTT. The incubations were for 1 hour at 37 °C as indicated with the concentrations of PMSF and DTT being 5 and 2mM respectively (for details see method 1.3).

6 Anti-peptide Sera

Peptides corresponding to the N and C- termini and 2 internal sequences of the L3 23-kDa protein were synthesised, coupled to HSA and injected into rabbits as described in method 3. The antisera were assessed for specific reactions with their respective peptides in ELISA reactions and against purified Ad2 using Western blotting (**fig.2.9**). All the antisera reacted well with the peptide against which they were raised; but not with unrelated control peptides. The N-terminal anti-peptide sera had a particularly high titre, reacting at a 1:3000 dilution. The antisera all reacted with a band of apparent mobility 20-kDa in the Western blots. Again the N- and C- terminal peptide antisera gave the strongest reactions with this band. It is assumed, therefore, that the 20-kDa band is the L3 23-kDa protein. Alignment of the blots with stained gels revealed that the band was located in between the VII and pVII bands (results not shown). The results also suggest that the 20-kDa protein is intact, containing its N- and C-termini. This observation conflicts with the suggestion by Chatterjee and Flint (1987) that the protein is autocatalytically processed and will be investigated further in chapter 3.

The antisera will be of great use in monitoring the purification of the L3 23-kDa

putative protease.

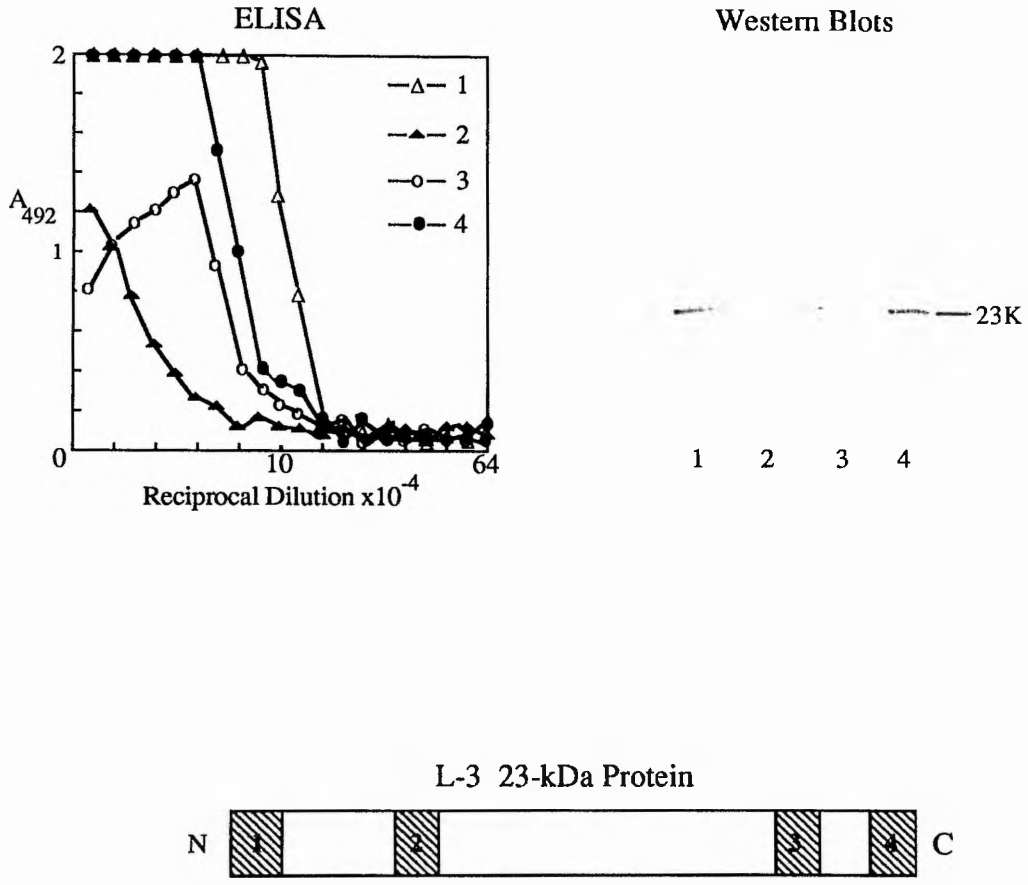


Fig. 2.9 Preparation of Anti-peptide Sera Against the L3 23-kDa Protein.
Four peptides corresponding to sequences from the 23-kDa protein were synthesised, coupled to albumin and injected into rabbits (see method 3). The schematic diagram of the protein shows the positions of the 4 peptides in its primary sequence. A graph showing the reaction of each of the resulting antisera against the appropriate peptide in an ELISA test is shown above alongside Western blots of the antisera against purified Ad2. 30µl of purified Ad2 were separated on a 12.5% SDS polyacrylamide gel and electroblotted onto nitrocellulose paper which was probed with antisera 1 and 4 (1:1000) and antisera 2 and 3 (1:500).

7 Extraction of Ad2 Protease Activity

When disrupted, purified Ad2 is centrifuged for 15 minutes at 100000g the protease activity and most of the viral proteins including the L3 23-kDa protein were found to be in the pellet. This is as a result of the many protein-protein and DNA-protein interactions that take place within the virion. The aim of this section was to extract the protease activity in a soluble form from the virus to enable its purification.

The purified Ad2 was extracted for 30 minutes at 4°C with 2M NaCl, 2M urea or 2M urea followed by 2M NaCl/2M urea as described in method 4.1. The supernatants and the pre-centrifugation samples from each extraction were assessed for protease activity using the peptide assay and analysed by SDS-PAGE (methods 1 and 2). **Table 2.7** shows the amount of protease activity in each sample and reveals that whilst 2M urea is able to increase the

Extraction Method	Total Activity	Soluble Activity
50mM Tris/HCl pH8	1	0.05
2M urea	4	0.1
2M NaCl	2.5	0.8
2M urea then 2M urea, 2M NaCl	3	1.5

Table 2.7 Extraction of Ad2 Protease from Purified Virions
The above table shows the amount of protease extracted from disrupted, purified Ad2 using a variety of treatments. The 2M urea and 2M NaCl were made up in 50mM Tris/HCl, 5mM EDTA, pH7.5 and the activities of the total extracts and supernatants were assessed using the peptide assay as described in method 1.3. Incubation times were for 1 hour at 37°C. The values in columns 2 and 3 are the protease activities relative to the activity in the total 50mM Tris/HCl pH8 extract which is given a value of 1.

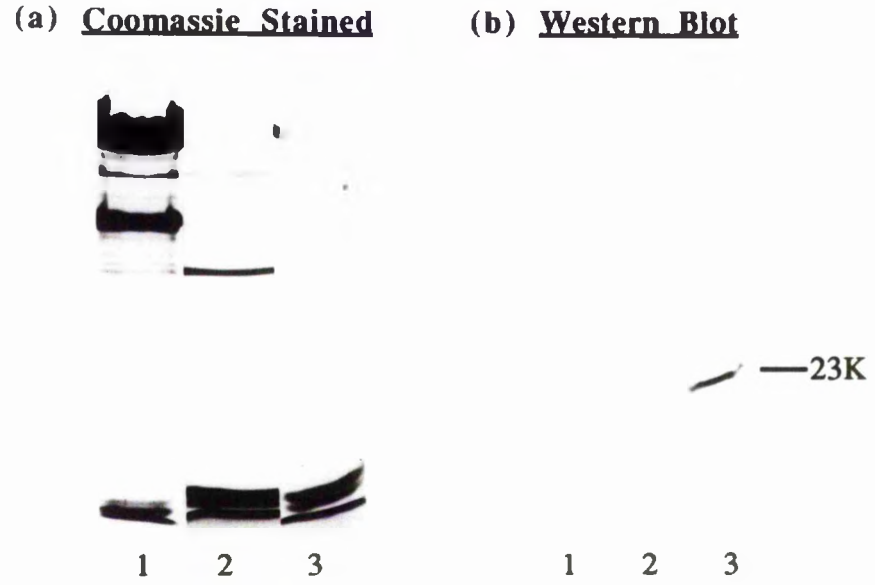


Fig. 2.10 Extraction of the L3 23-kDa Protein from Ad2
Disrupted, purified Ad2 was extracted with 2M urea (*lane 1*), 2M NaCl (*lane 2*) or 2M urea then 2M urea, 2M NaCl (*lane 3*) as described in method 4.1. Two 15µl samples of each supernatant were separated on 12.5% SDS polyacrylamide gels. One gel was stained with Coomassie blue (a) and the other was electroblotted onto nitrocellulose paper and probed with the L3 23-kDa N-terminal peptide antiserum at a dilution of 1:1000 (b).

protease activity in the total extract it does not solubilise the enzyme. In contrast, the total activity in the 2M NaCl extract was less than in the 2M urea extract; but a significant proportion of this was solubilised. Extraction of the virus with 2M urea followed by 2M urea/2M NaCl gave the greatest amount of protease activity in the supernatant and was the chosen method of extraction. In addition to solubilising the enzyme this method also provides a good initial purification step by removing many of the other viral proteins (**fig.2.10a**). The Western blot in **fig.2.10b** shows that significantly more of the L3 23-kDa protein is present in the urea/NaCl supernatant.

In conclusion the Ad2 protease activity can be readily solubilised from purified virions using 2M NaCl and the inclusion of 2M urea increases the yield. The L3 23-kDa protein is present in the soluble active fraction so additional purification steps should allow us to show whether or not the 23-kDa putative protease co-purifies with activity.

8 Purification of the Ad2 Protease

Ad2 was extracted with 2M urea followed by 2M urea/ 2M NaCl and the resulting supernatant was applied to an FPLC S75 gel filtration column equilibrated with 50mM Tris/HCl, 5mM EDTA, 1M NaCl, pH8 as described in the method 4.2. Fractions were collected and assayed for protease activity using the peptide and ^{35}S assays with incubation times of 30 minutes and 3 hours respectively. The peptide assay was as described in method 1.3 except that 10µl of each fraction were assayed and the assay buffer was made 1M with respect to NaCl. 4µl of each fraction were tested for protease activity using the ^{35}S Ad2ts1 assay (method 1.1) Activity was detected in both assays in fractions 13-17 with maximum activity in fractions 14 and 15 (see **fig.2.11**). The fractions were desalted using the FPLC fast desalting column equilibrated with 25mM ammonium bicarbonate, pH8 and then dried down separately and analysed by SDS-PAGE. The stained gel showed that fractions 14-16 contained a single band of apparent molecular weight 20-kDa. The corresponding Western blot showed that this band reacted strongly with the antisera raised against the N-terminal peptide of the Ad2 23-kDa putative protease (**fig.2.11**). Thus the 23-kDa protein has been purified to homogeneity and is active in both the peptide and ^{35}S assays. This shows conclusively that the L3 23-kDa protein is the adenovirus protease. Calibration of the column using molecular

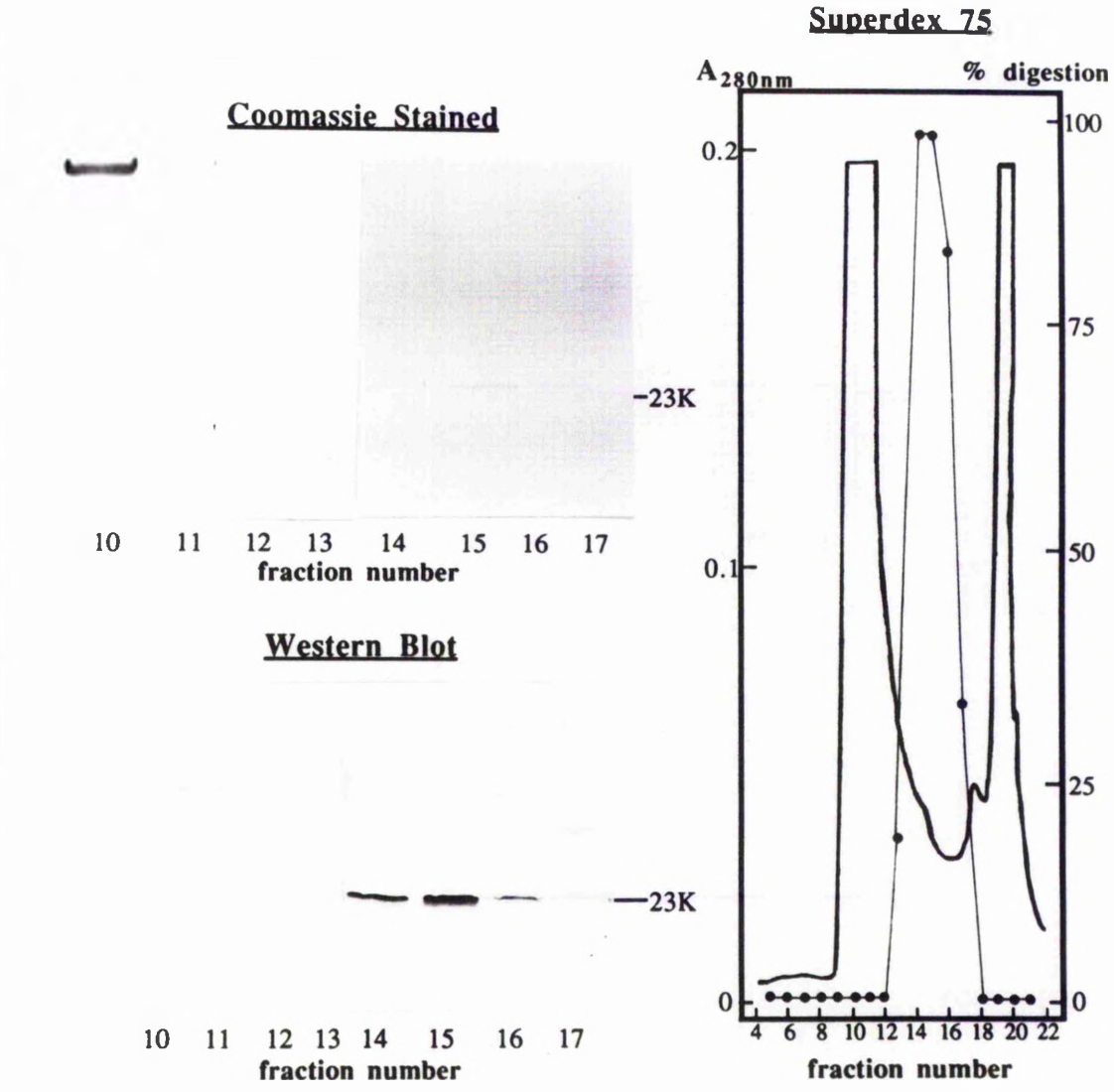


Fig. 2.11 Purification of the L3 23-kDa Protease from Ad2
Disrupted, purified Ad2 was extracted with 2M urea followed by 2M urea/NaCl as described in method 4.2 and the resulting supernatant was applied to an FPLC Superdex 75 column. The profile of the FPLC separation is given above, with the values for the % digestion of the peptide substrate MSGGAFSW by 10µl of each fraction superimposed onto the trace. Fractions 10-17 were desalted, dried and analysed on 12.5% SDS polyacrylamide gels. One gel was stained with Coomassie blue and the other was electroblotted onto nitrocellulose paper and probed with a 1:1000 dilution of the L3 23-kDa N-terminal peptide serum.

weight standards revealed that the protease eluted between SBTI (21-kDa) and myoglobin (17-kDa). This shows that the protease is active as a monomer.

In other experiments the 23-kDa protein was purified in the same way except that the S75 column was equilibrated in 50mM Tris/HCl, pH8, a low salt buffer. In these experiments activity was detected using the ^{35}S assay; but not in the peptide assays. The buffer used for the

peptide assays did not contain NaCl. Thus it appears that the salt may be required to stabilise the purified protease.

9 Effect of Salt on the Purified Protease

Disrupted, purified Ad2 was extracted with 50mM Tris/HCl, 5mM EDTA, 2M NaCl, 2M urea pH7.5 as described in the method 4.3. Half of the sample was applied to the FPLC desalting column equilibrated in 50mM Tris/HCl, pH8 and the other half to the column equilibrated in the same buffer made 1M with respect to NaCl. Peptide assays were set up as described in method 4.3 and **fig.2.12** shows that the protease activity was lost from the protein fraction if the column was run in low salt. Some of this activity could be recovered if the fraction was made 1M with respect to salt immediately after the desalting step. In contrast, when the separation was carried out in high salt protease activity could be detected in the main protein fraction using the peptide assay.

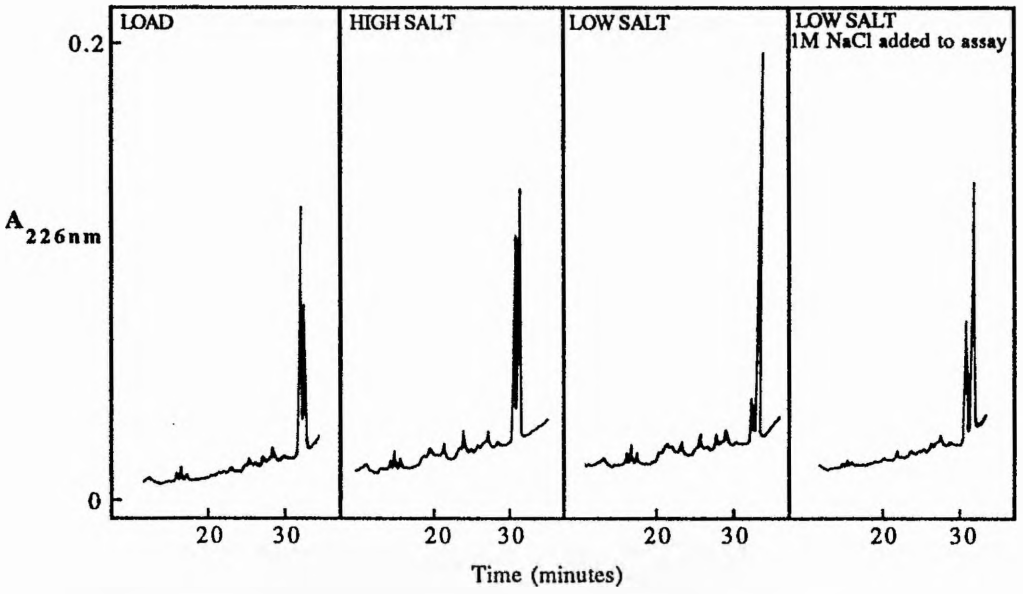


Fig. 2.12 Effect of NaCl on Ad2 Protease Activity
Disrupted purified Ad2 was extracted with 2M urea followed by 2M urea, 2M NaCl and applied to an FPLC desalting column as described in method 4.3. Two separations were carried out; one in which the column was equilibrated with a buffer containing 1M NaCl (high salt) and the other in which the column was equilibrated with a buffer containing 0M NaCl (low salt). 10µl of the main protein fractions were assayed for protease activity in each case as described in method 4.3. HPLC profile 1 shows the digestion of the peptide MSGGAFSW by the pre-column sample; profiles 2 and 3 show the digestion by the main protein fractions collected from the desalting column equilibrated with the high and low salt buffers respectively; profile 4 shows the digestion of MSGGAFSW by the main protein fraction from the low salt separation when 1M NaCl was added back to the fraction immediately after it eluted from the column. Peptide assays were carried out according to method 1.3 and incubations were for 1 hour at 37° C.

In conclusion the L3 23-kDa protein is the adenovirus protease, but in a pure assay system high salt is required for stability. The purified protein was able to cleave the ^{35}S labelled pVII from Ad2ts1 extracts in low salt; but in this case many other viral and cellular components are present. It would appear that one of these is required to stabilise the protease activity.

10 Comparisons of the L3 23K Protein from Different Serotypes

Armed with the knowledge that the L3 23-kDa protein is the protease and not simply an essential cofactor it is now possible to look at its amino acid sequence to try and identify the important catalytic residues. The gene has been sequenced from 9 different serotypes including 2 bovine viruses revealing that the protein is remarkably well conserved between the serotypes. The sequences which are all contained in the NBRF protein sequence data bases were aligned using the programme 'Pileup' from the Wisconsin computing package (chapter 1, method 5). **Fig.2.13a** shows the amount of variation at each residue in the sequence. As can be seen there are stretches of amino acids that are highly conserved among all the serotypes and the N-terminal half of the molecule is particularly well conserved.

Based on the assumption that the adenovirus protease is a cysteine or serine protease, the sequence was examined for conserved His, Asp/Glu and Ser/Cys residues. The positions of these residues in the Ad2 sequence are shown in **fig.2.13b**. Crystallography, active site labelling or site directed mutagenesis would be required to ascertain which residues are required for proteolysis, but it may be possible learn more from sequence alignments and secondary structure predictions.

11 Database Searches

The Ad2 protease sequence was searched against the NBRF protein sequence database and a subset of the database containing all known protease sequences using the method of Coulson *et al.* (1987). These searches revealed no significant homologies with any known cellular or viral proteases.

It has been shown that the secondary structure of proteins is conserved to a greater extent than primary structure and that comparison of secondary structures is a good means of testing whether or not 2 enzymes are related. Gorbalenya has used secondary structure

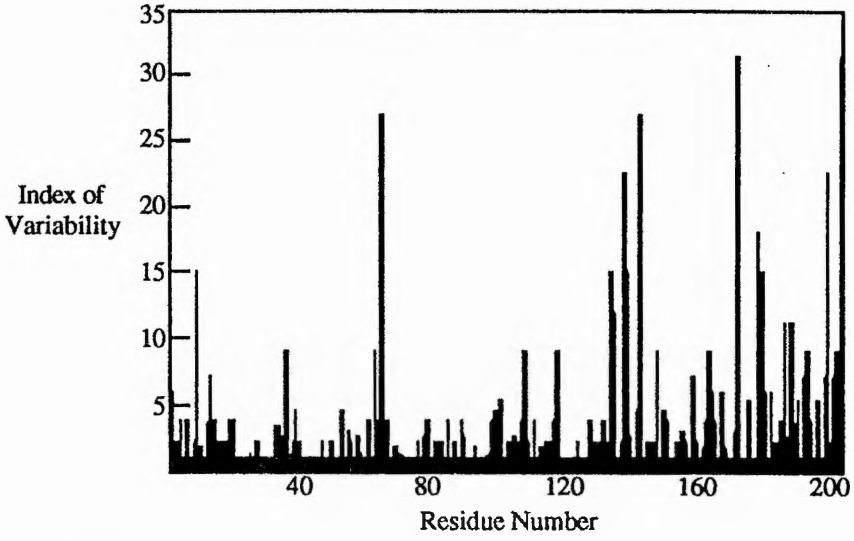


Fig. 2.13a Conservation of Protease Sequence
Histogram showing the distribution of conserved regions in the adenovirus protease sequences from 7 human and 2 bovine serotypes. The index of variability is calculated by first dividing the number of amino acids found at a given position by the frequency of the most common. Where the amino acids at one position are all of one group (+ve charge, -ve charge, large apolar etc) the value is divided by 2 to give the final index. A high value indicates a high variability at that position.

```
  1 MGSSEQELKA IVKDLGCGPY FLGTYDKRFP GFVSPHKLAC AIVNTAGRET
 51 GGVHWMAFAW NPRSKTCYLF EPFGFSDQRL KQVYQFEYES LLRRSAIASS
101 PDRCITLEKS TQSVQGPNSA ACGLFCCMFL HAFANWPQTP MDHNPTMNLI
151 TGVPNSMLNS PQVQPTLRRN QEQLYSFLER HSPYFRSHSA QIRSATSFCH
201 LKNM
```

FIG 2.13b Sequence of the Adenovirus 23 kDa Protease
The sequence of the 23 kDa protease from adenovirus type 2 is shown with the histidine, cysteine, serine and aspartic acid residues conserved in all 7 human and 2 bovine serotypes indicated by an asterisk. Derived from sequence alignments created with the Pileup program in the Wisconsin package (Devereaux *et al.*,1984).

alignments to compare the adenovirus protease with other viral and cellular proteases without success (A.E. Gorbalenya, personal communication). In conclusion, the adenovirus protease appears to be different from all the cellular and viral proteases studied to date and the elucidation of its crystal structure is required to establish its mechanism and classification with certainty.

DISCUSSION

Two complementary assay systems have been developed for the adenovirus protease. The first is based on a method previously described by Bhatti and Weber (1978) and utilises ^{35}S labelled, Ad2 infected cell extracts as the substrate. The extracts contain endogenous activity which can be removed by heating to 60°C if required. In some instances it is preferable to study the cleavage of pVII by the endogenous protease, since this is more akin to the *in vivo* situation than the use of heat treated Ad2 or labelled Ad2ts1 extracts. The second assay utilises the synthetic octapeptide, MSGGAFSW, as a substrate. The digestion of this peptide, which is derived from the established cleavage site in one of the natural substrates of the protease, pVI, can be readily monitored by HPLC or TLE.

The lack of a pure source of enzyme for the most part of this project has prevented kinetic analysis using either the peptide or natural substrates. Preliminary results reveal, however, that the K_m value for the cleavage of MSGGAFSW by the Ad2 protease is 8mM (G.Kemp, personal communication). Although this indicates that the affinity of the enzyme for the peptide is low, the K_m value is similar to those calculated for peptide substrates of other viral proteases (Kay and Dunn, 1990). The reason for the low affinity is probably because elements of secondary and tertiary structure are involved in the binding. In the natural substrates amino acids outwith the P_4-P_4' region will be required to present the scissile bond in the appropriate conformation. The peptide substrates, on the other hand, are relatively flexible and so will only adopt the correct conformation for short periods of time.

The advantage of the peptide assay over the ^{35}S assay is that it involves a pure substrate free of any other cellular or viral components. It is also faster, more cost effective and in the future could be used to study the kinetics of inhibitors of the adenovirus protease. The ^{35}S assay, on the other hand, utilises a natural substrate of the protease and is more suitable for screening crude extracts that may contain cellular proteases. As far as possible, the 2 assays were used in parallel to characterise the protease and in the main the results obtained were in complete agreement.

The only major discrepancy came when the effects of pH and ionic strength were examined (**fig.2.5**). Using the peptide assay a slightly broader pH optimum was observed and activity was found to increase with NaCl concentration; in contrast to the findings with the ^{35}S assay, where no cleavage of pVII took place above 0.5M NaCl. The narrower pH optimum and the inhibitory effects of high ionic strength, found in the ^{35}S assay, are probably a function of the influence of pH and ionic strength on the structure of the pVII protein, as well as on the enzyme. It appears then that the peptide assay is able to operate under more extreme conditions of pH and ionic strength, presumably because the peptide substrate has fewer conformational restraints.

Several reasons could account for the stimulatory effect of salt observed in the peptide assay (**fig.2.5b**). Perhaps the most likely of these is that the salt releases the protease from associations with other proteins or DNA in the disrupted virus, since extraction with 2M NaCl was found to be an effective way of solubilising protease activity (**table 2.7**). The apparent requirement of the purified protein for high salt in the peptide assay, however, suggests that it must also play some role in stabilising the enzyme in an active form (**fig.2.12**). Studies on the cleavage of peptide substrates by purified avian retroviral proteases and cellular aspartic proteases reveal that high salt not only increases activity, but, in some cases, also alters the cleavage site preferences of the enzymes (Kotler *et al.*,1989). The mechanisms behind these phenomena remain to be elucidated.

The neutral pH optimum of the Ad2 protease and its activity in the presence of the chelating agent, EDTA, would normally rule it out of the aspartic and metallo protease classes respectively. Furthermore, pepstatin, which is the diagnostic inhibitor for aspartic proteases, did not inhibit protease activity in either assay system and no divalent ions were found to have a stimulatory effect on its activity (**table 2.6** and **fig.2.6**). On the contrary, Cu^{+2}, Ni^{+2}, Fe^{+2} and Zn^{+2} were all found to inhibit the Ad2 protease, even at concentrations as low as 10μM. Processing of the picornavirus polyproteins has been shown to be inhibited by 1mM $ZnCl_2$, but no mechanism has been suggested for this inhibition (Butterworth and Korant, 1974). A number of divalent ions including those of zinc, copper and nickel are known to interact with surface exposed imidazole and thiol groups of proteins (Porath *et al.*,1975), so the

inhibition could be as a result of direct binding to an active site cysteine or histidine residue. The aspartic protease of HIV-1 has been shown to be inhibited by Cu^{+2} at concentrations of 25μM and its effects have been considered in some detail. It is suggested that the surface exposed Cys_{67} residue forms a weak interaction with His_{69} and that together they could constitute metal binding site (Karlstrom and Levine,1991). Interestingly cellular aspartyl proteases, including pepsin, have not been shown to be inhibited by copper, or any of the other transition elements at neutral pH, so these observations open the way for the design of substrate analogue inhibitors containing metal ions (Lunblad and Stein,1969; Karlstrom and Levine,1991). The design of such compounds has recently been reported and phenylnorstatine derivatives have been synthesised that inhibit the HIV-1 protease with K_i's of 3-24μM (Raju and Deshpande,1991). Presumably the inhibition of the Ad2 protease by zinc, copper, iron and nickel must be mediated by surface exposed cysteine or histidine residues, possibly at the active site.

Having eliminated the adenovirus protease from the aspartic and metallo classes, the effects of a range of serine and cysteine protease inhibitors were examined. The enzyme had been previously classified as a chymotrypsin-like serine protease based on reports that it is inhibited by DFP, PMSF, TPCK and TLCK (Bhatti and Weber,1979b; Tremblay *et al.*,1983). In agreement with previous studies TPCK and TLCK were both found to partially inhibit the protease in the ^{35}S assay (**fig.2.7**). In contrast, the classical serine protease inhibitors, PMSF and 3,4-DCI, failed to inhibit the Ad2 protease in the ^{35}S assay despite rapid and complete inhibition of trypsin in parallel control experiments (**table 2.6**). In the peptide assay partial inhibition was observed for PMSF, but only when DTT was excluded (**fig.2.8b**). It has been shown that PMSF, TPCK and TLCK are also effective inhibitors of cysteine proteases. Their inhibition by PMSF, however, is readily reversed by DTT and whilst TPCK and TLCK can discriminate between members of the serine protease class, they are much less specific inhibitors of cysteine proteases (Whittaker and Perez-Villasenor,1968). Thus the inhibitory effects of TPCK, TLCK, 3,4-DCI and PMSF on the adenovirus protease are much more in line with it being a cysteine protease than a serine protease.

The complete inhibition of the Ad2 protease by 1mM iodoacetate, NEM or DTDP and

by 0.1mM PCMB in both assay systems adds further weight to this argument (**table 2.6** and **fig.2.7**). Iodoacetate, NEM and DTDP at neutral or alkaline pH are general thiol reagents, reacting with the S^- ion, so their effect could be at any free thiol group in the protease. Inhibition by DTDP at pH5, however, suggests the presence of a thiol group with a very low pK. This can be found in the active site of cysteine proteases and indeed DTDP at pH3-5 is regarded as a specific active site titrant for cysteine proteases (Willenbrock and Brocklehurst, 1985). Inhibition by 0.1mM PCMB is also regarded as a good indication that an active site thiol group is present, as is the activation of an enzyme by DTT or cysteine (Barrett, 1977). The activity of the Ad2 protease increased significantly in the presence of 0.5-2mM DTT or cysteine as can be seen in the graph in **fig.2.8a**.

Although the results presented here strongly suggest that the adenovirus protease is a member of the cysteine class, the lack of inhibition by E64, leupeptin and cystatin (**table 2.6**) means that a definitive diagnosis cannot be made. E64 is a peptide epoxide that inhibits the papain-like cysteine proteases, whilst leupeptin (N-acetyl-Leu-Leu-Arg-al) inhibits most cysteine proteases. The arginine in leupeptin means that it will also inhibit serine proteases with a trypsin-like substrate specificity (Salvesen and Nagase,1989). Assuming that the adenovirus protease is a cysteine protease, its lack of inhibition by E64 and leupeptin could be explained by its unusual substrate specificity, described in chapter 1. The large molecular weight inhibitors, cystatin and SBTI, also failed to inhibit the protease (**table 2.6**). These, however, tend to be more selective within the cysteine and serine classes respectively and so their lack of effect is not surprising.

The conclusion from these inhibitor studies is that whilst the adenovirus protease does not appear to be a serine protease, its inhibitor profile differs from the classical papain-like cysteine proteases. Interestingly, inhibitor studies carried out on the rhinovirus 3C protease by Orr *et al.*, (1989) show very similar results. The only significant difference is that the 3C protease is inhibited by the 2 peptide aldehydes, leupeptin and chymostatin, specific for cysteine and chymotrypsin-like proteases respectively. Thus, like the adenovirus protease, the inhibitor profile of 3C does not fit exactly into either the serine or the cysteine class. As mentioned previously, the picornaviral 2A and 3C proteases belong to the recently established

group of trypsin-like cysteine proteases. Their evolutionary and mechanistic relationship with serine proteases in conjunction with an active site thiol group could provide a plausible explanation for their unusual inhibitor profiles. One interpretation of the results presented here is that the adenovirus protease also belongs to the subclass of cysteine proteases that are related to trypsin. In order to test this hypothesis further, the Ad2 protease sequence must be aligned with other cellular and viral proteases.

A major aim of this project was to establish with certainty that the L3 23-kDa protein is the adenovirus protease. This was achieved with the aid of antipeptide sera raised against synthetic peptides corresponding to the N- and C- termini of the 23-kDa protein. The antisera both reacted with a protein from purified Ad2 that had an apparent molecular weight of 20-kDa (**fig.2.9**). These results suggest that the 23-kDa putative protease is present in the virus in an intact form and that it is not cleaved at A_{46}-G_{47} as suggested by Chatterjee and Flint (1987). It cannot be ruled out, however, that the anti-peptide sera cross react with other regions of the protein and this possibility will be considered in chapter 3.

Extraction of protease activity from purified Ad2 in a soluble form required extraction with 2M urea followed by 2M urea/2M NaCl and even after such treatment not all the activity was released (**table 2.7**). This indicates that the protease must form strong associations with viral proteins and/or DNA. The NaCl/urea supernatant was purified further by gel filtration using an FPLC S75 column and the resulting fractions were tested for protease activity using the ^{35}S and peptide assays. Analysis of the active fractions by SDS-PAGE revealed that they contained a single band of apparent molecular weight 20-kDa and that this reacted with the antisera raised against the 23-kDa N-terminal peptide (**fig.2.11**). It has, therefore, been shown that the L3 23-kDa protein of Ad2 is indeed the virally encoded protease. Furthermore, the retention time of the protease reveals that it has an apparent molecular weight of 20-kDa, showing that it is active as a monomer. Surprisingly, when the Ad2 protease was purified by gel filtration in 50mM Tris/HCl, 5mM EDTA, pH8 (low salt) it was active in the ^{35}S assay, but not in the peptide assay (**fig.2.12**). It is concluded that some component of the virus is required to activate or stabilise the protease. Purification of the enzyme results in its inactivation unless carried out in the presence of high salt, which stabilises the protease in an

active conformation. Protease purified in low salt can be reactivated in the ^{35}S assay since this system includes other viral and cellular factors. The possibility that another component is required to activate the adenovirus protease will be investigated in chapter 4.

Having established that the L3 23-kDa protein is the adenovirus protease, its sequence was analysed to see if it resembled any of the established viral or cellular proteases. The gene which has been sequenced in 9 different adenovirus serotypes is extremely well conserved (Houde and Weber,1990a); but differences do exist, so computer alignments can be used to identify the important residues (**fig.2.13a**). Both serine and cysteine proteases contain active site histidine residues, so it is significant that there is only one conserved histidine residue, at position 54, in the protease sequence. Indeed, it would seem likely that His_{54} is one of the catalytic residues in the protease. The fact that His_{54} is followed by a tryptophan residue, is also of interest, in view of the inhibition of the Ad2 protease by a number of divalent ions including copper and nickel (**fig.2.6**). This is because the dipeptide sequence His-Trp has been shown to have an unusually high affinity for Ni(II), Co(II) and Cu(II), as demonstrated by its use as a polypeptide tag to facilitate the purification of genetically engineered proteins (Smith *et al.*,1988; Sassenfeld,1990). An attractive explanation for the inhibition of the adenovirus protease by a number of the divalent transition elements is that the accessible His-Trp dipeptide at the active site of the enzyme is a metal binding site. It is more difficult to predict the identities of the other catalytic residues, but assuming that the enzyme is a cysteine protease there are 3 conserved cysteine residues in the sequence. These are located at positions 104, 122 and 126, with the neighbouring amino acids being particularly well conserved in the case of Cys_{104} and Cys_{122}. In papain the nucleophilic cysteine residue is located 125 residues to the N-terminal side of the active site histidine. Since there are no conserved cysteine residues on the N-terminal side of His_{54} it seems unlikely that the adenovirus protease is a papain-like enzyme. The inhibitor profile of the Ad2 protease is identical to that of the rhinovirus 3C protease which is a member of the subclass of cysteine proteases that are related to trypsin. One of the 3 conserved cysteine residues at position 104, 122 or 126 could be the nucleophilic residue in such an enzyme. Although unlikely, based on its inhibitor profile, it

cannot be ruled out that the Ad2 protease is a serine protease and in this case the conserved Ser_{183} or Ser_{95} could be the active site residue. There are a number of conserved aspartate and glutamate residues that could act as the third residue in the catalytic triad. Set against this, however, is the fact that computer searches have failed to find any significant homologies between the Ad2 protease and trypsin-like enzymes, or any other proteases for that matter. Thus the classification of the adenovirus protease remains in doubt. Site directed mutagenesis and the elucidation of its crystal structure will be necessary to determine whether it is a trypsin-like cysteine protease or whether it has a novel mechanism. Expression of the active protease in large amounts must be carried out in order to meet these ends and this is the primary aim of the next chapter.

CHAPTER 3

Expression of the Ad2 L3 23 kDa Protein, Autocatalytically Processed ?

The aim of this chapter is to purify active Ad2 protease in large amounts and to establish whether or not the enzyme is autocatalytically processed. The low levels of expression of the 23-kDa protein during adenovirus infection make its purification from the virus, in sufficient quantities, almost impossible. In the previous chapter the total yield of protease obtained from 6 litres of infected Hela cells was only 2μg. It was, therefore, deemed necessary to use *E.Coli* and baculovirus systems to express the Ad2 protease. Such systems have enabled considerable advances to be made in the structural and mechanistic characterisation of other viral proteases including those of HIV and poliovirus (Graves *et al.*,1988; Baum *et al.*,1991). In the longer term and beyond the scope of this project the incorporation of the 23-kDa cDNA into an appropriate vector would provide the means to carry out systematic site directed mutagenesis, designed to identify the amino acids constituting the active site and substrate binding pockets of the enzyme.

AUTOCATALYTIC PROCESSING OF THE AD2 PROTEASE?

The expression of the 23-kDa protein in several systems should give a clearer picture as to what, if any, post-translational modifications it requires. In a report by Chatterjee and Flint (1987) it was proposed that the 23-kDa protein is activated by limited proteolysis to give a 19-kDa protease and that both species are phosphorylated on a tyrosine residue. They suggested further that the cleavage is at the Ala_{46}-Gly_{47} bond and that autocatalytic processing is involved. The evidence for this came from parallel labellings of wild type Ad2 and Ad2ts1 infected cells with ^{32}P at non-permissive temperatures. Analysis of the mobilities of the labelled proteins by SDS-PAGE revealed a band at 23-kDa in the purified Ad2ts1 particles. This was absent in wild type Ad2; but these virions contained several phosphorylated bands, missing from the ts1 particles, including one with an apparent mobility of 19-kDa. The authors

also reported that viral proteins with mobilities of 23-kDa and 19-kDa, in Ad2ts1 and wild type Ad2 respectively, were labelled with ^{3}H DFP *in vitro.* The latter is a potent inhibitor of serine proteinases and is reported to bind specifically to active site serine residues (Barrett,1977). Solely based on this evidence the authors were able to conclude that the 23-kDa protein in Ad2ts1, encoded by the L3 mRNAs, is proteolytically processed to give a 19-kDa protein, which is the active adenovirus endoprotease. Further support for this theory came from Houde and Weber (1990a) who expressed the Ad2 23-kDa putative protease as a fusion product with protein A using the vector pRIT2T. They claimed that the purified fusion protein was capable of autocatalytic cleavage and that it cleaved its natural viral substrates as well as foreign proteins. The expression of an apparently active enzyme in *E.Coli* led them to conclude that mammalian specific post-translational modifications are not required for Ad2 protease activity.

In chapter 2, the active Ad2 protease was found to have an apparent molecular weight of 20-kDa and to react with antisera raised against peptides corresponding to its N- and C-termini. This suggests that the 20-kDa species contains all 204 amino acids, coded for by the L3 23-kDa open reading frame and so contradicts the proposals that the enzyme is proteolytically processed (Chatterjee and Flint,1987; Houde and Weber,1990a). The specificity of the antipeptide sera will be confirmed in this chapter by testing them against deletion mutants of the 23-kDa protein which lack either residues 1-18 or 181-204.

<u>BACTERIAL AND BACULOVIRUS EXPRESSION SYSTEMS</u>.

Three sets of vectors will be used to express the 23-kDa putative Ad2 protease:

(1) **pGEX-2T**, an *E.Coli* vector that expresses foreign proteins as fusion products with glutathione transferase under the control of the beta-galactosidase promoter.

(2) **pET3c**, an *E.Coli* vector that expresses foreign proteins under the control of the T7 promoter.

(3) **pAcRP23** and **pVL1393**. These expression vectors are used to incorporate the genes of foreign proteins into the baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV) so that they are under the control of the very strong promoter of the polyhedrin gene.

The background to these vector systems will be discussed in turn.

pGEX-2T

The construction and detailed description of the pGEX series of vectors is reported by Smith and Johnson (1988) and a map of pGEX-2T is given in **fig.3.1.** The plasmids have been constructed so that they direct the synthesis of foreign proteins in *E.Coli* as fusions with the C-terminus of Sj26, a 26-kDa glutathione S-transferase (GST). The pGEX-2T vector has

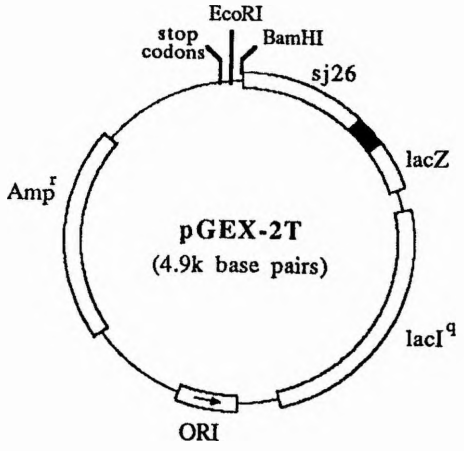


Fig. 3.1 Map of the pGEX-2T vector.
The map of the pGEX-2T plasmid is shown with the BamHI and EcoRI cloning sites marked.

been engineered so that the coding sequence for a thrombin cleavage site (LVPR-GS) is included, in frame, at the 3' end of the GST gene. The expression of the fusion proteins is under the control of the lacZ promoter and since the plasmids also encode for the lac repressor (lacI), the fusion proteins will only be expressed in the presence of the inducer molecule isopropyl-β D-thiogalactopyranoside (IPTG). Thus the synthesis of a foreign protein, that may be toxic to the *E.Coli*, can be switched on by the addition of IPTG after the cells have reached the required level of growth. The vectors include an ampicillin resistance gene, enabling the positive selection of transformed cells. Attachment to the carrier molecule GST means that proteins can be readily purified by affinity chromatography on a glutathione Sepharose column and then be cleaved from the carrier using thrombin. Smith and Johnson (1988) report that although insolubility is a frequent characteristic of foreign proteins expressed in *E.Coli* the majority of GST fusion proteins expressed, to date, are at least partially soluble.

In conclusion the pGEX system appears to offer several advantages over other fusion

protein-producing vectors (Sassenfeld,1990) and seems a good system to verify the report by Houde and Weber (1990a) that the 23-kDa protein is capable of autocatalytic processing. It was also the system chosen for the expression of the N- and C- terminal deletion mutants of the 23-kDa protein.

pET Vectors

The pET series of vectors are comprehensively described by Rosenberg *et al.* (1987) and Studier *et al.*(1990) and are examples of T7 expression systems that take advantage of the highly specific nature of the T7 RNA polymerase. The plasmids used for cloning and expressing cDNAs are based on the multi-copy plasmid pBR322 and carry a T7 promoter followed by a unique cloning site. A map of pET3c is given in **fig.3.2.** For efficient expression the vectors must be transformed into an *E.Coli* strain that expresses the T7 RNA

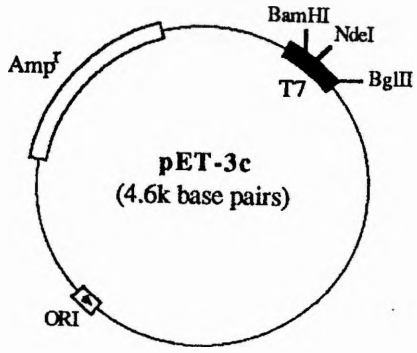


Fig. 3.2 Map of the pET3c vector
A map of the pET3c vector with the position of the NdeI cloning site indicated.

polymerase such as DH1(DE3) and BL21(DE3). The bacteriophage DE3 present in these strains is a lambda derivative that has been engineered to contain the gene for the T7 RNA polymerase under the control of the lacUV5 promoter. It also carries the lacI gene so that the T7 polymerase gene, normally switched off, can be activated by the addition of the inducer molecule IPTG. The pET vectors have been used successfully to express many proteins including the poliovirus 3C protease (Baum *et al.*,1991). Under favourable circumstances the resources of the cell are devoted almost entirely to the production of target proteins so that they

can constitute the majority of the total cell protein.

Baculovirus Expression Systems.

One of the problems with prokaryotic expression systems, including those described above, is that they are not able to carry out the post-translational modifications required by many eukaryotic proteins. In addition proteins are frequently incorrectly folded to an extent that makes them biologically inactive. These problems have led to the increasing use of recombinant eukaryotic expression systems to express cloned genes accurately. One such system involves the infection of insect cells, in tissue culture, with recombinant baculoviruses and has been used to produce high yields of biologically active proteins, including interleukin-2, *c-myc* and influenza haemagglutinin (reviewed in: Miller,1988).

Several features of the baculovirus infectious cycle make them amenable to genetic manipulation. The virus usually used is Autographa californica nuclear polyhedrosis virus (AcNPV) which has an 128-kDa double stranded, circular genome and a biphasic infectious cycle. The first phase involves the formation of nuclear capsids that bud through the cellular membrane to produce extracellular virus. The second phase involves high level expression of a 32-kDa protein, termed polyhedrin, that is responsible for packaging mature virions within the nucleus in large occlusion bodies. The occluded viruses are released after cell death and are only required for lateral transmission of the virus. The significance of this is that a gene switched on late in infection produces a protein in large amounts that is not required for infection in tissue culture. This means that the polyhedrin gene can be replaced with a foreign gene so that the latter is placed under the control of the strong polyhedrin promoter. This is achieved by cloning the foreign genes into specially designed transfer vectors such as pAcRP23 (Possee and Howard,1987) and pVL1393 (Invitrogen Corporation). The maps of these vectors are given in **fig.3.3** where it can be seen that they both contain cDNA from the AcNPV genome, encompassing the polyhedrin gene, in which unique restriction sites have been engineered downstream from the polyhedrin promoter. In the case of pVL1393, the promoter is followed by a mutated initiation codon and the bases coding for the N-terminal amino acids of the polyhedrin and then the restriction sites for insertion of foreign cDNAs. This works on the principle that the transcription enzymes will find the polyhedrin promoter

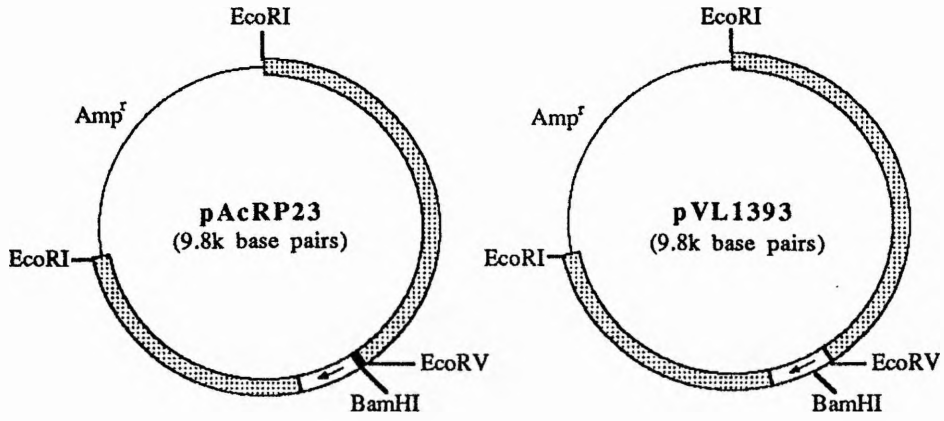


Fig.3.3 Maps of the pAcRP23 and pVL1393 vectors
The polyhedrin promoter is indicated by the arrow and the shaded boxes represent the flanking baculovirus DNA sequences. The location of the BamHI cloning site is marked in each case.

untampered with and will move along to the initiation ATG of the foreign gene, located only a few bases downstream.

The foreign gene contained within the recombinant transfer vector is introduced precisely into the AcNPV genome by homologous recombination. The recombinant virus can then be purified by plaque purification as described by Summers and Smith (1987). Since the frequency of such recombination events is typically 0.1-1% the identification of recombinant virus against a background of parental virus is a time consuming step. Possee and Howard (1987) engineered the AcNPV DNA so that it contained the lacZ gene coding for β-galactosidase in place of the polyhedrin to give AcRP23-lacZ. Recently Kitts *et al.*(1990) noticed that this DNA has a unique restriction site, Bsu36I in the lacZ gene. The recombinant baculovirus DNA can, therefore, be linearised by cutting with Bsu36I and the linear form has been shown to be recombinogenic; but non-infectious. Kitts *et al.* (1991) showed that recombination of the linearised AcRP23-lacZ with an appropriate transfer vector produces a circular genome, competent for replication. The system has the additional advantage that the virus expressing the foreign gene can be readily distinguished from AcRP23-lacZ in that it does not produce β-galactosidase.

It is hoped that expression of the 23-kDa protein in higher eukaryotic cells will reveal whether or not it is phosphorylated or post-translationally modified in some other way.

METHODS

1 BASIC CLONING METHODS.

E.Coli and baculovirus expression systems were utilised to express the Ad2 23-kDa protein. This involved the insertion of the L3 23-kDa gene into 4 vectors, namely pET3c (Studier *et al.*,1990), pGEX-2T (Smith and Johnson,1988), pAcRP23 (Possee and Howard,1987) and pVL1393 (Invitrogen Corporation). A number of standard cloning techniques were used in each case and these will be outlined first before the specific details of each construct are described. Many of the techniques used were based on methods from the Cloning Manual (Maniatis *et al.*,1982).

Restriction enzymes and DNA modifying enzymes were all obtained from Amersham, Promega or Gibco-BRL unless otherwise stated. All solutions were sterilised by autoclaving or sterile filtering. 10x enzyme buffers were stored in 0.5ml aliquots at -20°C.

1.1 Plasmid Preparation.

Bacteria were grown in Luria broth (LB) prepared by dissolving 10g of bacto-tryptone (Difco) and 5g of yeast extract (Difco) in 1 litre of 10mM NaCl and autoclaving the solution. All plasmids used in this project contained the ampicillin resistant gene and so ampicillin was added to the Luria broth at 0.1 mg/ml (LB/Amp).

5ml of LB/Amp were inoculated with 5µl of an *E.Coli* strain transformed with the required plasmid. The culture was grown for 10-18 hours at 37°C. 2ml of the stationary phase culture were inoculated into 1 litre of LB/Amp and incubated with shaking for 16-20 hours at 37°C. Cells were harvested by centrifugation at 10000rpm (12000g) for 5 minutes in a Beckman J2-21 centrifuge using the JA14 rotor. The pellet was resuspended in 18ml of 50mM glucose, 25mM Tris/HCl, 10mM EDTA, pH8 and 2ml of freshly prepared 20mg/ml lysozyme in 0.2M Tris/HCl pH8. A 5 minute incubation at room temperature was followed by the addition of 40ml of fresh 0.2M NaOH, 1% SDS and a second 5 minute incubation; this time on ice. 20ml of 5M potassium acetate, pH4.8 (60ml of 5M potassium acetate, 11.5ml of glacial acetic acid and 28.5ml of H_2O) were added and the mixture was left on ice for a further 15 minutes. The extract was centrifuged for 5 minutes as before (4°C) and the supernatant

was filtered through 2 layers of cheesecloth. Plasmid DNA was precipitated with 0.6 volumes of isopropanol for 15 minutes at room temperature and the suspension centrifuged at 12000g using the JA17 rotor in the Beckman J2-21 centrifuge for 5 minutes. The pellet was resuspended in 20ml of 70% ethanol, and centrifuged as just described, before the supernatant was discarded and the dry pellet was redissolved in 6ml 10mM Tris, 1mM EDTA, pH8. Caesium chloride (7.2g) and 100µl of 10mg/ml ethidium bromide were dissolved in the DNA suspension which was then allowed to warm to room temperature before being centrifuged at high speed in a microcentrifuge for 5 minutes. The supernatant was pipetted into 4, bell-top quickseal tubes (Beckman) which were filled, balanced and sealed according to the manufacturers instructions. Samples were centrifuged at 100000g for 12-18 hours at 20°C in a Beckman UL-55 centrifuge. The plasmid band was removed and extracted with 5x1 volume of caesium chloride saturated butan-2-ol to remove ethidium bromide. Two volumes of H_2O were added to dilute out CsCl, followed by 0.6 volumes of isopropanol. The plasmid DNA was left to precipitate for 20 minutes at room temperature before being pelleted by centrifugation for 5 minutes at high speed in a microcentrifuge. The DNA pellet was dissolved in 400µl of H_2O and was precipitated for 10 minutes at room temperature after adding 40µl 3M sodium acetate and 1ml of ethanol. The suspension was centrifuged as before, the pellet washed with 70% ethanol, dried under vacuo and finally dissolved in 500µl of H_2O and stored in aliquots at -70°C. The concentration of the plasmid was determined by measuring its absorbance at 260nm and analysing it by agarose gel electrophoresis (see 1.2, below).

1.2 Analysis of DNA.

(a) *Concentration of DNA.*

The concentrations of DNA samples were determined by measuring the absorbance of the sample or an appropriate dilution at 260 nm and applying the formula:

(i) For single stranded DNA 1 absorbance unit indicates a concentration of 50µg/ml.

(ii) For double stranded DNA 1 absorbance unit indicates a concentration of 40µg/ml.

(iii) For oligonucleotides 1 absorbance unit indicates a concentration of 20µg/ml.

The presence of any protein contaminants was determined by measuring the absorbance

at 260 nm and 280 nm. The ratio between the readings at 260nm and 280nm should be 1.8 for pure DNA solution. Values less than this indicate that there are protein or phenol contaminants in the preparation.

(b) *Agarose Gel Electophoresis.*

DNA samples were separated by agarose gel electrophoresis in the GNA100 submersible rig (Pharmacia) according to the manufacturers instructions. The preferred buffer system was 60mM Tris/acetic acid, 3mM EDTA, pH8 (TAE) and the agarose (Type II, medium EEO) was from Sigma. 0.3-2% agarose in TAE was used depending on the size of DNA fragments to be separated. 0.7% agarose gels were suitable for most separations. Depending on the number of DNA fragments in a particular sample and the size of gel used the amount of DNA loaded per track was 0.05-1μg. Samples were mixed with 1/6 volume of 0.25% xylene cyanol, 0.25% bromophenol blue, 30% glycerol (DNA loading dye) and applied directly to the gels. Phage lambda DNA cut with the restriction enzyme HindIII (Boehringer) was used as the size marker on gels. Gels were run for 1-2 hours at a constant current of 100mA until the dye front had migrated 3/4 the length of the gel and then were stained for 5 minutes in 0.5μg/ml ethidium bromide in TAE. DNA bands were examined and photographed on a UV transilluminator.

1.3 Purification of DNA.

(a) *Phenol Chloroform Extraction.*

DNA solutions were mixed with an equal volume of 1:1 phenol/chloroform (redistilled phenol equilibrated with 100mM Tris/HCl, 1mM EDTA, pH8) and centrifuged for 10 minutes at high speed in a microcentrifuge. The aqueous layer was stored on ice while the organic phase was re-extracted with an equal volume of 10mM Tris/HCl, 1mM EDTA, pH8 as before. The aqueous phases were pooled and extracted with 2x1 volume of chloroform with the centrifugation step as before. DNA was precipitated from the aqueous phase for 5 minutes at room temperature following the addition of 1/10 volume of 3M sodium acetate and 2.5 volumes of ethanol. The DNA was pelleted by centrifugation for 5 minutes in a microcentrifuge and then dried *in vacuo* and dissolved in the required amount of H_2O.

(b) *Extraction of DNA from Agarose Gels.*

Double stranded DNA fragments (>100 base pairs) were purified from agarose gels using the Geneclean II Kit (Stratagene) according to the manufacturers instructions.

DNA was separated on the appropriate % of agarose gel as described in method 1.2(b). The gels were viewed for the minimum time in UV light and the desired DNA band cut out and melted in 2.5 volumes of sodium iodide at 55°C for 3-5 minutes.

5µl of "glass milk" were added and the suspension was left at room temperature for 20 minutes with occasional shaking before being centrifuged for 1 minute in a microcentrifuge. The pellet was washed with 3x500µl of "new wash", air dried and resuspended in 40µl of H_2O. The suspension was incubated for 3 minutes at 55°C and then centrifuged for 5 minutes in a microcentrifuge. The DNA was in the supernatant.

(c) *Desalting on G-50 Sephadex Columns.*

G-50 Sephadex (Sigma) was swollen in 10mM Tris/HCl, 1mM EDTA, pH8 and packed into a 1ml syringe, plugged with glass wool. The resin was equilibrated with 10ml of the same buffer and then spun for 5 minutes at 2000g (IEC CENTRA-3R centrifuge) to remove the buffer. 50-200µl of DNA solution were loaded onto the column which was centrifuged as before and the desalted DNA was collected into an eppendorff tube.

1.4 Polymerase Chain Reaction (PCR).

(a) *Preparation of Oligonucleotide Primers.*

Oligonucleotide primers, synthesised on an Applied Biosystems 381A synthesiser, were resuspended in 200µl of H_20, 20µl 3M sodium acetate and 2.2µl 1M magnesium acetate and then precipitated with 3 volumes of cold ethanol. Samples were centrifuged for 5 minutes in a microcentrifuge, the precipitation procedure was repeated and the final pellet was air dried and dissolved in 400µl of H_20. The oligonucleotide concentration was determined by measuring its absorbance at 260nm (the mass of 1 base is 330 daltons) and a 100µM solution was prepared.

(b) *Preparation of Adenovirus DNA.*

Ad2, twice purified by caesium chloride centrifugation (chapter 1, method 3) was dialysed against 10mM Tris/HCl, 1mM EDTA, 0.6% SDS, pH7.5 and then proteinase K was

added to 100μg/ml. The sample was incubated for 2 hours at 50°C and the viral DNA purified by phenol/chloroform extraction (method 1.3a). The resulting DNA pellet was resuspended in 1mM Tris/HCl, 0.1mM EDTA pH 8.

(c) *Materials.*

Taq polymerase was from Northumbria Biologicals Limited and the 4 deoxynucleotide triphosphates (dATP, dTTP, dGTP and dCTP) were from Pharmacia. A stock solution of the nucleotide triphosphates (4dNTPs) was prepared so that the concentration of each was 1mM. The 10x PCR buffer comprised 0.1M Tris pH8, 15mM $MgCl_2$, 0.5% Tween 20 and 0.5% NP40.

(d) *Setting up PCR Reactions.*

Reactions were set up in triplicate in siliconised, sterile eppendorff tubes as follows:

H_2O	67μl
10x PCR	10μl
4dNTPs	10μl
Ad2 DNA (0.25μg)	10μl
Oligonucleotide 1 (100μM)	1μl
Oligonucleotide 2 (100μM)	1μl
Taq polymerase	1μl

Tubes were overlayed with 50μl of paraffin oil and placed in a programmable heating block. The program used for all the PCR reactions in this project was:

STEP1 (melting)	94°C for 1.5 minutes
STEP 2 (annealing)	60°C for 1.5 minutes
STEP 3 (elongation)	72°C for 2 minutes

This cycle was repeated 30 times and was followed by a final step of 7 minutes at 72°C.

(e) *Analysis of PCR Products.*

5μl of DNA were pipetted from under the paraffin layer and analysed by agarose gel electrophoresis (method 1.2b). The PCR products were separated from primers and free nucleotides by either purifying the DNA from gels or passing the mixture through a 1ml Sephadex G-50 column.

1.5 DNA Modifying Enzymes.

(a) *Restriction Endonuclease Digestions.*

Enzymes were supplied along with an appropriate 10x buffer. A typical reaction was for 2 hours at 37°C and comprised 0.1-5µg of DNA, 1-5µl of 10x buffer, 1-5µl of 1mg/ml acetylated BSA and 1-25 units of restriction endonuclease with the reaction volume being made up to 10-50µl with H_2O.

If DNA was to be digested by 2 restriction enzymes that were active in the same buffer then the digestions were carried out simultaneously. This not being the case, then the enzymes were added sequentially with an intermediate desalting step on a G-50 Sephadex spin column.

(b) *Phosphatase Treatment of DNA.*

Calf intestinal alkaline phosphatase (CIAP) was obtained from Amersham and the 10x CIAP used was 0.5M Tris/HCl, pH9, 10mM $MgCl_2$, 1mM $ZnCl_2$, 10mM spermidine. Typically 5µg of linearised plasmid, 5µl of 10xCIAP and 0.5 units of CIAP in a final volume of 50µl were incubated for 60 minutes at 37°C. The reaction was stopped by the addition of 1µl of 0.5M EDTA and heating for 20 minutes at 65°C. CIAP was removed by phenol/chloroform extraction (method 1.3a).

(c) *Ligation Reactions.*

Insertion of genes into linearised plasmid DNA was carried out using T4 DNA ligase. The 10x ligase buffer used was 0.4M Tris/HCl pH7.5, 0.1M $MgCl_2$, 0.1M DTT, 0.5mg/ml acetylated BSA. For plasmids cut with a single restriction enzyme 5' phosphates were first removed by CIAP to prevent self ligation. The ligation reaction was for 16 hours at 13°C and comprised 1µg cut plasmid, 0.5µg DNA insert, 1µl 10x ligase, 1µl 10mM ATP and 2 units of T4 DNA ligase. The reaction was stopped by storing at -20°C.

(d) *Polymerase Reactions.*

The Klenow fragment of *E.Coli* DNA polymerase I was used to fill in recessed 3' ends of double stranded DNA. The 10x Klenow buffer used was 0.5M Tris/HCl, 0.1M magnesium sulphate, 1mM DTT, 0.5mg/ml acetylated BSA, pH7.2.

T4 DNA polymerase was used to fill in double stranded DNA with protruding 3' ends.

The 10x buffer for this enzyme was 0.33M Tris/Acetate pH7.9, 0.66M potassium acetate, 0.1M magnesium acetate, 5mM DTT, 1mg/ml acetylated BSA. A solution was made 1mM with respect to each of the 4 deoxynucleotide triphosphates (dATP, dTTP, dGTP and dCTP) and stored at -20°C. One unit of the appropriate polymerase enzyme was added to 2µg of DNA, 2µl of 10x buffer and 2µl of 4dNTPs in a final volume of 20µl. Incubations were for 30 minutes to 1 hour at room temperature for Klenow polymerase or 30 minutes at 37°C in the case of T4 polymerase. The Klenow reaction was stopped by the addition of 1µl of 0.5M EDTA and the T4 polymerase reaction by heating at 70°C for 5 minutes. Both enzymes could be removed by phenol/chloroform extraction (method 1.3a).

(e) *T4 Polynucleotide Kinase.*

T4 polynucleotide kinase (Amersham) was required to phosphorylate dephosphorylated 5' ends of DNA and the 5x kinase buffer used was 0.5M Tris/HCl pH7.6, 0.1M $MgCl_2$, 50mM DTT, 1mM spermidine, 1mM EDTA.

Dephosphorylated DNA (1-50pmoles of 5' ends) was added to 10-20 units of T4 kinase, 5µl of 10mM ATP, 10µl of 5x kinase and made up to 50µl with H_2O. The reaction was allowed to proceed at 37°C and stopped by adding 2µl of 0.5M EDTA prior to extraction with phenol/chloroform (method 1.3a).

1.6 Transformation of *E.Coli* with Plasmid DNA.

The strain of *E.Coli* normally transformed was JM101. Other strains of *E.Coli* used were DH1(DE3), DH5 and BL21(DE3). The procedure followed for transformations was the same for all the strains. The selected bacterial strain was inoculated into 5ml of LB and incubated at 37°C overnight with shaking. 500µl of the stationary phase culture were inoculated into 20ml of LB and grown at 37°C with shaking until they reached 0.15 absorbance units at 600nm (1-2 hours). Cells were placed on ice and 1.5ml aliquots centrifuged in a microcentrifuge for 2 minutes. Each cell pellet was resuspended in 500µl of cold 10mM MOPS, 10mM rubidium chloride, pH7 (Buffer A), centrifuged as before and the pellet resuspended in 500µl of cold 10mM MOPS, 50mM calcium chloride, 10mM rubidium chloride, pH6 (buffer B). Cells were left on ice for 1 hour, centrifuged as before and the pellet

suspended in 200µl of buffer B. 3µl of DMSO (frozen in aliquots at -20°C and not reused) were added to each tube followed by 0.2µg of DNA and the cells were incubated on ice for 30 minutes. 4ml of LB were added to the cells which were incubated at 37°C with shaking. After 2 hours cells were centrifuged at 2000g and the pellet was resuspended in 1ml of LB. 50-200µl were plated out on LB/Amp plates (see below) with an L-shaped glass rod. Plates were incubated, agar side up, at 37°C overnight and then stored at 4°C until required.

Typically, control transformations were set up (a) no DNA added, (b) uncut plasmid and (c) cut, ligated plasmid (no insert). The control 'no DNA' transformed cells were plated on LB plates and LB/Amp plates.

Preparation of LB and LB/Amp Plates.

15g of Bacto-agar (Difco) were added to 1 litre of LB and autoclaved. The agar was cooled to 50°C and ampicillin was added to 100µg/ml where required. 30ml of agar were poured into 85mm petri dishes and any bubbles were removed by flaming with a bunsen burner. The plates were allowed to set and could be stored for up to 1 week at 4°C. Before use the plates were dried for 15 minutes in a UV cabinet at 45°C.

1.7 Mini Preparations of Plasmid DNA from Single Colonies.

Using a sterile toothpick individual colonies were picked from LB/Amp plates and grown in 3ml of LB/Amp overnight at 37°C. 1ml of each overnight culture was centrifuged in a microcentrifuge for 30 seconds and the pellet resuspended in 100µl of H_2O. An equal volume of 2x lysis buffer (0.1M NaOH, 2% SDS, 10mM EDTA) was added, the samples were vortexed and boiled for 2 minutes. 100µl of 1M magnesium acetate were added, the samples were vortexed and left to stand on ice for 2 minutes prior to being centrifuged for 30 seconds as above. The supernatant was NOT removed and 50µl of 5M potassium acetate were added, the samples were vortexed and left on ice for 2 minutes and centrifuged as before. The supernatants were transferred to sterile tubes containing 0.6ml of 95% ethanol, mixed and left on ice for 2 minutes. Samples were centrifuged for 2 minutes in a microcentrifuge and the resulting pellets were washed with 500µl of 70% ethanol and centrifuged as before. The pellets were dried in *vacuo* and dissolved in 40µl of 0.1mM EDTA, 10mM Tris/HCl pH 8. One unit of RNase (Boehringer) was added to each sample and the DNA was analysed by

restriction endonuclease digestion and agarose gel electrophoresis. Typically 5µl of DNA was used for each restriction digest.

1.8 Replica Plating of Transformed *E.Coli* Colonies.

Identical 50 square grids were drawn on an 82mm Hybond-N filter (Amersham) and the base of an 85mm petri dish. The latter was taped to the base of an LB/Amp plate keeping it in the same orientation as the filter, which was placed directly onto the agar of a second LB/Amp plate. Using sterile toothpicks individual colonies were picked and streaked onto a grid on the filter and the corresponding grid on the master plate. Both plates were incubated, agar side up, at 37°C overnight and stored at 4°C until required.

(a) *Preparation of Filters.*

The nylon filters were lifted from the agar using forceps and placed sequentially (colony side up) on filter paper soaked in (1) 0.1M HCl for 1 minute, (2) 0.5M NaOH for 15 minutes, (3) 1M Tris/HCl, pH 7.5 for 5 minutes and (4) 1.2M NaCl for 15 minutes. The filter was then air dried for 1 hour and placed colony side down on a UV transilluminator for 2 minutes.

(c) *Labelling of DNA Probes.*

DNA probes were labelled by random priming using the Klenow polymerase. The 5x oligonucleotide labelling buffer (OLB) used was 0.25M Tris/HCl pH8.1, 10mM DTT, 25mM $MgCl_2$, 0.2M KCl and the unlabelled dNTP stock was 0.4mM with respect to dCTP, dGTP and dTTP. 0.2µg of dsDNA in 3.75µl of H_20 were denatured by boiling for 3 minutes and then placing on ice. 1.5µl OLB, 0.25µl 1mg/ml acetylated BSA, 0.75µl unlabelled dNTPs, 10µCi [α-^{32}P] dATP and 1 unit of Klenow were added and the 7.5µl reaction mixture was left for 2.5 hours at room temperature.

0.5ml of TEN (10mM Tris/HCl pH7.5, 1mM EDTA pH8, 100mM NaCl) and 100µl of 1M NaOH were added to the labelled DNA and left at room temperature for 10 minutes. The solution was then neutralised by the addition of 100µl of 1M HCl.

(d) *Hybridisation.*

The solutions required for the hybridisation reaction were 3M NaCl, 0.3M sodium citrate (20xSSC) and 1% BSA, 1% Ficoll, 1% polyvinyl pyrrolidone, 0.02% azide

(50xDenhardts). The latter was stored at 4°C. Single stranded DNA was prepared by sonicating salmon sperm DNA or calf thymus DNA at a concentration of 2mg/ml in H_2O and storing in 1ml aliqouts at -20°C. Immediately prior to use the DNA was boiled for 3 minutes and then placed on ice. The prehybridisation and hybridisation solutions were made up as follows:

Pre-hybridisation buffer
1.5ml 20xSSC
3ml H_2O
0.5ml 50xDenhardts
25µl 20% SDS
50µl denatured DNA

Hybridisation buffer
3ml 20xSSC
2ml 50xDenhardts
2.8ml H_2O
20µl 0.5M EDTA pH8
200µl 1M Tris/HCl pH7.5
250µl denatured DNA
500µl 10% SDS

The nitrocellulose filter (prepared above) was placed in a sealed plastic bag with 5ml of prehybridisation buffer and incubated in a shaking water bath at 65°C for 3 hours. The filter was transferred to a fresh bag containing 8.8ml of hybridisation buffer, 0.1% sodium tetrapyrophosphate. The ^{32}P labelled probe was added, the bag sealed and the filters hybridised at 65°C overnight. Filters were washed for 3x30 minutes in 500ml of 2x SSC, 0.25% SDS at 65°C to remove free probe. They were then blotted dry and exposed to Fuji X-ray film for 1-5 hours.

1.9 Storage of Transformed *E.Coli* strains.

Transformed bacteria were grown in LB/Amp overnight and the stationary phase cultures were stored at -70°C in 15% glycerol.

2 CLONING AND EXPRESSION OF 23-KDA PROTEIN IN *E.COLI*

The 23-kDa protein was expressed as a fusion protein with glutathione transferase using the vector pGEX-2T and as an individual entity using the pET3c vector. In addition to expressing the intact protein 3 deletion mutants were also expressed.

2.1 Cloning and Expression of the 23-kDa Protein using pGEX-2T.

Three plasmids were constructed by incorporating all or part of the gene encoding the Ad2 23-kDa putative protease into the vector pGEX-2T as follows:

pGEX23K(entire 23-kDa sequence)

pGEX23KΔN (codons for N-terminal 18 amino acids deleted)

pGEX23KΔC (codons for C-terminal 23 amino acids deleted)

In each case the cDNA insert encoding the desired region of the 23-kDa protein was copied from Ad2 DNA using PCR. The oligonucleotide primers used to do this were:

(1) 5'-CCCC**GGATCC**AGTGAGCAGGAACTGAAAGCCATTGTC-3'

(2) 5'-GCGC**GGATCC**TGGGGGGTAAATAATCACCCGAGAGTGTACAA-3

(3) 5'-GCGC**GGATCC**CCATATTTTTTGGGCACCTATGACAAGCGC-3'

(4) 5'-GCGC**GAATTC**TTAGTGGCGCTCCAGGAAGCTGTAGAG-3'

Each oligonucleotide primer was designed to have 4 bases at its 5' end followed by the desired restriction site, this being a BamHI site in primers (1) - (3) and an EcoRI site in primer (4). The sites recognised by these enzymes are highlighted above. Primers (1) and (3) were taken from the +strand of Ad2 and primers (2) and (4) from the -strand. The inserts for the 4 plasmids were copied from Ad2 DNA by PCR according to method 1.4 using the following pairs of oligonucleotide primers:

pGEX23K - (1) and (2)

pGEX23KΔN - (3) and (2)

pGEX23KΔC - (1) and (4)

The PCR products were purified from agarose gels and digested with either BamHI or BamHI and EcoRI as appropriate. The products were then purified from agarose gels and ligated to pGEX-2T x BamHI or pGEX-2T x BamHI, EcoRI. Ligation products were transformed into JM101 cells and plated onto LB/Amp plates (methods 1.5 and 1.6).

Screening of Colonies

For each construct 6 colonies were selected at random and grown overnight in 3ml of LB/Amp at 37°C. 50µl of each densely grown culture were inoculated into 5ml of LB/Amp and bacteria were grown at 37°C with shaking until they reached 0.6 absorbance units at 650nm. To 1ml of each culture 4µl of 0.1M β-D-isopropyl-thiogalactopyranoside (IPTG) were added and the cells were induced for 3 hours at 37°C with shaking. Cultures were harvested by centrifugation for 1 minute in a microcentrifuge, the cell pellets were washed with 1ml of cold 50mM Tris/HCl, 1mM EDTA, pH8 and resuspended in 200µl of the same buffer made

0.5% with respect to NP-40. Samples were sonicated for 4 x15 seconds on ice and 15µl of each were analysed by SDS-PAGE and Western Blotting (chapter 2, method 2).

2.2 Cloning and Expression of the 23-kDa Ad2 gene using pET3c.

This work was carried out in the Brookhaven National Laboratory under the supervision of Dr C.W. Anderson.

The 23-kDa putative Ad2 protease had previously been expressed in pET3 and the construction of the resulting plasmid, pT7AD23K6 is described by Anderson (1990).

(a) *Construction of pT7AD23K12*

The cDNA fragment encoding the 23-kDa protein was prepared by digestion of pT7AD23K6 with the restriction endonucleases NdeI and PstI. The resulting DNA fragment was purified from an agarose gel and ligated to pET3c that had been digested with the same restriction enzymes. The ligation product was then transformed into DH1(DE3) cells and plated onto LB/Amp plates.

(b) *Construction of pT7AD23K10 (N-terminal deletion mutant)*

A plasmid pT7AD23K10 was constructed, encoding the 23-kDa gene with a 27 base pair deletion covering the coding region for amino acids 5-13 of the protein. This was achieved by copying the DNA from the linearised plasmid, pT7AD23K6 x EcoRV, using the oligonucleotide primers:

(5) 5'-CCCCAT**ATG**GGCTCCTCAGATCTTGGTTGTGGGCC-3'

(6) 5'-GGCCCTTTCGTCTTCAAG-3'

Primer (5) matches with the +strand across the NdeI site (highlighted) at the initiation codon (underlined) through to the TCA codon for Ser_4 of the 23-kDa protein, then with the bases coding for Asp_{14} to Pro_{19} of the same protein. Primer (6) matches with the opposite strand at a site just distal to the EcoRI site of pET3.

The PCR product was purified from an agarose gel, digested with the restriction endonucleases NdeI and KpnI and then purified a second time from an agarose gel. The plasmid pT7AD23K6 was digested with the same restriction enzymes and the cut vector was separated from the insert by agarose gel electophoresis. The cut vector and PCR product were ligated together and then transformed into DH1(DE3) cells and plated onto LB/Amp plates.

(c) *Screening Colonies*

For each construct 6 colonies were selected and the nature of their transformants assessed by carrying out mini preparations of their plasmid DNA followed by selected restriction enzyme digestions (EcoRI, NdeI,BglII/NruI for pT7AD23K12 and EcoRI,NdeI, KpnI/NdeI for pT7AD23K10).

(d) *Expression of Cloned Proteins*

Correctly constructed plasmids were transformed into BL21(DE3) cells in order to express the 23-kDa protein and its deletion mutant and the resulting strains were grown up in LB/Amp overnight at 37°C. 5µl of the stationary phase cultures were inoculated into 100ml of LB/Amp or M9/Amp. The M9 minimal medium was prepared by mixing 5ml of 20x M9 (10g NH_4Cl, 30g KH_2PO_4 and 60g Na_2HPO_4 made up to 500ml with H_2O, filtered through a 0.2µ nylon filter and autoclaved), 2ml of 20% glucose (sterile filtered) and 0.1ml of 1M $MgSO_4$ (sterile filtered) and making up to 100ml with sterile H_2O.

The 100ml cultures were grown overnight with shaking at 30°C and then transferred to a 37°C shaker. The absorbance of the cultures at 600nm was monitored and on reaching 0.4 units the expression of the proteins was induced by the addition of IPTG to 0.4µM. At 0, 1, 2 and 3 hours after the addition of the IPTG, 1ml samples were withdrawn from the cultures and labelled for 5 minutes at 37°C with 20µCi of ^{35}S methionine. The bacteria were sedimented by centrifugation for 5 minutes in a microcentrifuge and washed x2 with PBS, pH6.2. They were then disrupted by boiling in 100µl of SDS gel loading buffer. 15µl of each sample was analysed by SDS-PAGE and autoradiography.

Three hours post induction the cells were harvested by centrifugation for 5 minutes at 10000g and then washed with 50mM Tris/HCl, pH8, 5mM EDTA, 4% glycerol. The pellet was resuspended in 5ml of the same buffer and incubated with 100µg/ml lysozme for 15 minutes at 37°C. The lysed cells were frozen and thawed 3 times in a dry ice /ethanol bath and then incubated for a further 15 minutes in the presence of 10mM $MgCl_2$ and 100µg/ml DNaseI. Samples were centrifuged at 10000g for 20 minutes at 4°C and the supernatants were assessed for protease activity using the ^{35}S Ad2ts1 assay (chapter 2, method 1.1). The pellets were

washed x2 in 50mM Tris/HCl pH8, 5mM EDTA, 4% glycerol and then solubilised by heating to 90°C in 1ml SDS-PAGE loading buffer. The proteins were then purified by gel filtration and reverse phase chromatography as described by Anderson (1990). The composition of the proteins were verified by tryptic peptide and protein sequence analysis by R.Greene (Anderson, 1990).

2.3 Purification of 23-kDa Protein from *E.Coli*.

10µl of pT7AD23K12 transformed BL21(DE3) cells were inoculated into 10ml of LB/Amp and grown overnight at 37°C. 5ml of the densely grown culture were added to 1 litre of M9/Amp and grown with shaking at 37°C until the absorbance of the culture at 600nm reached 0.4-0.6. The culture was made 0.4µM with respect to IPTG and induced for 4-6 hours at 37°C with shaking. The cells were harvested by centrifugation for 5 minutes at 12000g in a JA14 rotor using a J2-21 Beckman centrifuge and then washed with 200ml of 50mM Tris/HCl, 5mM EDTA, 1mM β-mercaptoethanol, 4% glycerol pH8. The pellet was then resuspended in 50ml of the same buffer and stored at -70°C until required.

0.5ml of a 1mg/ml lysozyme solution were added to the 50ml cell suspension and incubated for 15 minutes at 37°C before freeze/thawing x3 using liquid nitrogen and a 37°C water bath. 0.5ml of 1M $MgSO_4$ and 0.25ml of 2mg/ml DNaseI were then added to the lysed cells, mixed and incubated for a further 15 minutes at 37°C. The sample was centrifuged for 20 minutes at 12000g (4°C) using a JA17 rotor in a J2-21 Beckman centrifuge and the resulting supernatant was dialysed for 2 hours against 3x2 litres of 50mM Tris/HCl pH8 at 4°C. It was then filtered through a 0.2µ nylon filter and loaded onto a 25ml DEAE-Sepharose-CL6B column equilibrated with 50mM Tris/HCl pH8. The flow rate was 1ml/minute, the detection wavelength 280nm and the full scale deflection was 2 absorbance units. 4ml fractions were collected for 90 minutes or until the absorbance at 280nm had returned to the baseline. The column was then washed with 200ml of 1M NaCl and stored in 0.02% azide. The flow-through fractions contained the 23-kDa protein and were pooled and solid ammonium sulphate was added slowly to the pooled fractions in order to make them 1.2M with respect to the salt. The sample was then filtered through a 0.2µ nylon filter and

loaded onto an FPLC phenyl Superose HR5/5 column equilibrated with 50mM Tris/HCl, 1.2M ammonium sulphate pH8. The flow rate was 0.5ml/minute, the full scale deflection 1 absorbance unit and the detection wavelength 280nm. Once the sample was loaded the following gradient was applied to the column:

time (minutes)	%B
0	30
10	30
15	45
30	45
45	90
45.1	100
50	100
50.1	30
60	30

The buffers used were (A) 50mM Tris/HCl, 1.7M ammonium sulphate pH8 and (B) 50mM Tris/HCl pH8. Fractions were collected and 5µl of each were analysed by SDS-PAGE and Western blotting as described in chapter 2, method 2.

3 CLONING AND EXPRESSION OF AD2 23-kDa PROTEIN USING RECOMBINANT BACULOVIRUSES.

The 23-kDa putative protease gene was cloned into two different baculovirus vectors in order to compare expression levels.

3.1 Cloning of the 23-kDa Protein into Baculovirus Vectors

(a) *Cloning of 23-kDa protein into pAcRP23*

The DNA fragment encoding the 23-kDa protein was obtained by digesting the plasmid pT7AD23K12 with NdeI and EcoRV, treating it with Klenow polymerase and purifying it from an agarose gel. The vector pAcRP23 (Possee and Howard,1987) was linearised with BamHI, treated with Klenow polymerase and calf intestinal alkaline phosphatase (CIAP), phenol/chloroform extracted and then ligated to the purified NdeI-EcoRV fragment. The ligation product was transformed into JM101 cells and colonies were screened as described in method 3.2 (below).

(b) *Cloning of the 23-kDa protein into pVL1393*

A fragment representing the entire coding region for the Ad2 putative protease was copied from Ad2 DNA by PCR using the following oligonucleotide primers:

(7) 5'-GCGCGGATCCTATAAATATGGGCTCCAAGTGAGCAGGAACTGAA
AGCCAT-3'

(2) 5'-GCGCGGATCCTGGGGGGTAAATAATCACCCGAGAGTGTACAA-3'

Both primers include a recognition site for BamHI close to their 5' end and primer (7) matches exactly with the +strand from the initiation codon of the 23-kDa protein (underlined). The sequence TATAAAT is included upstream from the initiation codon since this is reported to give increased levels of transcription (Possee and Howard,1987).

The resulting PCR product was gel-purified, digested with BamHI and then gel purified for a second time. The vector pVL1393 (Invitrogen) was also digested with BamHI before treatment with CIAP and phenol/chloroform extraction. The PCR insert and vector were then ligated together and the ligation product was transformed into JM101 cells and the colonies were screened as described in 3.2.

3.2 Screening of Colonies.

For each of the constructs one hundred colonies were selected at random and replica plated as described in method 1.8 Filters were hybridised with a specific ^{32}P labelled probe, prepared by random priming of the BamHI digested PCR product prepared in 3.1. Positive colonies were identified by autoradiography and 6 positives were selected for each construct. The orientations of the protease gene were established by BglII, EcoRI digestions of the mini-plasmid preparations from each colony. Large scale plasmid preparations of the two correctly constructed vectors, termed pAc23K and pVL23K respectively, were carried out ready for transfections with baculovirus DNA.

3.3 Insect Cell Culture.

Spodoptera frugiperda cells (sf9) were maintained in suspension and monolayer cultures at 28°C in TC100 medium (Gibco-BRL) supplemented with 10% fetal calf serum (FCS) and 50μg/ml gentamycin. For routine maintenance cells were passaged every 2-3 days.

(a) *Suspension Cultures*

For suspension cultures when the cell count reached $2x10^6$ cells/ml, 80% of the culture was discarded and replaced with the same volume of TC100/10% FCS.

(b) *Monolayer Cultures*

Monolayer cultures were maintained in $75cm^3$ plastic bottles and on reaching confluence the medium was decanted and the cells were resuspended in 10ml of TC100/10% FCS. 2ml of the cell suspension were transferred to a second $75cm^3$ bottle which was made up to 25ml with TC100/10% FCS.

(c) *Freezing of sf9 Cells*

When the cell count of a suspension culture reached 10^6 cells per ml, 10ml aliquots were centrifuged for 5 minutes at 1000g (4°C) and the cell pellets were gently resuspended in 1ml of 90% FCS, 10% DMSO (10^7 cells/ml). The 1ml aliquots were placed in cryogenic tubes and cooled at -20°C for 1 hour before being transferred to a -70°C freezer overnight and then placed under liquid nitrogen for long term storage.

(d) *Thawing of sf9 cells*

A 1ml aliquot of frozen sf9 cells was thawed rapidly at 37°C then placed on ice prior to being added to 9ml of FCS and centrifuged at 1000g for 5 minutes at 4°C. The cell pellet was resuspended in 5ml of cold TC100/10% FCS and placed at 28°C in a $25cm^3$ plastic flask. After 12 hours fresh medium was added to the cells and on reaching confluence (3-5 days) the medium was decanted, the cells were suspended in 5ml of TC100/10% FCS and transferred to a $75cm^3$ flask along with 20ml of the same medium. On reaching confluence cells were passaged as in (b) above. To set up a fresh spinner culture the medium was decanted from 2 confluent $75cm^3$ bottles of sf9 cells and the cells were resuspended in TC100/10%FCS, transferred to a spinner flask and made up to 200ml with TC100/10%FCS.

3.4 Preparation of Baculovirus DNA.

500ml of sf9 cells were grown in suspension to a density of 10^6 cells per ml in TC100/10%FCS at 28°C as above and the cells were infected at a multiplicity of 1 p.f.u. per cell with AcNPV or AcRP23-lacZ virus (Possee and Howard,1987). The infections were allowed to proceed at 28°C for 5-6 days until all the cells showed signs of infection (nuclei full of inclusion bodies in the case of AcNPV and large round, vacuolar cells in the case of AcRP23-lacZ). At this stage cells were harvested by centrifugation at 2000g for 10 minutes at

4°C. The supernatant, into which the virus is released, was stored at 4°C until required and the cell pellet discarded. The supernatant was centrifuged at 100000g for 30 minutes at 4°C in a Beckman L8-55M ultracentrifuge. The viral pellet was air dried and then resuspended in 4.5ml of 0.6% SDS, 10mM Tris/HCl pH7.5, 10mM EDTA, made 100µg/ml with respect to proteinase K and incubated for 2 hours at 50°C. The DNA solution was extracted with 1 volume of phenol and then with 2x1 volume of chloroform prior to being precipitated for 20 minutes in a dry ice/ethanol bath following the addition of 10ml of ethanol. The DNA was pelleted by centrifugation at 2000g for 10 minutes and washed with cold 90% ethanol before being dried by rotary evaporation and resuspended in 200µl of 1mM Tris/HCl pH8, 0.1mM EDTA. Incubation for 15 minutes at 65°C aided resuspension and the DNA was then stored at 4°C.

3.5 Transfections.

(a) *pAc23K and AcNPV DNA*

5ml of sf9 spinner cells (10^6 cells per ml) were transferred to a 25 cm^3 flask and left at room temperature to form a monolayer.

0.5µg of AcNPV DNA and 2µg of pAc23K were mixed and made up to 50µl with ddH_2O in a siliconised eppendorff tube prior to the addition of 50µl of lipofectin (BRL). On gentle mixing with a pipette tip a fine precipitate appeared (a heavy precipitate should not form at this stage) and the mixture was left for 10 minutes at room temperature.

The media was decanted from the 25cm^3 flask and the sf9 cell monolayer was washed with 5ml of Optimem (Gibco-BRL) before 3ml of fresh Optimem were added and the transfection mixture was pipetted dropwise and evenly over the cell monolayer. The bottle was rocked gently and then placed at 28°C for 4 hours, after which the Optimem was removed and the cells were washed with 5ml of TC100/10% FCS and left in 5ml of the same medium at 28°C for 4 days. At this stage occlusion bodies were present in the nuclei of some of the cells so the culture was harvested by centrifugation at 2000g for 10 minutes and the supernatant stored at 4°C.

(b) *pVL23K and AcRP23-lacZ*

0.5μg of AcRP-lacZ were linearised by incubation with 10 units of the restriction endonuclease, Bsu36I for 2 hours at 37°C. The enzyme was inactivated by heat treatment at 65°C for 15 minutes.

0.5μg of AcRP23-lacZ x Bsu36I and 2μg of pVL23K were transfected into sf9 cells exactly as described in (a) except that no occlusion bodies formed in cells since AcRP23-lacZ DNA does not encode for the polyhedrin.

3.6 Purification of Recombinant Baculovirus.

Two methods were used in parallel to purify recombinant baculoviruses encoding the 23-kDa Ad2 protein.

(a) *Plaque Purification*

6ml of sf9 cells (10^6 cells per ml) were pipetted into each of twelve $5cm^3$ petri dishes and left overnight at 28°C to form monolayers.

A dilution series of the supernatants from the 4 day transfections (prepared in 3.5) were set up in 1ml of TC100 (no FCS) going from 10^{-1} to 10^{-5} of neat stock. Fresh tips were used for each dilution to prevent carry over.

The medium was removed from the sf9 monolayers in the petri dishes and 100μl of the appropriate virus dilution was pipetted gently onto the cells and allowed to cover the entire monolayer (duplicate plates were set up for each dilution). The plates were left for 1 hour at 28°C in a sealed container. Meanwhile a 3% solution of low melting point agarose (Sigma) was autoclaved, cooled to 37°C and mixed with an equal volume of TC100/10% FCS (also at 37°C). When recombinant virus was prepared from AcRP23-lacZ DNA, 1/100 volume of a 50mg/ml solution of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) in DMF was also added to the agarose.

4ml of the 1.5% LMP agarose were pipetted gently over the cells in each petri dish so that the entire monolayer was covered and the agarose was allowed to set before being overlayed with 4ml of TC100/10% FCS and left at 28°C. After 5 days the medium was removed from above the agarose and replaced with 4ml of 0.05% Neutral Red in TC100/10% FCS for 6 hours. The stain was removed and the plates left overnight at 28°C.

Clear plaques (AcNPV - no occlusion bodies in cell nuclei; AcRP23-lacZ - not blue) were picked through the agarose using pipette tips, suspended in 50µl of TC100 and stored at 4°C until required. To check that the plaques were produced by recombinant virus expressing the 23-kDa protein, a 96-well plate was set up with 100µl of sf9 spinner cells (10^6 cells per ml) in each well. The cells were left for at least 1 hour at room temperature to form monolayers and then the medium was aspirated off. 20µl from each plaque stock were pipetted into separate wells (duplicate wells for each stock) and the plate was placed at 28°C for 1 hour after which 150µl of TC100/10% FCS were added and the plate was returned to 28°C for 60 hours. Supernatants were transferred to the equivalent wells of a second plate and stored at 4°C, whilst the cells were suspended in 15µl of Laemmli electrophoresis buffer and 5µl of gel loading buffer. Samples were boiled for 2 minutes and analysed by SDS-PAGE and Western Blotting (chapter 2, method 2).

A second round of plaque purifications was then carried out using the supernatant corresponding to the sample containing the largest amount of the 23-kDa Ad2 protein and the process was repeated until a homogeneous stock was obtained (only clear plaques present).

(b) *Microtitre Plate Assay for Recombinant Virus*

This assay is of greatest value if the amount of recombinant virus obtained from the transfection is small compared to the parental virus. This was the case when AcNPV DNA was used.

2.5ml of sf9 spinner cells and 2ml of TC100/10% FCS were pipetted into each of 6 sterile tubes. 500µl of the transfection supernatant (prepared in 3.5) was added to the first tube to give a 10^{-1} stock, the tube was mixed by inverting it several times and 500µl was then transferred to the second tube (10^{-2} stock). The process was repeated until a 10 fold dilution series from 10^{-1} to 10^{-6} had been set up and then 180µl of the stocks were pipetted into the wells of a 96-well plate as follows. The 10^{-1} stock was added to all the wells in columns 1-2, the 10^{-2} stock to the wells in columns 3-4 and so on through to the 10^{-6} stock in the wells of columns 11 and 12. The plate was sealed and placed at 28°C for 4 days after which the virus containing supernatants were transferred from each well to the corresponding well in a second

plate using a fresh tip for each transfer. The master plate was stored at 4°C.

To the cells in each well of the original plate 100µl of freshly prepared 0.5M NaOH, 1% SDS were added and left for 10 minutes at room temperature. The lysed cells were neutralised by the addition of 150µl of 0.15M Tris/HCl pH8, 0.3M NaCl, 0.3M acetic acid. A piece of Hybond-N Filter (Amersham), cut to the same size as the plate and soaked in H_2O was labelled for orientation and placed in the vacuum dot blot apparatus (BRL). In order to seal the apparatus and ensure efficient suction it was necessary to wet the plastic sections with H_2O before assembly. 150µl from each well of the microtitre plate were transferred to the corresponding well of the dot blot apparatus and the vacuum was applied to suck the cell lysates through the filter. 2x200µl of 0.15M Tris/HCl pH8, 0.3M NaCl was then washed through each well and the filter was removed from the apparatus and left to air dry for 1 hour. DNA was crosslinked to the filter for 2 minutes on a UV transilluminator and hybridised with the appropriate probe (methods 1.8 and 3.2). A well from the highest dilution that gave a signal was selected and purified further either by setting up a second microtitre plate or by plaque purifying it and following the procedure outlined in (a) above.

3.7 Preparation of Recombinant Virus Stocks

5ml of sf9 spinner cells (10^6 cells per ml) were transferred to to a $25cm^3$ flask and left at room temperature for 1 hour to form a monolayer. The medium was poured out of the flask and replaced with 1ml of TC100 (no FCS) containing 50µl of plaque purified recombinant virus stock. The virus was allowed to absorb onto the cells for 1-2 hours at 28°C before being removed and replaced with 5ml of TC100/10%FCS. The flask was returned to the 28°C incubator and the infection was allowed to proceed for 4-5 days or until most of the cells were lysed. The media containing the virus was decanted from the flask and clarified by centrifugation at 1000g for 10 minutes at 4°C.

A monolayer of sf9 cells was set up in a $75cm^3$ flask using 25ml of a spinner culture as above. The monolayer was infected with 500µl of the virus stock harvested from the $25cm^3$ flask in 2ml of TC100, as described previously and then replaced with 25ml of TC100/10%FCS. The infection was allowed to proceed for 4-5 days and the virus stock

harvested as before.

A 500ml spinner culture (10^6 cells per ml) was infected with 20ml of the virus stock prepared in the $75cm^3$ flask. The infection was allowed to proceed for 5-6 days until most of the cells were lysed and then the medium containing the virus was clarified by centrifugation for 10 minutes at 1000g (4°C). 1ml aliquots of the stock were placed at -70°C for long term storage and the remainder of the stock was stored at 4°C. The titre of the stock was determined by carrying out plaque assays as described in method 3.6a, with duplicate plates being set up for ten fold dilutions of the stock ranging from 10^{-1} to 10^{-9}.

3.8 Expression of the Recombinant 23-kDa Protein in sf9 cells.

For maximum yields of the recombinant protein it is necessary to carry out high multiplicity infections in excess of 1p.f.u. per cell.

(a) *24 well plates*

24 well plates were used to carry out timed infections and monolayers were set up in each well by the addition of 1ml of a 10^6 cells per ml spinner culture of sf9 cells. After the monolayers had formed, the medium in each well was replaced with 200µl of the virus stock (10^7 p.f.u. per ml) or TC100 and the plate was returned to the 28°C incubator for 1-2 hours. The virus was then removed and 1ml of TC100/10%FCS was added to each well and the infection was allowed to proceed. Cells were harvested at 0, 24, 36, 48, 72 and 96 hours post infection by transferring the contents of the appropriate well to an eppendorff tube, centrifuging the cells at low speed in a bench microcentrifuge for 3 minutes and washing the cell pellet x2 with PBS. The cell pellet was then resuspended in 80µl of 50mM Tris/HCl, 1mM EDTA, 0.5% NP40 and 20µl of SDS gel loading dye and 10µl of the sample was analysed by SDS-PAGE and Western blotting (chapter 2, method 2).

(b) *$25cm^3$ and $75cm^3$ flasks.*

Monolayer cultures of sf9 cells were set up in the $25cm^3$ and $75cm^3$ flasks as described in method 3.7. The cells were infected with 0.3 and 1ml of a 10^7 p.f.u. per ml virus stock in a total volume of 1 and 2ml of TC100 respectively and the infection procedure was as described in method 3.7. The cells were harvested at 60 hours post infection by centrifugation at 1000g

for 5 minutes (4°C) and washed x2 with ice cold PBS, pH6.2. The cell pellets were stored at -70°C until required.

(c) *Suspension Cultures.*

500ml cultures of sf9 cells were grown in suspension until they reached a density of $3x10^6$ cells per ml. The cells were then pelleted by centrifugation at 1000g for 5 minutes and resuspended in 15ml of a 10^8 p.f.u. per ml stock of recombinant virus. TC100 was added to a final volume of 50ml and the cells were infected for 1-2 hours at 28°C. 450ml of TC100/10% FCS were added to the culture and the infection was allowed to proceed for 60 hours at 28°C. The cells were harvested and washed as described previously and the cell pellet was stored at -70°C.

3.9 Purification of the 23-kDa protein from insect cells.

The washed cell pellet from an infection of a 500ml suspension culture of sf9 cells (method 3.8c) was suspended in 5ml of ice cold 50mM Tris/HCl, pH8 and homogenised on ice with 20 strokes of the tight fitting pestle (pestle B) in a Dounce homogeniser. The nuclei were sedimented for 5 minutes at low speed in a microcentrifuge and the supernatant (cytoplasmic extract) was clarified by centrifugation for 20 minutes at 26000rpm (30000g) in a Beckman TL100 ultracentrifuge set at 4°C.

The resulting cytoplasmic extract was dialysed against 50mM Tris/HCl pH8 for 2 hours at 4°C and purified using DEAE and phenyl Superose columns essentially as described for the recombinant *E.Coli* protein (method 2.3). The only difference was that a 5ml DEAE Sepharose CL6B column was used and 1ml fractions were collected. 10µl of the fractions collected from the DEAE and phenyl Superose colums were analysed by SDS-PAGE and Western blotting according to chapter 2, method 2.

4 PROTEIN ESTIMATION

(a) *Absorbance at 280nm*

In order to get a very rough estimate of the amount of protein present in a sample its absorbance at 280nm was read and it was assumed that an absorbance of 1 corresponded to 1mg/ml of protein in the solution.

(b) *Bradford Test*

When required the protein concentrations of samples were estimated using the method of Bradford (1976). Protein samples, in a total volume of 10µl, were mixed with 1ml of Bradford's reagent (0.1g Coomassie blue G250, 100ml of orthophosphoric acid and 50ml of ethanol made up to 1 litre with ddH_2O) and incubated at room temperature for 5 minutes. The absorbance of the samples were measured at 595nm and the protein concentrations calculated by comparison with a standard curve constructed using BSA.

(c) *By SDS-PAGE analysis and Coomassie blue staining.*

To estimate the concentration of an individual protein in a heterogeneous mixture the sample was analysed by SDS-PAGE and Coomassie staining (chapter 2, method 2). The intensity of the individual protein band was compared to bands on a standard gel separating 0.05-1µg of soya bean trypsin inhibitor (SBTI).

5 PREPARATION OF POLYCLONAL ANTISERA

A polyconal antiserum was raised in rabbits against the recombinant *E.Coli* expressed 23-kDa protein, purified according to method 2.3. 100µg of purified protein in 1ml of PBS pH6.2 were mixed with 1ml of complete Freund's adjuvant and injected into Dutch rabbits as described in chapter 2, method 3.2. Booster injections of 100µg of protein in 1ml of PBS pH6.2, mixed with 1ml of incomplete Freund's adjuvant, were given at 2 week intervals. Test bleeds and assessment of the antiserum were carried out according to chapter 2, method 3, except that the ELISA plates were coated with 50µl of a 10µg/ml solution of purified *E.Coli* expressed 23-kDa protein.

6 PROTEASE ASSAYS

The ^{35}S methionine labelled Ad2ts1 and the peptide assays for the Ad2 protease were carried out essentially as described in chapter 2, methods 1.1 and 1.3.

RESULTS

1 Expression of the Ad2 Protease as a Fusion Protein in *E.Coli*

The plasmids pGEX23K, pGEX23KΔN and pGEX23KΔC were constructed and transformed into JM101 cells, as described in the method 2.1. The resulting cells were induced with IPTG for 3 hours, harvested by centrifugation, disrupted by treatment with lysozyme and analysed by SDS-PAGE. As a control JM101 cells, transformed with the parent vector, pGEX-2T, were treated in a similar fashion. Duplicate gels were prepared and one was stained with Coomassie blue (**fig.3.4a**), whilst the other was electroblotted onto nitrocellulose paper and probed with a polyclonal antisera raised against the 23-kDa protein (**fig.3.4b**). The stained gel reveals that whilst the pGEX-2T transformed cells over express a protein of apparent molecular weight 26-kDa (lane 2), the cells transformed with pGEX23K, pGEX23KΔN and pGEX23KΔC express proteins of apparent molecular weight 44-46-kDa (lanes 3,5 and 6). In each case, this corresponds to the fusion protein of GST with part or all of the Ad2 protease sequence. Lane 4 shows the pGEX23K extract after a 2 hour incubation with thrombin. Looking at the corresponding lanes in **fig.3.4b** it can be seen that there is no reaction between the anti-23K sera and any of the proteins in the GST expressing cells (lane2); but that bands corresponding to the GST-23K, GST-23KΔN and GST-23KΔC fusion proteins are present in lanes 3,5 and 6 respectively. A band that co-migrates with the Ad2 purified protease (lane 1) appears when the GST-23K containing extract is cleaved with thrombin (lane 4).

The level of expression of the GST-23K fusion protein is extremely high (20mg/litre). Unfortunately, however, the majority of this is present as insoluble inclusion bodies, making it difficult to purify the protein by affinity chromatography using a glutathione Sepharose column (data not shown).

As can be seen from lanes 3, 5 and 6 in **fig.3.4b** there is a significant amount of proteolytic degradation of all the fusion proteins. It is possible that the 23-kDa protein is cleaving the fusion protein, as was observed by Houde and Weber (1990a). On the other hand *E.Coli* proteases could be responsible for the degradation. When the inductions and

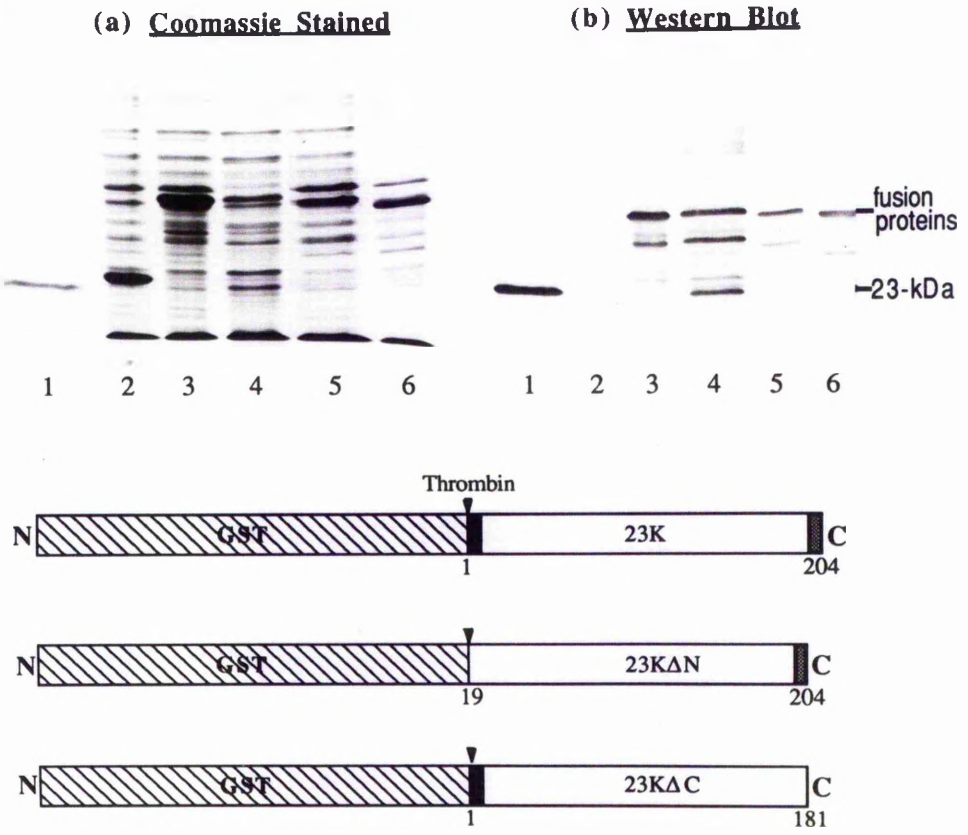


Fig.3.4 Expression of the Ad2 Protease as a fusion protein with GST
Samples were analysed on 12.5% SDS polyacrylamide gels and gel (a) was stained with Coomassie blue whilst gel (b) was electroblotted onto nitrocellulose paper and probed with a polyclonal antisera raised against the 23-kDa protein (1:500 dilution). The samples were IPTG induced JM101 cells transformed with *lane 2*, pGEX-2T; *lane 3*, pGEX23K; *lane 5*, pGEX23KΔN and *lane 6*, pGEX23KΔC. Purified Ad2 23-kDa protein is in *lane 1*, whilst *lane 4* contains the pGEX23K transformed extract incubated with thrombin.

extractions were carried out in the presence of 1mM $ZnCl_2$, a potent inhibitor of the Ad2 protease, the same degree of digestion was observed. Furthermore, incubations of the fusion proteins with disrupted Ad2 did not result in any further degradation (results not shown). Therefore, in this case, *E.Coli* proteases would appear to be responsible for the proteolysis of the GST-23K fusion protein.

2 Specificity of the Antisera Raised against the N- and C- Termini of the Ad2 Protease.

The *E.Coli* extracts prepared above containing the GST-23K, GST-23KΔN and GST-23KΔC fusion proteins were separated by SDS-PAGE alongside protease purified from Ad2. The proteins from 2 identical gels were transferred onto nitrocellulose paper and one was probed with the N-terminal peptide antiserum (**fig.3.5a**), whilst the other was probed with the

C-terminal peptide antiserum (**fig.3.5b**). Both the antisera reacted with the purified Ad2 protease and with the GST-23K fusion protein (lanes 1 and 2). The N-terminal peptide antiserum reacted with GST-23KΔC (lane 4), but not with GST-23KΔN (lane 3). Since the only difference between GST-23K and GST-23KΔN is the first 18 amino acids of the 23-kDa protease, the N-terminal peptide antiserum is specific and does not cross react with any other region of the 23-kDa protein. In a similar vein, the C-terminal peptide antiserum reacts with GST-23KΔN (lane 3), but not with GST-23KΔC (lane 4), so must be specific for the C-terminal residues of the 23-kDa protease.

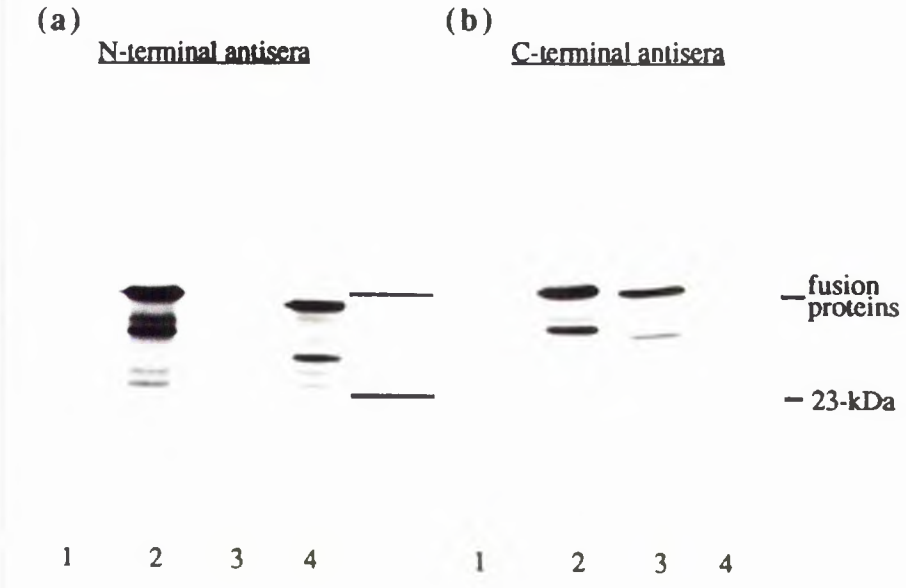


Fig. 3.5 Specificity of 23-kDa N- and C-terminal peptide antisera
Duplicate 12.5% SDS-polyacrylamide gels were prepared with *lane 1*, purified Ad2; and induced JM101 extracts containing *lane 2*, GST-23K; *lane 3*, GST-23KΔN and *lane 4*, GST-23KΔC. The proteins were electroblotted onto nitrocellulose paper and probed with (a) the 23-kDa N-terminal peptide antiserum and (b) the 23-kDa C-terminal peptide antiserum. Both antisera were used at a dilution of 1:1000.

In chapter 2 it was shown that the 23-kDa protein is the Ad2 protease. The results described above show conclusively that the *intact* 23-kDa protein is the protease and that the enzyme is not autocatalytically cleaved at the Ala_{46}-Gly_{47} bond as proposed by Chatterjee and Flint (1987) and Houde and Weber (1990a).

3 Expression of the Ad2 Protease in *E.Coli* using pET Vectors

The cloning of the Ad2 protease and a deletion mutant lacking amino acids 5 to 13 into

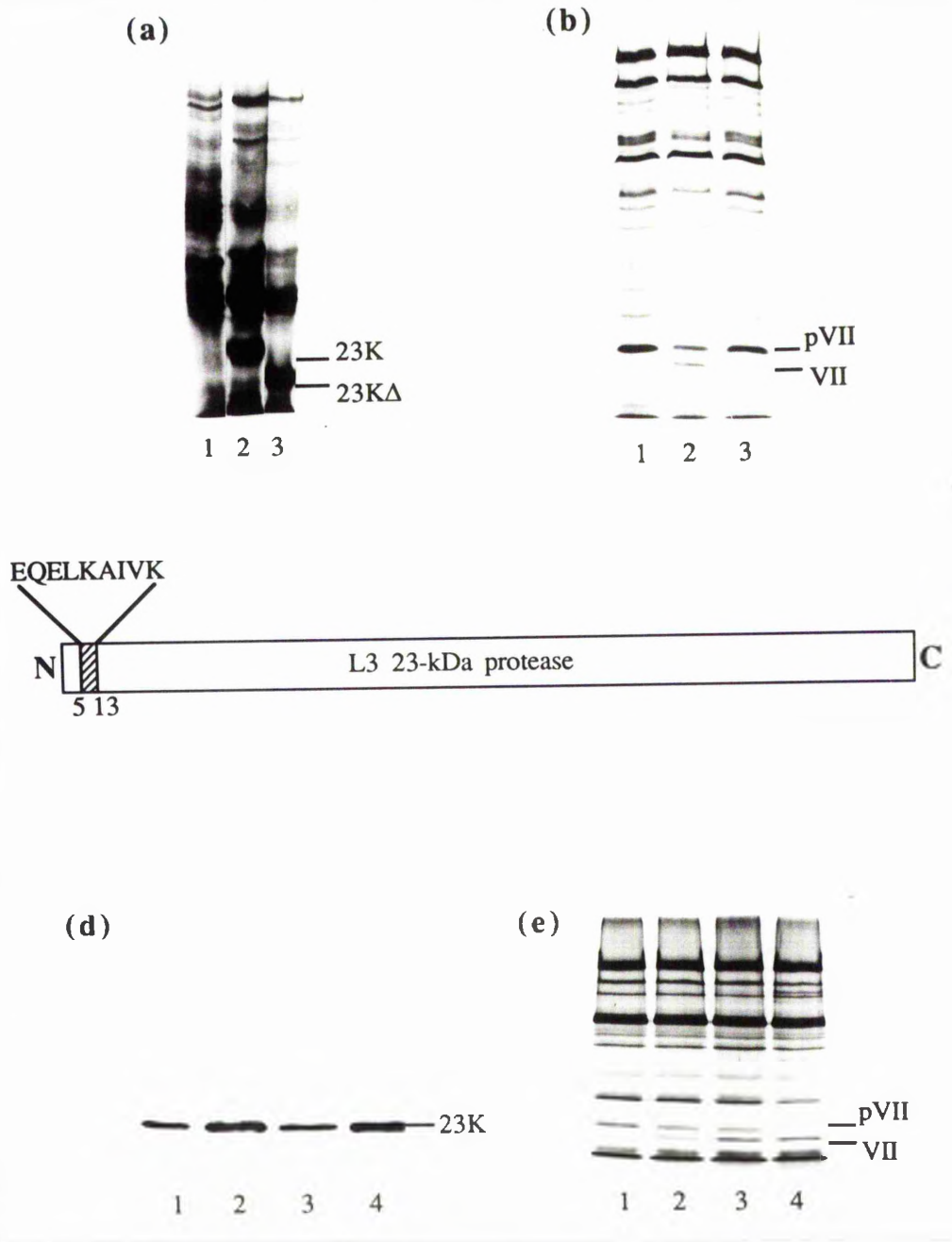


Fig. 3.6 Expression of the Ad2 23-kDa protease in *E.Coli*
(a) Autoradiogram of a 12.5% polyacrylamide gel showing ^{35}S methionine labelled transformed BL21(DE3) cells. The labelling was for 5 minutes; 3 hours after the addition of IPTG to the cultures. Bacteria were transformed with *lane 1*, pET3c; *lane 2*, pT7AD23K12 and *lane 3*, pT7AD23K10. **(b)** Autoradiogram of a 12.5% polyacrylamide gel showing the results of the ^{35}S Ad2ts1 protease assays of the IPTG induced, transformed BL21(DE3) cell extracts. 10µl of substrate were incubated with 5µl of the transformed bacterial extracts for 6 hours at 37°C. The samples analysed were BL21(DE3) cells transformed with *lane 1*, pET3c; *lane 2*, pT7AD23K12 and *lane 3*, pT7AD23K10. **(c)** Schematic diagram showing the amino acids deleted from the 23-kDa protease in the protein expressed by pT7AD23K10 transformed bacteria. **(d)** Western blot showing the amounts of soluble 23-kDa protein present in pT7AD23K12 transformed BL21(DE3) cells induced for 4 hours with IPTG in *lane 1*, M9 (25°C); *lane 2*, M9 (37°C); *lane 3*, LB (25°C) and *lane 4*, LB (37°C). The proteins were probed with the 23-kDa N-terminal antipeptide serum at a dilution of 1:1000. **(e)** Autoradiogram of a 12.5% polyacrylamide gel showing the results of the ^{35}S Ad2ts1 assays to assess the amount of protease activity in the samples analysed in (d). The lane order is the same as for the Western blot in (d) and the incubations were for 3 hours at 37°C.

the pET3 vector was carried out under the supervision of Dr Carl Anderson at the Brookhaven National Laboratory, Long Island, New York (Anderson, 1990). The plasmids pT7AD23K10 and pT7AD23K12 were constructed and transformed into DH1(DE3) and BL21(DE3) cells as described in the method 2.2. The transformed BL21(DE3) cells were labelled for 5 minutes with ^{35}S methionine after being induced for 3 hours with IPTG. The samples were then analysed by SDS-PAGE and autoradiography, and **fig.3.6a** shows that proteins of apparent molecular weights 20-kDa and 17-kDa are being expressed in the pT7AD23K12 and pT7AD23K10 transformed cells respectively (lanes 2,3). The sequences of the 2 proteins were later verified by tryptic digestion and N-terminal sequencing by R. Greene (Anderson, 1990).

Extracts of the pT7AD23K10 and pT7AD23K12 transformed cells were prepared by lysozyme and DNase I treatment as described in the methods section and assessed for activity using the ^{35}S assay. **Fig.3.6b** shows that the intact Ad2 23-kDa protein, expressed from pT7AD23K12, specifically cleaves pVII to VII (lane 2). No digestion of pVII to VII took place when ^{35}S Ad2ts1 was incubated with the pT7AD23K10 transformed extracts (lane 3). The incubations were for 6 hours at 37°C. Thus the deletion of residues 5-13 from the 23-kDa protein appears to result in a total loss of activity and it is concluded that these residues must play an essential structural or catalytic role.

The use of M9 medium instead of LB is preferable for labelling studies and also makes protein purification easier. The yields of soluble 23-kDa protein in M9 and LB were compared at 2 induction temperatures, 25°C and 37°C, and as can be seen from the Western blot in **fig.3.6d** more protein was obtained at 37°C, compared to 25°C in each case. The amount of the 23-kDa protein produced was only slightly greater in LB than in M9 at both temperatures and this is reflected in the activity of the extracts as assessed by the ^{35}S assay (**fig.3.6e**). The assays were carried out according to chapter 2, method 1.1 and the incubations were for 3 hours at 37°C Based on these results and the advantages of M9, it was decided to carry out future inductions at 37°C in the minimal media.

4 Purification of the Ad2 Protease from *E.Coli*

The soluble yield of the 23-kDa protein obtained from a 4 hour induction of pT7AD23K12 transformed cells was 1mg per litre of M9 media. The cells were disrupted by

lysozyme and DNase I treatment and then centrifuged, as described in the methods section. The supernatant was dialysed against 50mM Tris/HCl pH8 and loaded onto a DEAE sepharose column equilibrated with the same buffer. Most of the proteins bound to the column; but the 23-kDa protease eluted in the flow-through (**fig.3.7a**). The single anion exchange step produced an excellent purification as can be seen from the Coomassie blue stained gel. The flow-through fractions from the DEAE column were pooled, made 1.2M with respect to ammonium sulphate and loaded onto an FPLC phenyl-Superose column. The 23-kDa protein bound tightly to the hydrophobic column and was eluted with the low salt buffer B (**fig.3.7b**). As can be seen from the Coomassie blue stained gel the second column not only provided a second purification step, but also served to concentrate the protein. The flow-through from the phenyl-Superose column had a high absorbance at 280nm; but when the fractions were analysed by SDS-PAGE no proteins appeared to be present (results not shown).

A polyclonal antiserum was raised against the purified 23-kDa protein as described in the method 5. The antiserum reacted well in Western Blotting at a dilution of 1:500 (see **fig.3.4b**).

5 Activity of the Purified 23-kDa Protein Expressed in *E.Coli*

A small scale purification from 50ml of induced, pT7AD23K12 transformed BL21 cells was carried out. **Fig.3.8** shows the profile of the DEAE Sepharose separation (5ml column) alongside a Western blot of the flow through fractions. The corresponding fractions were assayed using the ^{35}S assay and peptide assays. Incubation times were 3 hours for the ^{35}S assays and 1-12 hours for the peptide assays. The peak of protease activity, as determined by the ^{35}S Ad2ts1 assay (**fig.3.8**) was found to coincide with the fractions containing the 23-kDa protein. As can be seen the activity of the purified protein (lanes 4-6) is considerably greater than that in the crude extract (lane 1). Presumably the purification of the protease from *E.Coli* extracts has removed some inhibitory factor. The same fractions were tested for protease activity using the peptide assay with very different results. Although MSGGAFSW was cleaved rapidly by the crude extract, digestion did not take place at the Gly-Ala bond and it was concluded that *E.Coli* proteases were responsible. In contrast no protease activity was detected in the flow through fractions, containing the 23-kDa protein. The peptide assays were carried

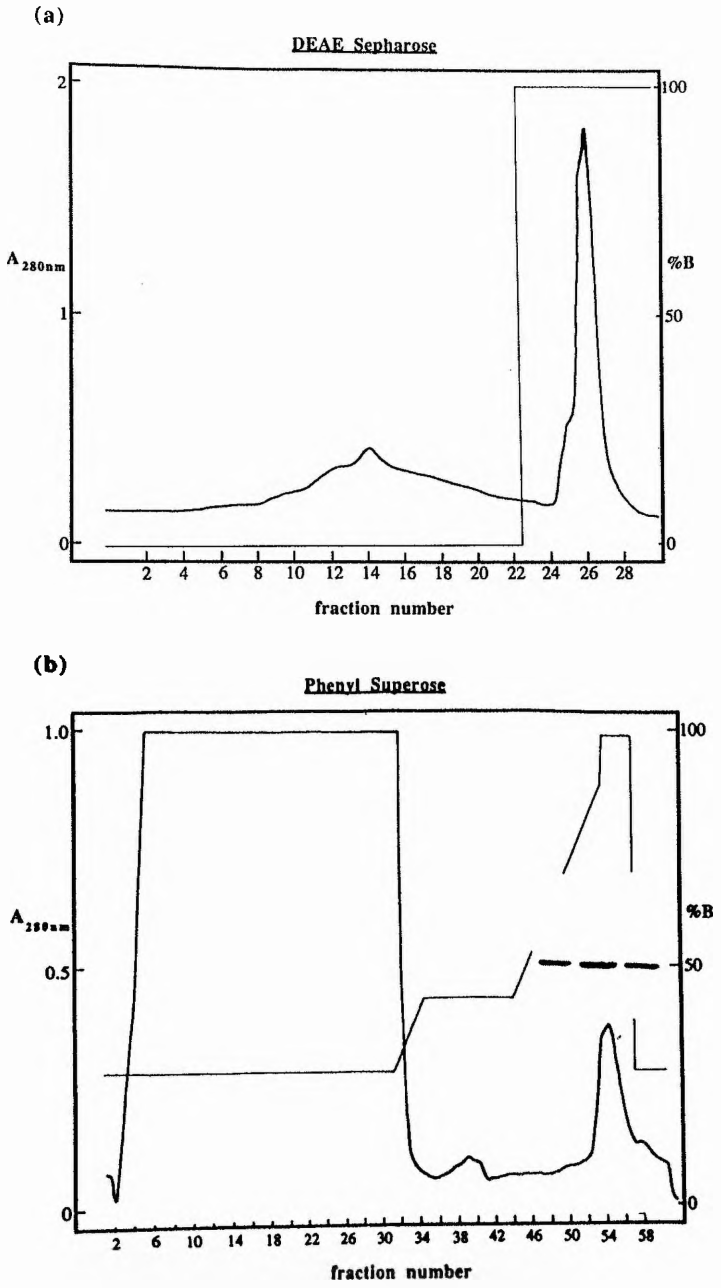


Fig. 3.7 Purification of the 23-kDa protease from *E.Coli*
The recombinant 23-kDa protein was purified from pT7AD23K12 transformed BL21(DE3) cells as described in method 2.3 and result 4. The elution profiles from (a) DEAE Sepharose and (b) phenyl Superose columns are shown with photographs of Coomassie stained gels of the fractions containing the 23-kDa protein superimposed. The left hand lane in the gel in (a) shows the DEAE column load and the remaining lanes correspond to flow-through fractions. The Coomassie stained gel of fractions 53-55 eluted from the phenyl Superose column is included in (b).

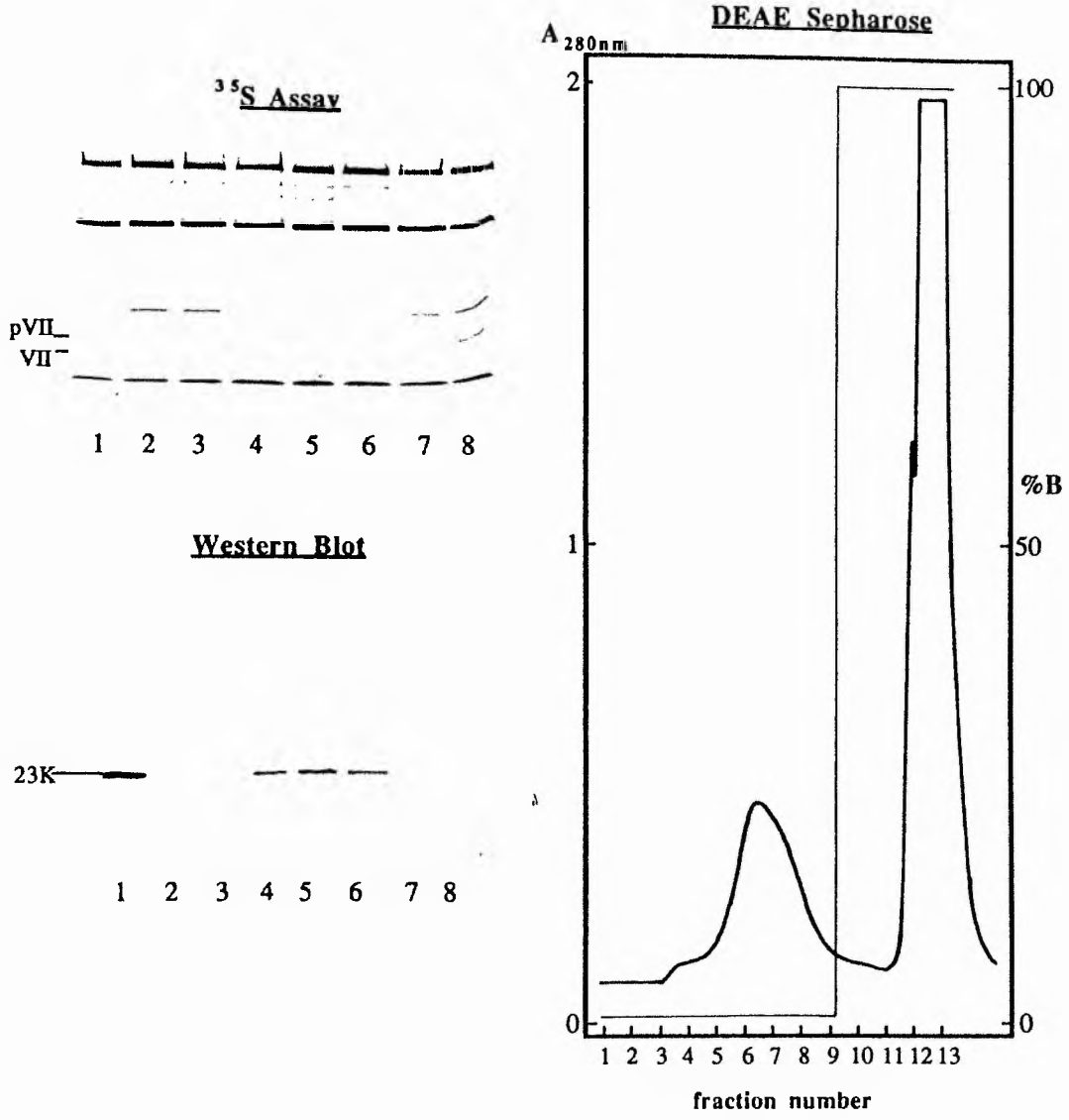


Fig. 3.8 Activity of the 23-kDa protein expressed in *E.Coli*
The elution profile of *E.Coli* expressed 23-kDa protein on a DEAE Sepharose column (see result 5). The fractions were assessed for protease activity using the ^{35}S Ad2ts1 assay and for the presence of the 23-kDa protein by Western blotting with a 1:1000 dilution of the 23-kDa N-terminal peptide antiserum. *Lane 1* in each case is the column load and *lanes* 2-8 correspond to fractions 2-8.

out in the presence and absence of 1M NaCl, 5mM EDTA and 2mM DTT; but still no activity was detected in the 23-kDa peak fractions.

It is, therefore, concluded that some other factor, be it protein or DNA, must be present in the ^{35}S Ad2ts1 extracts that acts as a cofactor. One possibility is that a post-translational modification, such as phosphorylation, is required for activity and that this cannot be carried

out in the *E.Coli*.. It was, therefore, decided to express the 23-kDa protein using the eukaryotic baculovirus expression system.

6 Expression of the Ad2 Protease using Recombinant Baculovirus Vectors

The 23-kDa gene was cloned into 2 vectors, pAcRP23 and pVL1393, and the correct orientations were determined by EcoRI/BglII digestions of small scale plasmid preparations as described in method 3. The plasmids pAc23K and pVL23K were transfected into sf9 cells along with AcNPV or AcRP23-lacZ x Bsu36I DNA. Recombinant viruses expressing the 23-kDa protein were purified and high titre stocks were prepared (method 3). These were then used to infect sf9 cells with a view to purifying the 23-kDa protein.

The expression of the 23-kDa protein was found to be optimal at 48-72 hours post infection and the expression levels using the pAc23K and the pVL23K derived viruses were virtually identical (**fig.3.9**). High multiplicity infections of 1 litre of sf9 cells for 60 hours with the pVL23K derived recombinant baculovirus resulted in the expression of 1mg of soluble 23-kDa protein.

Cell extracts were disrupted by homogenisation in 50mM Tris/HCl, pH8 and clarified by centrifugation. The supernatant contained most of the 23-kDa protein and was put through the purification schedule developed for the *E.Coli* expressed protein (see above). Unfortunately, however, a number of other proteins from the infected sf9 cells did not bind to the DEAE column and even after the phenyl-Superose step the 23-kDa protein was not completely pure (**fig.3.10**). It will, therefore, be necessary to design a better purification protocol for the baculovirus expressed protein in the future. In the meantime, it was decided to investigate the activity of the partially purified baculovirus expressed 23-kDa protein in both protease assays.

7 Activity of the Ad2 Protease Expressed in Insect Cells.

The cytoplasmic extract from 100ml of sf9 cells, infected with the pVL23K derived recombinant baculovirus, was prepared as described previously and dialysed against 50mM Tris/HCl pH8. The sample was loaded onto a DEAE Sepharose column and the resulting fractions were assessed for the presence of the 23-kDa protein by SDS-PAGE and Western blotting (**fig.3.11**). The same fractions were also tested for protease activity in the ^{35}S and

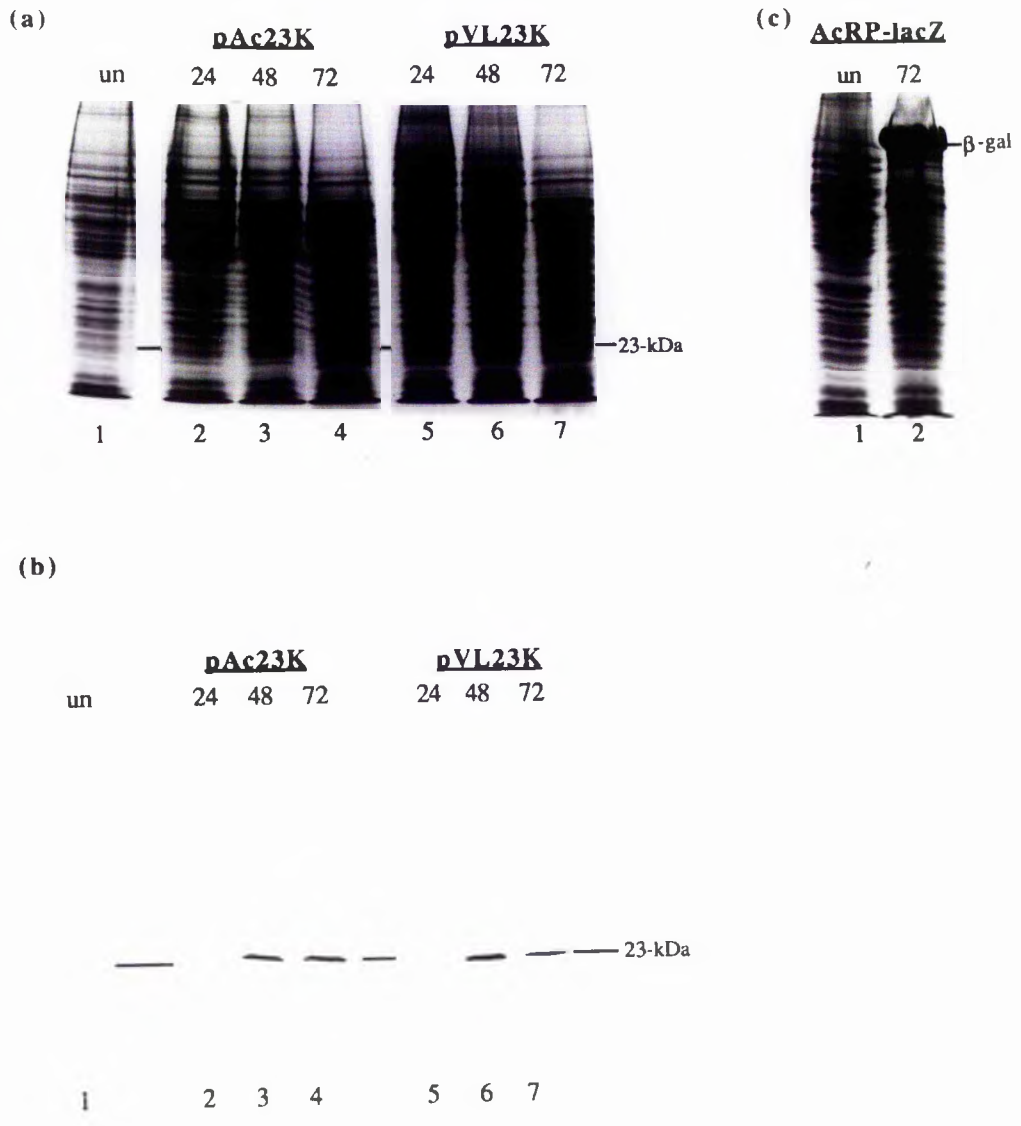


Fig. 3.9 Expression of the Ad2 23-kDa protein in insect cells
sf9 cells were infected with recombinant baculoviruses for 24-72 hours as indicated and analysed on 12.5% polyacrylamide gels. (a) Coomassie blue stained gel and (b) Western blot probed with the 23-kDa N-terminal peptide antiserum at a dilution of 1:1000. In each case the samples are: *lane* 1, uninfected cells; *lanes* 2-4, cells infected with the recombinant baculovirus derived from the plasmid pAc23K and *lanes* 5-7, cells infected with the recombinant baculovirus derived from the plasmid pVL1393 (for details see method 3). A Coomassie blue stained gel of cells infected with the virus AcRP23-lacZ is shown in (c) as a control.

peptide assays with incubation times of 3 and 12 hours respectively. The results were very similar to those obtained for the *E.Coli* expressed material (compare **fig.3.9** and **3.11**), in that the fractions in the flow-through, containing the 23-kDa protein, were active in the ^{35}S Ad2ts1 assay; but not in the peptide assay. Again, more protease activity was detected in the

DEAE flow through (lanes 3-7) than in the crude extract loaded onto the column (lane 1). Thus the partially purified baculovirus expressed 23-kDa protein is not able to cleave MSGGAFSW and it appears likely that some cofactor must be involved.

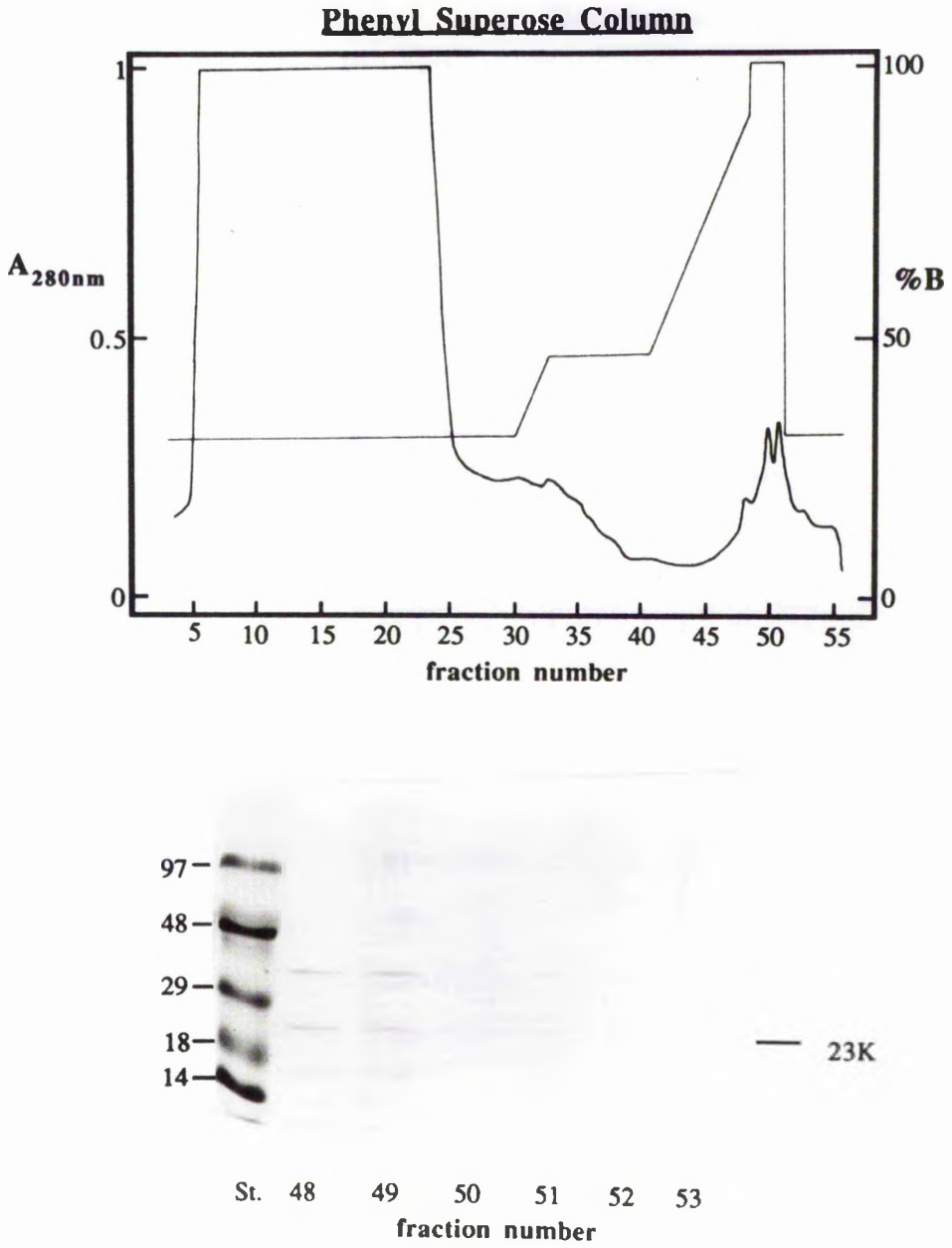


Fig. 3.10 Purification of the 23-kDa protease from insect cells
The 23-kDa protein was partially purified from sf9 cells infected with the pVL23K derived, recombinant baculovirus using anion exchange and hydrophobic interaction chromatography as described in method 3. The elution profile from the phenyl Superose column and a Coomassie stained 12.5% SDS polyacrylamide gel of fractions 48-53 are shown. Molecular weight standards are in lane 1 and the molecular weights of the protein standards, in kDa, are indicated to the left of the gel.

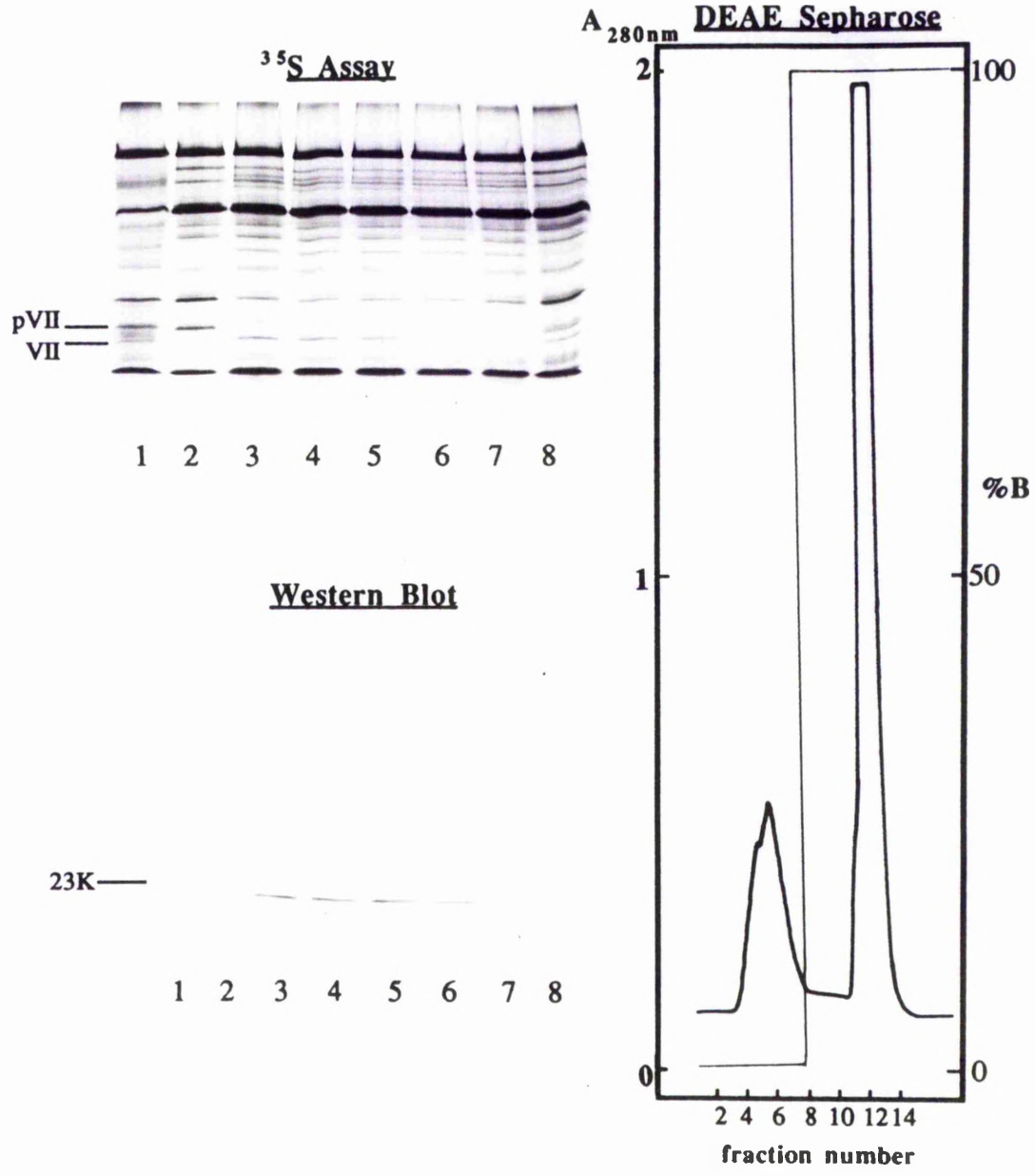


Fig. 3.11 Activity of the 23-kDa protein expressed in insect cells
The elution profile of the recombinant baculovirus expressed 23-kDa protein from a DEAE Sepharose column is shown. The column load (lane 1) and fractions 2-8 (lanes 2-8) were assessed for protease activity using the ^{35}S Ad2ts1 assay and for the presence of the 23-kDa protein by Western blotting using the 23-kDa N-terminal peptide antiserum (for details see method 3 and result 7).

8 Comparison of the Activities of the *E.Coli*, Baculovirus and Ad2 Expressed Proteases

The Coomassie stained gel and Western blot in **fig.3.12a** show the relative mobilities of the Ad2, *E.Coli* and baculovirus expressed 23-kDa proteins. As can be seen, they all migrate the same distance and have an apparent molecular weight of 20-kDa.

The incubation of 0.1μg of 23-kDa protease, from each source, with 10μl of ^{35}S Ad2ts1 extract revealed that digestion of pVII to VII took place in each case. The protease was extracted from Ad2 by NaCl/urea treatment as described in chapter 2, whilst it was partially purified from *E.Coli* and baculovirus extracts by anion exchange chromatography (see above). Although the treatment and levels of purity of the enzymes differed, the *E.Coli* expressed 23-kDa protein consistently showed less protease activity than the baculovirus and Ad2 expressed proteins in the ^{35}S assay (**fig.3.12b**). It is likely that the percentage of the *E.Coli* expressed protease that is folded correctly is considerably less than that for the Ad2 and baculovirus expressed proteins. Only the protease purified from Ad2 was able to cleave the peptide MSGGAFSW (**fig.3.12b**). The activities of the baculovirus and *E.Coli* expressed 23-kDa proteins were assessed in the peptide assay in the presence of 0-2M NaCl, 0-5mM EDTA and 0-2mM DTT; but no cleavage was detected even after 12 hour incubations at 37°C.

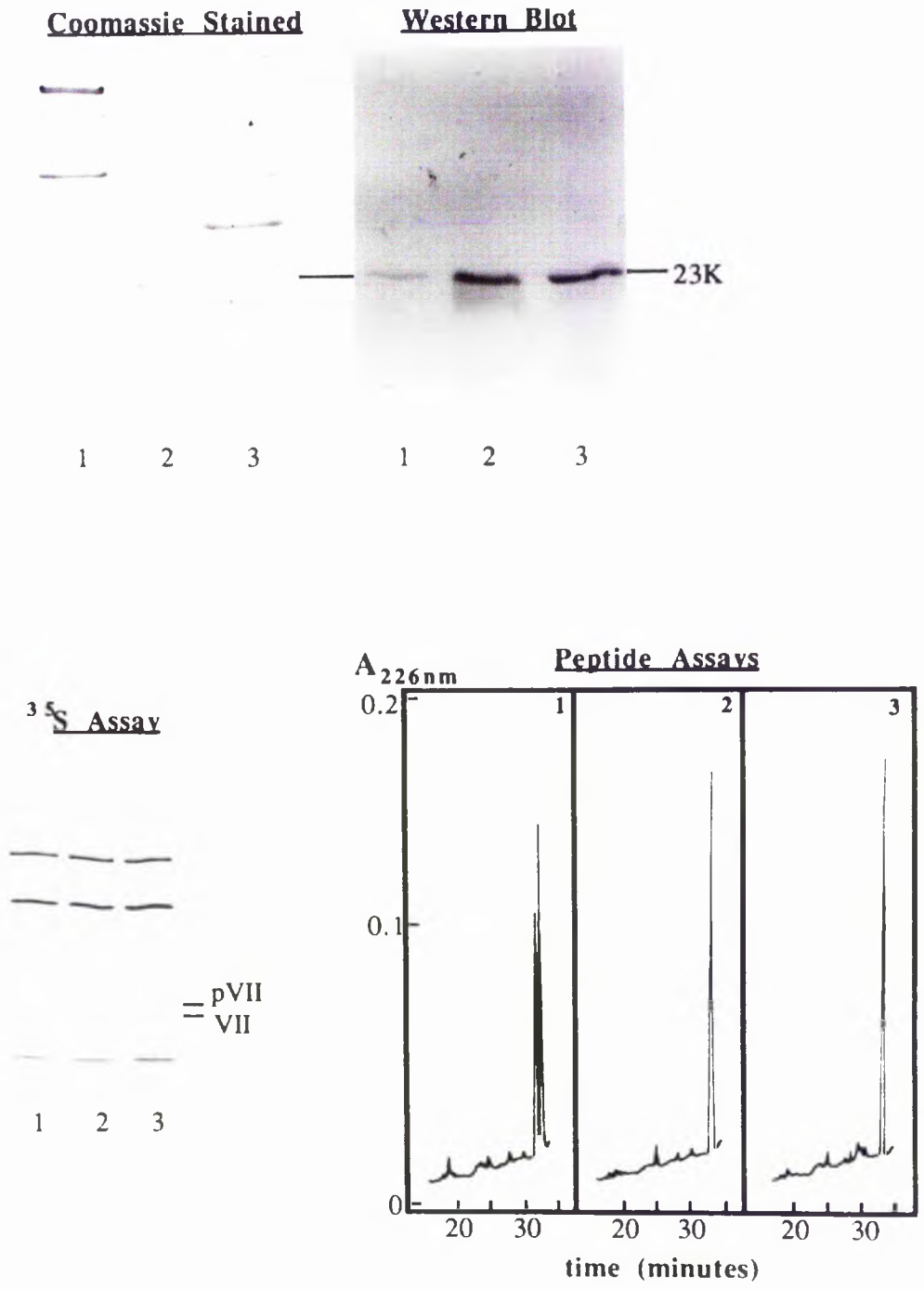


Fig. 3.12 Comparison of Ad2, *E.Coli* and baculovirus expressed protease
The activities of the 23-kDa protease expressed in (1) Ad2, (2) *E.Coli* and (3) baculovirus systems were compared using the peptide and ^{35}S Ad2ts1 assays (see result 8). The Coomassie blue stained gel and Western blot (probed with the 23-kDa N-terminal peptide antiserum) of the 3 samples are included above the autoradiogram and HPLC profiles of the ^{35}S Ad2ts1 and peptide assays respectively. The incubations were for 3 and 2 hours at 37°C in the ^{35}S and peptide assays respectively. The single peak in profiles 2 and 3 is the undigested peptide MSGGAFSW.

DISCUSSION

The first conclusion drawn in this chapter is that the previous reports, by Chatterjee and Flint (1987) and Houde and Weber (1990a), stating that the adenovirus protease is autocatalytically processed are incorrect.

It is shown clearly here that the antipeptide sera raised against amino acids 1-17 and 190-204 of the Ad2 protease are specific and do not cross react with any other sequences in the protein (**fig.3.5**). These antisera both react with the active protease which has a mobility between pVII and VII and has an apparent molecular weight of 20-kDa, when analysed by SDS-PAGE. This means that the 23-kDa protein from Ad2ts1 reported by Chatterjee and Flint (1987) to have a slower mobility than pVII, to be phosphorylated and to bind DFP, cannot be the L3 23-kDa protein. Thus the main basis for their conclusion that the Ad2 protease is autocatalytically processed is wrong. In the same report they describe a 19-kDa protein from wild type Ad2 which migrates between pVII and VII when analysed by SDS-PAGE. The protein was shown to bind DFP and to be phosphorylated and although no sequence evidence was presented, the possibility that this band is the intact Ad2 protease cannot be ruled out. The proposal by Houde and Weber (1990a) that the Ad2 protease is autocatalytically cleaved is based on the degradation of an *E.Coli* expressed fusion protein of protein A and the 23-kDa protein; but no sequence information regarding the cleavage site was given. It seems likely that what they observed was proteolysis by an *E.Coli* enzyme that co-purified with their fusion protein. Such degradation was observed for the GST-23K fusion protein prepared as part of this project (**fig.3.4b**).

EXPRESSION OF THE AD2 PROTEASE

The other major aim of this chapter was to express and purify large amounts of active Ad2 protease and 3 different systems were investigated in an attempt to do this. The insertion of the L3 23-kDa cDNA into the pGEX-2T vector resulted in the expression of a GST-23K fusion protein, at levels of around 20mg/litre, when transformed bacteria were induced for 4 hours with IPTG (**fig.3.4a**). Unfortunately, however, virtually all of the protein was present in insoluble inclusion bodies. The level of expression of the 23-kDa protease using the pET3c

vector was considerably lower, at 2mg/litre; but in this case 25-50% of the protein could be readily solubilised by DNaseI treatment of the lysed bacteria (**fig.3.7**). Finally the protease gene was inserted into the BamHI sites of 2 baculovirus vectors, pAcRP23 and pVL1393, and transfected into sf9 cells with AcNPV or AcRP23-lacZ x Bsu36I DNA. The resulting recombinant viruses were purified and each was found to express the Ad2 protease optimally at 48-72 hours post infection (**fig.3.9b**). The yields of protease from the pAcRP23 and pVL1393 derived viruses, at around 1mg/litre, were 2 orders of magnitude lower than the levels of expression that have been achieved for other proteins, such as β-galactosidase, in the baculovirus system (Possee and Howard,1987). As a control, the AcRP23-lacZ virus expressing β-galactosidase was used to infect sf9 cells in parallel experiments and **fig.3.9c** shows that large amounts of the foreign protein were expressed. It is also noted that the presence of the TATAAAT sequence upstream from the initiation codon in the pVL23K derived virus does not increase the level of expression of the protease. Thus it would appear that the 23-kDa protein might be toxic to the cells when present at high concentrations.

For the purposes of this project, only the pET3c and baculovirus systems will be considered further. In the future, however, it may be possible to solubilise/refold the GST-23K fusion protein from inclusion bodies and so take advantage of the high levels of expression of the protease achieved using the pGEX-2T vector.

PURIFICATION OF RECOMBINANT AD2 PROTEASE

The *E.Coli* expressed 23-kDa protein was purified to homogeneity using anion exchange and hydrophobic interaction chromatography, with yields of 0.5mg of pure protein being obtained per litre of M9/Amp (**fig.3.7**). The minimal media was preferred to Luria broth for bacterial cultures since cleaner protein purifications were achieved. The protein purified by this method was injected into rabbits in order to raise a polyclonal antiserum against the whole protein. The resulting antiserum reacted well in Western blots at a dilution of 1:500 (**fig.3.4b**).

The same purification procedure was followed for infected sf9 cell extracts; but was considerably less successful from 2 viewpoints. Firstly, the sample was not homogeneous and secondly the % yield was low, with only 0.2mg of protease being recovered from a 500ml

culture. In the future if recombinant baculoviruses are to be of use in the expression of the Ad2 protease, then a better purification schedule must be designed.

Cation exchange, chromato-focussing, gel filtration, single stranded DNA-Sepharose, PCMB-Sepharose and zinc chelating columns were tested for their abilities to bind the 23-kDa protease (data not included). The DNA affinity (see chapter 4), zinc chelating and PCMB-Sepharose columns all gave promising results in small scale purifications and could be of use in the future.

ACTIVITY OF THE RECOMBINANT AD2 PROTEASE

The full length 23-kDa protein, expressed in both *E.Coli* and baculovirus systems, was active when assessed in the ^{35}S Ad2ts1 assay. An N-terminal deletion mutant expressed by BL21(DE3) cells transformed with the plasmid pT7Ad23K10, however, was unable to cleave pVII->VII even after 12 hour incubations at 37°C. This shows that amino acids 5-13 are essential for activity and presumably must play an important structural or catalytic role. It also demonstrates that the pET3c system provides a good means for carrying out site-directed mutagenesis of the Ad2 protease in order to assess the importance of individual amino acids to the protease (Anderson,1990).

In contrast to their activity in the ^{35}S Ad2ts1 assay, neither the *E.Coli* nor the baculovirus expressed 23-kDa protein was able to cleave the peptide MSGGAFSW even after incubations for 6 hours at 37°C. It is noted that the cloned protease behaves in a very similar fashion to protease purified from Ad2 in the presence of low salt (chapter 2). The possibility that the *E.Coli* or baculovirus expressed protease would cleave the peptide in a high salt environment was investigated; but even in the presence of 1-2M NaCl no digestion took place. It appears then that in adenovirus infections the protease must be activated in some way and that once activated, if the enzyme is purified, then the presence of high salt is able to lock it in an active conformation. The addition of salt to the recombinant protease is not sufficient to activate it; but a factor present in the ^{35}S labelled Ad2ts1 infected cell extracts is.

CONTROL OF PROTEASE ACTIVITY

As discussed in the main introduction, the control of protease activity is an important feature in the life cycle of many viruses. This is particularly true for viruses whose proteases

are required to carry out maturation cleavages after the assembly of the virus particle. The HIV protease is activated by dimerisation and work by Krausslich (1991) has shown that premature activation of the enzyme is toxic to the cells and prevents assembly of the virions. Another example of activation involving a viral protease is the *trans* cleavage of p220 by the poliovirus 2A protease and eIF3. In this system neither 2A nor eIF3 on its own is able to process the purified p220 protein; but together they cleave the protein efficiently. This is a cooperative event, so much so, that it is not certain whether eIF3 activates 2A or 2A activates eIF3 (Wyckoff *et al.*,1990).

Although it has been demonstrated that the activity of the Ad2 protease is controlled, the mechanism of the activation remains to be established. The results in chapter 2 show that the enzyme is active as a monomer, ruling out dimerisation and as discussed in this chapter limited proteolysis is not involved. The fact that high salt is required to solubilise the protein from the virions and that its predicted isoelectric point is 8.6 would be consistent with it being a DNA binding protein. It has been shown that in the temperature sensitive mutant, Ad2ts1, the 23-kDa protein is not packaged in the virus at non-permissive temperatures (Anderson,1990). This taken along with the observation that most of the substrates of the adenovirus protease are DNA binding proteins means that the possibility that the enzyme binds to, or is activated by, DNA is worth considering. Finally in this regard, it has been shown that both the Sindbis virus capsid protein and the poliovirus 3C protease bind to RNA so there are precedents for viral proteases being associated with nucleic acids (Choi *et al.*,1991; Andino *et al.*,1990). Chapter 4 will consider further the mechanism by which the Ad2 protease activity is controlled.

CHAPTER 4

Control of the Ad2 Protease and Processing of the Pre-terminal Protein

INTRODUCTION

In the previous chapter it was shown that purified recombinant Ad2 protease is unable to cleave the peptide substrate, MSGGAFSW, and the conclusion was drawn that it is synthesised in an inactive form. It was also demonstrated that the enzyme was not activated by limited proteolysis or dimerisation and the suggestion was made that DNA-binding or phosphorylation may be involved. The first part of this chapter is devoted to finding out what controls the activity of the protease. It is hoped that this will enable the purified recombinant enzyme to be activated *in vitro* and so lead to a plentiful supply of active protease. This in turn would enable detailed kinetic, structural and mechanistic studies of the enzyme to be carried out.

The final section of this project will consider one of the natural substrates of the Ad2 protease, the pre-terminal protein (pTP). As mentioned previously, the pTP is the protein primer for adenovirus DNA replication and is covalently attached to the 5' end of the viral DNA. An overview of Ad2 DNA replication was given in the main introduction and so this discussion will concentrate on the 3 viral proteins, the pTP, polymerase and DBP, and the 2 cellular proteins, NFI and NFIII, required for replication.

Replication of adenovirus DNA can be divided into 2 stages, termed initiation and elongation, with the protein and divalent ion requirements of each being different (Leith *et al.*,1989). Initiation of replication involves the covalent attachment of a dCMP base to the β-hydroxyl group of Ser_{580} in the pTP. This process requires ATP and Mg^{+2} and is dependent

on the precise formation of a multiprotein complex at the origin of replication. The overall geometry and stability of the complex is dependent on a number of specific protein-protein and protein-DNA interactions. The expression of the replication factors using baculovirus and vaccinia vectors has enabled interactions between individual proteins to be considered and as a result our understanding of the molecular events involved in initiation has been enhanced considerably (Stunnenberg *et al.*,1988; Watson and Hay,1990; Bosher *et al.*,1990).

The Pre-terminal Protein (pTP)

The pTP is a 671 amino acid protein encoded by the E2 region and the covalent attachment of the dCMP to its serine residue at position 580 is probably catalysed by the adenovirus polymerase although an autocatalytic event cannot be ruled out (Stillman *et al.*,1981). Even the conservative mutation of Ser->Thr at residue 580 has been shown to prevent initiation, showing that the precise orientation of the β-hydroxyl group is critical (Pettit *et al.*,1989). Indeed studies have shown that the pTP is very sensitive to mutagenesis, with the residues at the N-terminus being absolutely required for replication (Pettit *et al.*,1989).

The pTP is expressed coordinately with the adenovirus polymerase (pol) early in infection and the 2 proteins associate to form a stable heterodimer which can only be dissociated in the presence of 1.7M urea (Lichy *et al.*,1982). The binding of the pTP-pol heterodimer to specific sequences at the origin of replication is the central step in the initiation process. It has been shown that whilst the pTP and pol form both specific and non-specific interactions with bases 1-18 of the Ad2 genome, the binding of the heterodimer is stronger and more specific. Indeed, recently Temperley and Hay (1992) have used DNA footprinting assays to show that the pTP-pol binds to bases 9-17 and significantly, these are conserved at the origin of replication in all the adenovirus serotypes.

Examination of the Ad2 pTP sequence reveals a potential nuclear localisation signal at amino acids 362-373 and its identity as such was confirmed by Zhao and Padmanabhan (1988). The binding of pol to the pTP means that this signal is responsible for directing both pol and pTP to the nucleus and so ensures that stoichiometric amounts of each are available for replication. In addition to its role in initiation, mutagenesis studies suggest that the pTP is required for elongation and although this may simply be as a structural component of the

replication machinery it has been proposed that the pTP acts as a helicase to unwind the dsDNA and release the non-template strand (Kenny *et al.*,1988).

As mentioned in chapter 1, the pTP is specifically cleaved to the 55-kDa TP via a 62-kDa intermediate and the processing is carried out by the adenovirus encoded protease. Moreover, 4 potential cleavage sites were identified in the pTP using the consensus sequence described earlier and the positions of these sites are shown in **fig.4.1a**. Octapeptides corresponding to the cleavage sites A and B were synthesised and found to be efficiently cleaved *in vitro*. Cleavage at both these sites which are separated by only 8 amino acids would be consistent with the observation that the 62-kDa iTP appears as a doublet when analysed by SDS-PAGE (Leith *et al.*,1989). The peptides corresponding to sites C and D were not cleaved by the Ad2 protease *in vitro* and it was suggested that the presence of a valine at P_1' in short peptides was not favourable. The similarity between site D and the cleavage sites in pVI and pVII (**fig.4.1b**) and the fact that this site is conserved in other serotypes (**fig.4.1a**) suggests that it is the cleavage site giving rise to the mature TP. The possibility that site C is also cleaved in Ad2, however, cannot be ruled out.

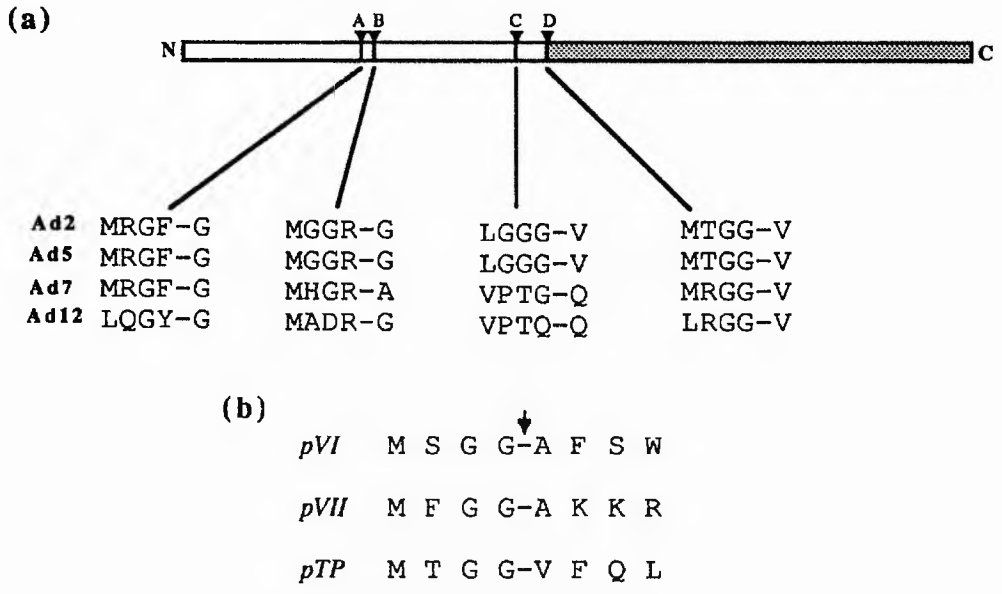


Fig. 4.1 Predicted Cleavage sites in the pre-terminal protein
(a) A schematic representation of the Ad2 pTP showing the locations of the 4 cleavage sites predicted in chapter 1. Listed below each site are the amino acids in the P_4 to P_1' positions at each of these sites in Ad2, Ad5, Ad7 and Ad12. Potential cleavage sites are found at A and D in all the serotypes; but not at B and C. (b) Comparison of the Ad2 pTP site D with the established pVI and pVII sites.

In recent reviews on replication little effort has been made to distinguish between the pTP and the TP with few, if any, references being made to the iTP (Challberg and Kelly,1989; Hay and Russell,1989; Stillman,1989; Salas,1991). Perhaps this is because it is generally accepted that the processing of the pTP is required for a packaging or uncoating event and does not take place until late in infection (Stillman *et al.*,1981). The possibility that the processing to the iTP and/or the TP could play a part in the control of DNA replication, however, should be considered. Although it has been shown that naked adenovirus DNA can act as a template for replication, DNA with the TP or pTP covalently attached to the 5' end is the preferred template. Originally it was thought that this was because the attachment of a protein at the 5' end protected the DNA from exonucleases; but this does not appear to be the case. Other possibilities include the suggestions that the TP is able to unwind the dsDNA, bind to the DNA to change its structure or form a complex with the incoming pTP-pol complex. In any case, it should be remembered that *in vivo*, although the TP-DNA complex is the template for the initial replication event, the templates for most of the subsequent rounds of replication are pTP-DNA complexes. It is possible that the TP-, iTP- and pTP-DNA complexes are indistinguishable in terms of their behaviour in replication or, on the other hand, there may be differences which are functionally significant.

Adenovirus Polymerase (pol)

The Ad2 polymerase is a 1198 amino acid protein with an apparent molecular weight of 140-kDa when analysed by SDS-PAGE (Field *et al.*,1984). It is the enzyme responsible for the synthesis of the viral DNA and although it is homologous to the cellular α and β polymerases, it differs from them in its insensitivity to the inhibitor aphidicolin. Indeed this property provides the basis for a simple assay for specific adenovirus polymerase activity (Shu *et al.*,1987). The Ad2 pol has been expressed in an active form using recombinant baculovirus and vaccinia systems enabling its purification in large amounts (Watson and Hay,1990; Stunnenberg *et al.*,1988). The Ad2 pol sequence contains 2 potential zinc fingers and it is possible that these form the attachment sites to the DNA (Chen *et al.*,1990). In addition to its specific association with the pTP and viral DNA, pol is also thought to bind to the DBP and NFI. It is clear, therefore, that as well as having an important enzymatic role, pol plays a

central part in maintaining the stablilty and overall geometry of the replication machinery. The specific binding of NFI to pol has been shown by a number of groups and is interesting in that it helps to explain why NFI enhances initiation of replication in Ad2 (Bosher *et al.*,1990; Chen *et al.*,1990; Mul and van der Vliet,1992). The association of DBP and pol, on the other hand, is less certain, but it is clear that at physiological concentrations the DBP enhances the initiation of replication (Leith *et al.*,1989). The fact that other non-specific ssDNA binding proteins are unable to substitute for the DBP has led Lindenbaum *et al.* (1986) to suggest that the effect must be as a result of direct binding of the DBP to pol.

The Adenovirus DNA Binding Protein (DBP)

Like the coding sequences for the pTP and pol, the DBP gene is in the E2 region; but the DBP is translated from a different primary transcript (E2a) and so is present in adenovirus infected cells in much larger quantities. Indeed, at early times in infection it is the most abundant viral protein present, with up to $5x10^6$ molecules per cell. This enables the DBP to be purified in reasonable quantities from adenovirus infected cells which have been blocked at the replication stage by the addition of hydroxyurea (Rosenfeld *et al.*,1987; Leith *et al.*,1989).

The DBP is a multifunctional protein that plays a key role in the regulation of adenovirus transcription, translation and DNA replication (Seiberg *et al.*,1989; Hay and Russell,1989). It is a single stranded DNA binding protein, that is also able to bind to double stranded DNA and to RNA. The affinity of the protein for single stranded DNA is considerably higher than for double stranded DNA and the complexes formed are much more stable (Cleghon and Klessig,1986). The Ad2 DBP has 529 amino acids and an expected molecular weight of 60-kDa, considerably lower than its apparent molecular weight of 72-kDa as determined by SDS-PAGE. The difference is because the protein is heavily phosphorylated with at least 17 potential phosphorylation sites (Linne and Philipson,1980; Russell *et al.*,1989). These are mainly serine and threonine residues located in the N-terminal third of the protein and their function is not clear, although one suggestion is that the N-terminal region of the DBP is concerned with defining the host range specificity of the virus (Anderson *et al.*,1983; Brough *et al.*,1985). The DBP contains an N-terminal and a C-terminal domain that are readily separated by proteolysis both *in vitro* and *in vivo*. It appears that the C-terminal 45-kDa

fragment is able to carry out most of the functions attributed to the DBP, including the DNA binding and it appears that this may be mediated by zinc fingers (Vos *et al.*,1988).

The DBP is absolutely required for the elongation stage of replication and EM studies show that it binds to the non-template strand of DNA to stabilise it in the single strand form (Kedinger *et al.*,1978). Although it has been shown that initiation of adenovirus replication can take place *in vitro* in the absence of DBP, there is a requirement for the protein for initiation in *in vitro* systems where physiological concentrations of the factors and templates are used, as well as *in vivo* (Kenny and Hurwitz,1988; Cleat and Hay,1989). It is proposed that the DBP binds to the dsDNA at the origin of replication and alters its structure in such a way as to enhance the binding of NFI to bases 19-39. NFI, however, is not required for Ad4 replication, but there is still a requirement for the DBP in initiation so the protein must have some other role, perhaps through a direct interaction with pol as discussed above (Temperley and Hay,1991).

Nuclear Factor I (NFI)

Nuclear Factor I (NFI) is a multifunctional cellular protein that is required for the replication of Ad2 DNA *in vivo* and has been shown to be identical to the CAAT transcription factor (CTF). NFI (or CTF) is the collective name for a group of proteins ranging in molecular weight from 55-65-kDa that are derived from a single gene by differential splicing (Santoro *et al.*,1988). Whilst the C-terminal domain of the proteins has been shown to be essential for transcriptional activation, the N-terminal 222 amino acids are sufficient for DNA binding, dimerisation and for Ad2 replication (Bosher *et al.*,1991). NFI binds as a dimer to the palindromic sequence located at bases 19-39 of the Ad2 genome and manipulation of the binding site has shown that its position is critical. Deletions or insertions of 1 or 2 bases are sufficient to nullify the helper effect of NFI in Ad2 replication. This is presumably because NFI interacts with pol and the relative positions of the 2 proteins are controlled by the spacing of the pTP-pol and NFI binding sites. As mentioned above the binding of NFI to the DNA is significantly enhanced by the presence of DBP (Cleat and Hay,1989; Stuiver and van der Vliet,1990). No direct interaction between the DBP and NFI has been shown and so it is thought that the cooperative binding results from the DBP's ability to alter the conformation of

the NFI binding site.

Nuclear Factor III (NFIII)

The presence of NFIII is not absolutely required for Ad2 DNA replication, but its binding to base pairs 40-51 of the Ad2 genome has been shown to cause a 2 to 3 fold increase in the initiation reaction. NFIII has a molecular weight of 90-95-kDa and also acts as a transcription factor, binding to the octamer sequence 5'-ATGCAAAT-3', present in the promoter and enhancer regions of a number of cellular genes. Its mutifunctional nature means that it has a number of pseudonyms including Oct-1, OTF-1 and NFA1. The protein has 743 amino acids, which include a DNA binding region termed the POU domain. This 160 residue fragment has been shown to be sufficient for the stimulation of adenovirus replication and contains the DNA binding sequence. As for NFI, additional regions of NFIII are required for transcriptional activation, suggesting that its modes of action in replication and transcription are distinct (Verrijzer *et al.*,1990). Despite the fact that the NFI and NFIII binding sites are adjacent in Ad2 DNA no interaction or cooperative binding has been shown between the 2 factors and it appears that they employ different mechanisms to enhance adenovirus DNA replication (Mul *et al.*,1990). Since no interaction has been shown between NFIII and NFI or pTP-pol, it is possible that NFIII acts indirectly by altering the structure of the DNA in such a way that initiation is favoured.

The realisation of some of the interactions and associations described above has led to a number of models for the formation of the pre-initiation complex in Ad2 and one such model is shown in **fig.4.2.** One version of the sequence of events is that the DBP binds to the dsDNA at the origin of replication to alter its structure in a way that enhances the binding of NFI to bases 19-39. NFI then binds as a dimer and acts as a docking mechanism for the pTP-pol heterodimer which binds specifically to bases 9-17. The presence of NFIII at bases 40-51 also appears to enhance the binding of pTP-pol, perhaps by changing the structure of the DNA. Additional interactions may also take place in the formation of the pre-initiation complex and among those that have been proposed are associations between DBP and pol and between pTP-pol and the TP attached to the 5' end of the template strand.

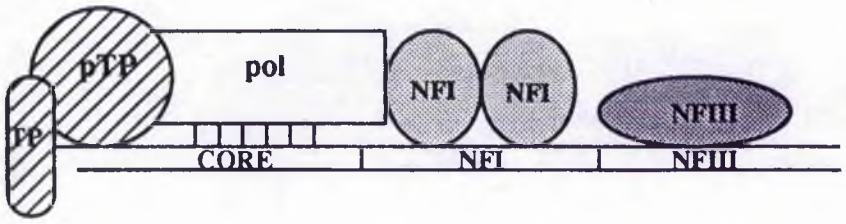


Fig. 4.2 Model for initiation of Ad2 DNA replication
Schematic diagram showing the origin of replication in the Ad2 genome and the binding sites for the cellular and viral replication factors. The core sequence comprises bases 9-17 conserved in all the adenovirus serotypes. NFI binds as a dimer to the palindromic sequence at bases 19-39, whilst the NFIII binding site is at bases 40-51. The Ad2 DBP also plays a role in stimulating initiation of replication; but its mode of action remains to be established so it is not included in the above diagram.

As mentioned previously the cellular factors NFI and NFIII do not appear to have any stimulatory effect in DNA replication of the group E virus, Ad4 and it has been proposed that in the case of Ad4 the pTP-pol heterodimer has a higher affinity for bases 9-17 at the origin of replication. This removes the requirement for the cellular factors. As yet the Ad4 pTP and pol genes have not been sequenced, but it has been shown that polyclonal antisera raised against the Ad2 pol does not react with pol from Ad4 (Temperley and Hay,1991).

In this chapter after the control of the Ad2 protease is considered the expression of the Ad2 pTP using baculovirus vectors will be described. It is hoped that this will form the basis for future studies on the roles of the pTP, iTP and TP in adenovirus DNA replication.

METHODS

1 DENATURED CALF THYMUS DNA SEPHAROSE CHROMATOGRAPHY.

Cytoplasmic extracts of sf9 cells, infected with the pVL23K derived recombinant baculovirus, were prepared and partially purified by anion exchange chromatography as described in chapter 3, method 3.9. The fractions containing the recombinant 23-kDa protein were applied to a 5ml single stranded calf thymus DNA-Sepharose column (Sigma) equilibrated with 50mM Tris/HCl, pH8. The flow rate was 0.5ml/minute and the column was washed with 25ml of equilibration buffer after the sample was loaded. The column was then eluted with 10ml each of the equilibration buffer made 0.2, 0.4 and 1M with respect to NaCl. 1ml fractions were collected and 15µl of each was analysed by SDS-PAGE as described in chapter 2, method 2.

2 GEL ELECTROPHORESIS DNA BINDING ASSAYS.

The method used for these assays is based on that described by Temperley and Hay (1992).

2.1 Labelling of DNA.

5µg of the plasmid pEX8, containing bases 1-600 of the Ad2 genome (Barrett *et al.*,1987) were cut with the restriction endonucleases EcoRI and PvuII to release a 466 base pair fragment corresponding to bases 1-466 of the Ad2 genome. The 5' overhang created by the EcoRI cut was labelled with [α-^{32}P] dATP using Klenow polymerase and the reaction was set up as follows: 5µg pEX8 x PvuII/EcoRI, 1.5µl 10xKGB (1M potassium glutamate, 0.25M Tris/acetate pH7.5, 0.5mg/ml BSA, 5mM β-mercaptoethanol), 4 units of Klenow polymerase, 30µCi [α-^{32}P] dATP made up to 25µl with H_2O. The reaction was allowed to proceed for 45 minutes at room temperature before the labelled fragment was purified by gel electrophoresis.

The gel was prepared by mixing 8ml of 10xTBE (0.5M Tris/borate, 10mM EDTA, pH8.3), 8.4ml of 30% acrylamide (29:1, acrylamide: bis-acrylamide), 32.6ml of H_2O, 4ml of freshly prepared 1% ammonium persulphate and 50µl of TEMED and pouring it into a gel cassette of width 0.75mm. 5µl of DNA loading dye (chapter 3, method 1.2) were added to the

sample and it was loaded onto the gel alongside DNA molecular weight markers. The electrophoresis buffer was 1xTBE and a constant voltage of 200V was applied to the gel for 2 hours, after which the gel was removed from the cassette, covered with cling film and exposed to X-ray film for 5 minutes. The labelled band was excised from the gel, cut up into small pieces and placed in an eppendorff tube along with 400µl of 10mM Tris/HCl pH7.5, 0.1M NaCl, 1mM EDTA. The DNA was eluted from the gel by shaking the tube on a mechanical shaker at 37°C overnight. The fragments of gel were removed by filtering the sample through glass wool and the DNA was precipitated by adding 1ml of ice cold ethanol and leaving it at -70°C for 30 minutes. The labelled probe was sedimented by centrifugation for 5 minutes at high speed in a microcentrifuge, air dried and resuspended in 1mM Tris/HCl pH8, 0.1mM EDTA such that each µl contained 100 counts per minute.

2.2 DNA binding Assays.

DNA binding reactions were incubated for 30 minutes at room temperature and each assay included 0.1ng of probe and 2µg of unlabelled DNA to prevent non-specific binding to the probe. Each reaction was set up as follows: 1µl of labelled probe, 1µg poly d(A-T), 1µg poly d(G-C), 1µl of 1mg/ml BSA, 0.1µg of 23-kDa protein (or pTP), 2.5µl of 0.5M Tris/HCl pH8 and was made up to 25µl with H_2O. Reactions were stopped by the addition of 5µl of DNA loading dye and analysed by gel electrophoresis as described in 2.1 (above). The resulting gel was dried and exposed to X-ray film overnight.

3 PREPARATION OF ANTISERUM AGAINST THE pTP.

A peptide including the N-terminal 15 amino acids of the Ad2 pre-terminal protein:

M A L S V N D C A R L T G Q S C

was synthesised as described in chapter 1, method 1. The antiserum was prepared by coupling the peptide to HSA and injecting the conjugate into rabbits as described in chapter 2, method 3.

4 CLONING OF THE pTP GENE INTO BACULOVIRUS VECTORS.

The methods used to clone the pTP gene into the baculovirus vector pVL1393 were essentially the same as those described for the 23-kDa gene in chapter 3, methods 1 and 3. A schematic diagram showing the construction of the pVLpTP vector is shown in **fig.4.3.**

A fragment representing the entire coding region for the pTP was constructed from 2 regions of the viral genome in 2 stages. The first step involved copying the coding region for the N-terminal third of the protein from Ad2 DNA using PCR as described in chapter 3, method 1.4. The oligonucleotide primers used were:

(8) 5'-GGGATCCATGGCCTTGAGCGTCAACGACTGCGCACGCCTCAC-3'

(9) 5'-CGAAAAACCTCTCGAGAAAGGCGTC-3'

Oligonucleotide (8) includes a BamHI site at its 5' end followed immediately by the codons for the N-terminal 3 amino acids of the pTP, coded for by the exon at map position 39 in the genome (see **fig.8**, p22). The remainder of oligonucleotide matches exactly with the start of the main coding region for the pTP at map position 26 on the reverse strand. Oligonucleotide (9) is identical to a 25 base sequence on the + strand within the main coding region of the pTP. The 800 base pair fragment synthesised in the PCR reaction was purified from an agarose gel and treated with T4 polynucleotide kinase and T4 polymerase. It was then ligated to the vector pUC13, which had been linearised with the restriction enzyme SmaI and treated with calf intestinal alkaline phosphatase (CIAP). DH5 cells were transformed with the ligated DNA and bacteria transformed with pUC13 containing the PCR fragment, were selected by replica plating. The probe used was the purified PCR fragment which was labelled by random priming with [α-^{32}P]dATP. Mini plasmid preparations from the 123 positive colonies were carried out and the resulting DNA was digested with the restriction enzyme BamHI. Only one of the resulting plasmids gave an 800 base pair fragment when digested with BamHI. The plasmid DNA (pUCpTP1) was cut with the restriction enzyme XhoI, then treated with CIAP and used in the next stage.

A plasmid comprising the Hind B fragment of Ad2 (base pairs 6231-11555), which includes the main coding region for the pTP, in pUC13 was supplied by R. Hay. This plasmid was cut with XhoI to release a fragment containing the C-terminal two thirds of the pTP gene. The XhoI fragment was purified from an agarose gel, ligated to the CIAP treated, pUCpTP1 x XhoI and then transformed into DH5 cells. Bacteria transformed with a plasmid containing the XhoI fragment were selected by replica plating. The probe used was the XhoI fragment, which had been labelled with [α-^{32}P]dATP by random priming, and mini plasmid preparations

were carried out for 6 of the positive colonies. KpnI/EcoRI digests were used to select the plasmids containing the XhoI fragment in the correct orientation (pUCpTP2).

The BamHI fragment containing the pTP gene was cut out of pUCpTP2 and ligated to pVL1393 x BamHI which had been treated with CIAP. The ligated DNA was transformed into DH5 cells and transformants were selected by replica plating and probing as described for the selection of pUCpTP2. Mini plasmid preparations were carried out for 6 positive colonies and the orientation of the pTP gene was checked by KpnI/EcoRV digests.

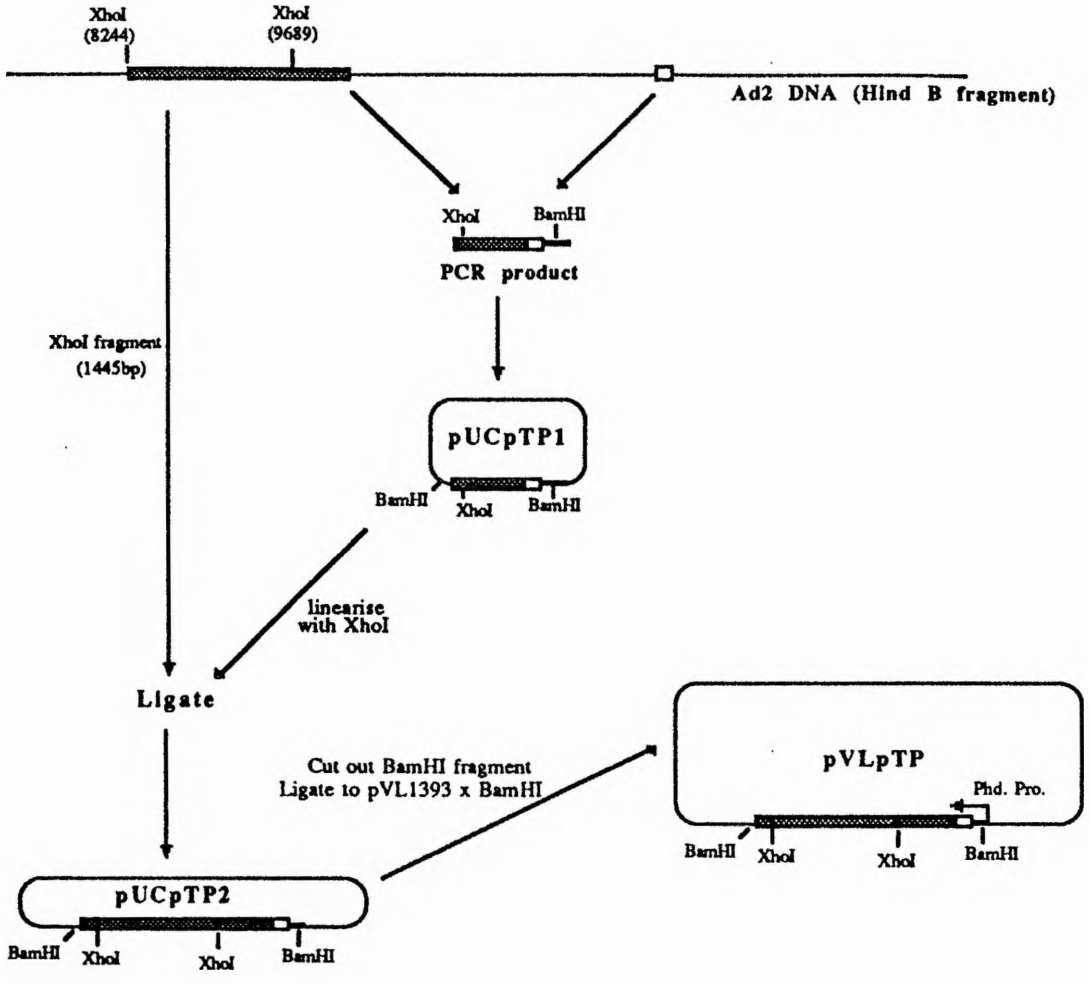


Fig. 4.3 Construction of the plasmid pVLpTP
A schematic diagram showing the construction of the plasmid pVLpTP as described in method 4. The polyhedrin promoter (Phd. Pro.) is marked. The open box (□) corresponds to the exon at position 39 coding for the first 3 amino acids of the pTP. The shaded box includes the main coding region for the pTP.

The resulting plasmid pVLpTP was transfected into sf9 cells with pAcRP23-lacZ x Bsu36I as described for the plasmid pVL23K in chapter 3, method 3.5. The recombinant virus was purified by 3 rounds of plaque purifications, a high titre virus stock was prepared and this was used to express the recombinant protein.

5 PURIFICATION OF THE RECOMBINANT pTP.

A 500ml culture of sf9 insect cells was infected with a recombinant baculovirus containing the pTP gene as described in chapter 3, method 3.8. Due to reasons that will become apparent later (result 6), the pTP recombinant baculovirus used was not that prepared in method 4; but instead was a virus prepared and supplied by Lars Rogge (University of Munich). The infection time was 72 hours and cells were harvested by centrifugation at 1000g for 5 minutes. After washing twice with PBS the cells were resuspended in 4ml of 25mM Hepes/KOH pH8, 5mM KCl, 0.5mM $MgCl_2$, 0.5mM DTT and disrupted on ice in a Dounce homogeniser with a type B pestle (20 strokes). Nuclei were sedimented by centrifugation for 5 minutes at high speed in a microcentrifuge and then resuspended in 4ml of the same buffer made 0.4M with respect to NaCl and incubated on ice for 15 minutes. Cell debris was removed by centrifugation for 5 minutes as before and then the extract was clarified by centrifugation at 100000g in a Beckman TL100 centrifuge. The supernatant was diluted to reduce the NaCl concentration to 0.2M and then applied to a 10ml single stranded calf thymus DNA-sepharose column equilibrated with 25mM Hepes/KOH pH8, 1mM EDTA, 2mM DTT, 0.2M NaCl and 10% glycerol. The column was washed with 50ml of the equilibration buffer and then eluted with 20ml of the same buffer made 0.6M with respect to NaCl. 1ml fractions were collected and the protein concentration of each was estimated by reading the absorbance at 280nm. 10µl of each fraction were analysed by SDS-PAGE. The fractions containing the pTP were pooled and 0.5ml of the resulting sample was made 1.2M with respect to KCl before being applied to an FPLC phenyl Superose HR5/5 column equilibrated with 0.1M potassium phosphate pH7, 2M KCl, 1mM DTT. The flow rate was 0.5ml/minute, the detection wavelength 280nm and the full scale deflection was set at 0.5. A linear solvent gradient going from 0-100% bufferB (0.1M potassium phosphate pH7, 1mM DTT, 10% glycerol) in 45 minutes was applied to the column and the pTP eluted at 90-100% B. 1ml fractions were

collected and 10µl of each were analysed by SDS-PAGE.

6 *IN VITRO* REPLICATION ASSAYS

The pTP was labelled *in vitro* using [α-^{32}P] dCTP in an initiation reaction. The Ad2 cores (template), uninfected cell extracts (source of NFI and NFIII), purified DBP and partially purified pTP/polymerase were provided by I.Leith and were prepared as described in Leith *et al.* (1989).

The initiation reactions were set up as follows: 4µl of Ad2 cores, 5µl of 5x buffer, 2.5µl of 20mM $MnCl_2$, 2.25µl of 40mM ATP, 8µl of uninfected HeLa cell extracts, 2µg of DBP, 2 units of pTP/pol and 10µCi [α-^{32}P]dCTP made up to 25µl with H_2O. The 5x buffer used in each case was 25mM Hepes/KOH pH7.5 made 0, 2.5, 5 or 25mM with respect to DTT. The incubations were for 90 minutes at 30°C after which the product was digested with micrococcal nuclease (Pharmacia, 5 units) in the presence of 10mM $CaCl_2$ for 30 minutes at 37°C. The reactions were stopped by the addition of 8µl of SDS-PAGE loading dye and the samples were boiled for 2 minutes prior to analysis on 12.5% acrylamide gels (see chapter 2, method 2).

In one experiment the action of the micrococcal nuclease was stopped by the addition of 1µl of 0.25M EDTA pH8 and 20µl of disrupted, purified Ad2 were added to 5 reactions, set up as above with the 5x buffer containing 5mM DTT. Samples were incubated for 0, 1, 2, 6 and 12 hours at 37°C before the addition of 10µl of SDS-PAGE loading dye, boiling and analysis by SDS-PAGE as above.

RESULTS

1 Effect of Inhibitors on *E.Coli* and Baculovirus expressed Ad2 protease

In chapter 2 the effects of a range of inhibitors on the native Ad2 protease were investigated using the peptide and ^{35}S assays. Here we consider the effect of a number of inhibitors on the protease activity of the recombinant 23-kDa protein expressed in *E.Coli* and baculovirus systems. As described in chapter 3, the recombinant material was not active in the peptide assay and so only the ^{35}S assay was used in this case. Purified *E.Coli* or partially purified baculovirus expressed protease was pre-incubated with a range of inhibitors for 10 minutes at room temperature prior to a 3 hour incubation with 10µl of ^{35}S labelled Ad2ts1 extracts at 37°C (chapter 2, method 1.1). The degree of digestion of pVII to VII, in each case, was determined by analysis of the samples by SDS-PAGE and autoradiography. The results obtained were identical for the *E.Coli* and baculovirus expressed material (**table 4.1**) and in virtual agreement with the inhibition profile described for the Ad2 protease in chapter 2.

Inhibitor	Activity in ^{35}S Assay
PMSF (1mM) DCI (0.1mM)	- -
Iodoacetate (1mM) PCMB (1mM) DTT (2mM)	inhibited inhibited 80% inhibited
o-phenanthroline (10mM) EDTA (10mM)	- -
pepstatin (2µM)	-
α_2-macroglobulin (2µM)	inhibited

Table 4.1 Effects of selected inhibitors on recombinant protease

The only difference observed was in the effect of the thiol reagent DTT. In chapter 2 it was reported that in the peptide assay purified Ad2 protease was activated up to 5 fold by the presence of 0.5-2.5mM DTT, whereas **table 4.1** shows that 2mM DTT inhibits the *E.Coli*

and baculovirus expressed protease.

2 Effect of DTT on Recombinant Ad2 Protease

The ^{35}S Ad2ts1 assay was used to investigate the effects of 0-5mM DTT on Ad2 protease activity. Two sources of protease were used, purified *E.Coli* expressed 23-kDa protein and purified disrupted Ad2 (**fig.4.4**). Incubation of the respective protease sources with 10µl of ^{35}S Ad2ts1 was for 3 hours at 37°C in the presence of 0, 1, 2, 5 and 10mM DTT. The samples were analysed by SDS-PAGE and autoradiography as described previously. The results show that the digestion of pVII-VII by the recombinant protease is reduced by 80% in the presence of 2mM DTT and at 5mM DTT virtually no digestion takes place (**fig.4.4**, lanes 4,5). The native Ad2 protease was active in the presence of 2mM DTT (lane 9) and only partial inhibition was observed at concentrations of 5 and 10mM (lanes 10,11), consistent with the results reported for the peptide assay (chapter 2).

Taken together, these results suggest that the cleavage of pVII by the recombinant protease is a 2 step process. The first step involves the activation of the protease and is

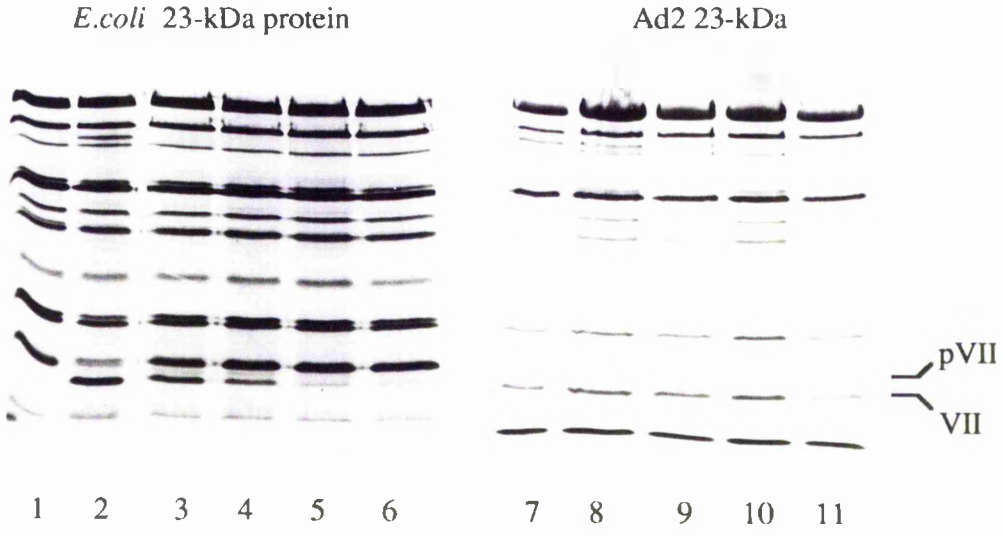


Fig. 4.4 Effect of DTT on *E.Coli* and Ad2 expressed protease activity
The effect of 0-10mM DTT on the activities of the *E.Coli* and Ad2 expressed 23-kDa protease were assessed using the ^{35}S Ad2ts1 assay. The assays were examined by SDS-PAGE and autoradiography. *Lane 1*, substrate alone; *lanes 2-6*, substrate incubated with *E.Coli* expressed protein at 0, 1, 2, 5 and 10mM DTT; *lanes 7-11*, substrate incubated with disrupted purified Ad2 at 0, 1, 2, 5 and 10mM DTT. Incubations were for 3 hours at 37°C.

inhibited by DTT and the second step is the proteolysis itself, which is activated by 0.5-2.5mM DTT. The protease from purified Ad2 is already in an active form, explaining why it is not inhibited by 1-2mM DTT in the ^{35}S assay. Thus the differential effect of DTT provides further evidence that the Ad2 protease is controlled in some way.

3 Effect of DNA and ATP on Recombinant Protease Activity

The fact that 1M NaCl is required to solubilise the Ad2 protease from purified virus suggests that it may be bound to DNA and raises the interesting possibility that the enzyme is activated on binding to a specific DNA sequence. Further evidence for this came from gel retention assays by Maria Grable, Pat Hearing and Carl Anderson which showed that under some conditions purified *E.Coli* expressed protease can bind to the packaging sequences at bases 190-300 of the Ad2 genome (C.Anderson, personal communication). It was decided, therefore, to see if the addition of Ad2 or calf thymus DNA to the partially purified recombinant baculovirus protease would be sufficient to activate it in the peptide assay. The assays were set up as described in chapter 2, method 1.3, except that the buffer used was 50mM Tris/HCl, 0.1M NaCl, pH8. Incubations were for 6 hours at 37°C. 1-5µg of purified Ad2 DNA, pEX8 (a plasmid containing bases 1-600 of the Ad2 genome) or calf thymus DNA (double or single stranded) were included in the peptide assay in the presence and absence of 1mM ATP. No digestion of the peptide, MSGGAFSW, took place in any of the assays, indicating that DNA alone is not sufficient to activate the recombinant Ad2 protease. Calf thymus DNA-Sepharose beads were also assayed in case activation involves a matrix; but again no activity could be detected in the peptide assay. It therefore appears likely that a viral or a cellular protein may be involved.

4 Activation of the Baculovirus Expressed Protease by an Ad2 Extract

As described in chapter 2, protease activity can be extracted from purified Ad2 using 2M NaCl, 2M urea. The aim of this section is to determine whether the DNA pellet from such a purification can be used to activate the baculovirus expressed protease. In order to remove as much protease activity as possible, the DNA pellet, prepared as described in chapter 2, method 4.2, was extracted further with 2x 4M Urea, 2x 4M NaCl and 2x 50mM Tris/HCl, pH8. It was then resuspended in 400µl of 50mM Tris/HCl, pH8 by vortexing and sonication (4x15s

on ice).

Peptide assays were carried out essentially according to chapter 2, method 1.3, with the following modifications. The source of protease was 30µl of a 23-kDa recombinant baculovirus extract, partially purified by anion exchange chromatography according to chapter 3, method 3.9 (0.2µg of 23-kDa protein per assay). 10µl of the resuspended Ad2 DNA pellet were included in some of the assays and the final volume was made up to 50µl with 50mM Tris/HCl pH8 in each case.

Surprisingly, even after such a rigorous washing procedure the Ad2 DNA pellet still contained a small amount of protease activity which could be detected by incubation with MSGGAFSW for 30 minutes at 37°C (**fig.4.5a**, profile 3). This shows that the protease must bind very tightly to some other viral component, be it protein or DNA. Examination of the protein content of 10µl of the extract by SDS-PAGE reveals that a number of proteins are present including the hexon and mu, albeit at low concentrations (**fig.4.5c**). As mentioned previously, incubation of the partially purified baculovirus expressed protease with the peptide MSGGAFSW did not result in any digestion of the peptide (**fig.4.5a**, profile 2). When the peptide was incubated with the same baculovirus extract, along with the Ad2 DNA extract, however, 60% of the peptide was digested in 30 minutes (**fig.4.5a**, profile 4). It is concluded, therefore, that some component of the Ad2 DNA extract is able to activate the cloned protease and that the addition of the recombinant protease to the peptide/ DNA extract is analogous to the addition of recombinant protease to the ^{35}S Ad2ts1 extracts. In order to verify this the effect of DTT on the recombinant protease in the peptide assay was considered.

Assays were set up as above, with 10µl of the Ad2 DNA suspension, 30µl of the baculovirus 23-kDa protein extract and 2.5µg of MSGGAFSW being incubated for 1 hour at 37°C in the presence and absence of 2mM DTT. The results were identical to those obtained in the ^{35}S assay in that the cleavage of MSGGAFSW by the baculovirus expressed protease was reduced considerably by the inclusion of 2mM DTT (**fig.4.5b**, profiles 3,4). The parallel control experiment in which the peptide, MSGGAFSW was incubated with 10µl of disrupted, purified Ad2 in the presence and absence of 2mM DTT showed a 5 fold increase in protease activity in the presence of the thiol reagent (**fig.4.5b**, profiles 1,2).

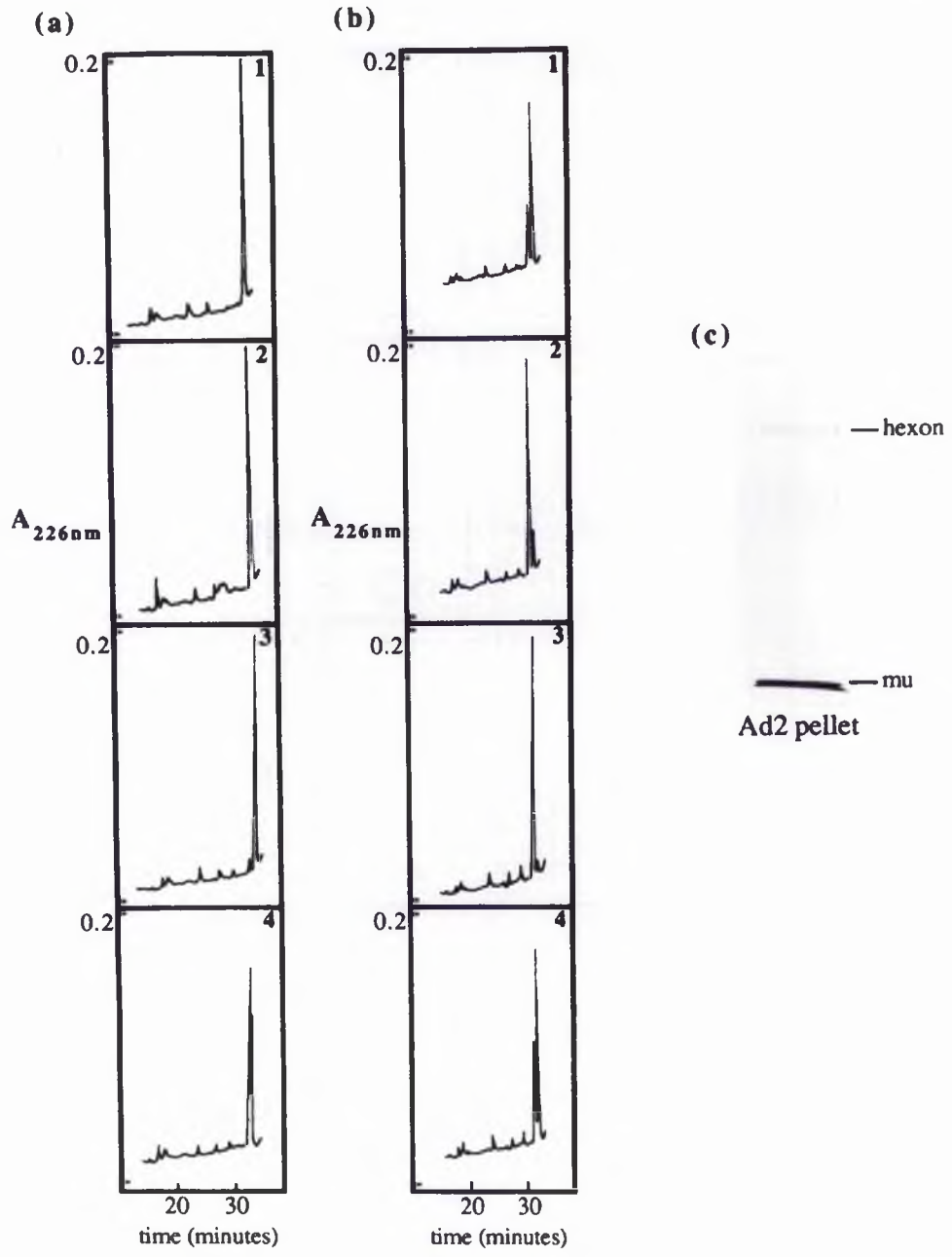


Fig. 4.5 Activation of the baculovirus expressed 23-kDa protein
The peptide assay was used to assess the activity of the recombinant baculovirus expressed 23-kDa protease as described in result 3. The HPLC profiles in (a) represent 30 minute incubations of 1.25μg of the peptide MSGGAFSW (1) alone; (2) with partially purified baculovirus 23-kDa protein; (3) with Ad2 DNA extract and (4) with baculovirus expressed 23-kDa protein and the DNA extract. The main peak in profiles 1, 2 and 3 is the undigested peptide MSGGAFSW. The HPLC profiles in (b) represent 1 hour incubations of 1.25μg of the peptide MSGGAFSW with (1) Ad2, 0mM DTT; (2) Ad2, 2mM DTT; (3) baculovirus 23-kDa protein and Ad2 DNA extract, 0mM DTT and (4) baculovirus 23-kDa protein and Ad2 DNA extract, 2mM DTT. The main peak in profile 3 is the digestion product AFSW. The Coomassie blue stained gel in (c) shows 10μl of the Ad2 DNA extract.

5 Binding of the Baculovirus Expressed Protease to DNA?

A cytoplasmic extract from 500ml of sf9 cells infected with a 23-kDa recombinant baculovirus stock was prepared and applied to a 5ml single stranded calf thymus DNA Sepharose column, equilibrated with 50mM Tris/HCl pH8, as described in method 1. The column was eluted with 10ml each of the same buffer made 0.2, 0.4 and 1M with respect to NaCl. 1ml fractions were collected and 10µl of each was analysed by SDS-PAGE, revealing that the 23-kDa protein did bind to the DNA column, but that it eluted at 0.2M NaCl (**fig.4.6a**). The fractions containing the 23-kDa protein were active when assessed for protease activity using the ^{35}S Ad2ts1 assay (results not shown) and the DNA-Sepharose column appears to provide a reasonable initial purification step in that it serves to concentrate the 23-kDa protein.

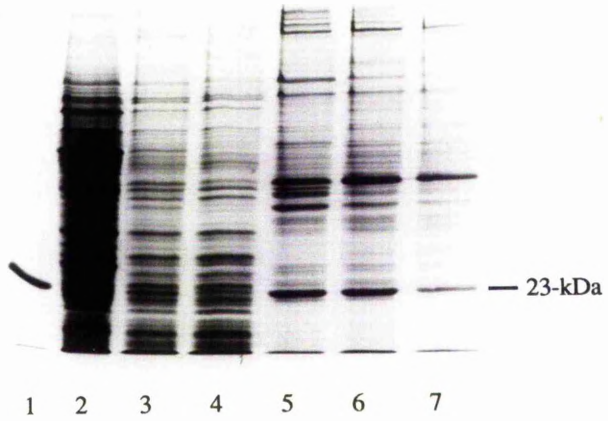


Fig. 4.6a Binding of the 23-kDa protein to DNA-Sepharose
A cytoplasmic extract from 500ml of sf9 cells infected with the pVL23K derived recombinant baculovirus stock was prepared and applied to a 5ml single stranded calf thymus DNA Sepharose column as described in method 1 and result 5. 10µl of the fractions eluted from the column were separated on a 12.5% SDS-polyacrylamide gel which was stained with Coomassie blue. *Lane 1*, purified *E.Coli* 23-kDa protein; *lane 2*, infected cell extract; *lane 3*, cytoplasmic extract; *lane 4*, column load and *lanes 5-7*, peak protein fractions eluted from the DNA Sepharose column with 0.2M NaCl.

To investigate the proposal that the 23-kDa protein binds specifically to a packaging sequence in the Ad2 genome (C.Anderson, personal communication), a ^{32}P labelled dsDNA probe was prepared comprising bases 1-466 of the Ad2 DNA as described in method 2. The source of 23-kDa protein was the pooled 0.2M NaCl fractions eluted from the single stranded

DNA Sepharose column. 5µl of the extract was included in each assay (0.1µg of 23-kDa protein) and tested for binding to the probe in the presence of 50mM EDTA, 1mM $MgCl_2$, 1mM $MnCl_2$, 0.1mM $ZnCl_2$ or 2mM DTT as described in method 2 (**fig.4.6b**). No binding was detected for any of the reactions, but the probe was retarded by the inclusion of 0.1µg of purified pTP as a control (lane 2, also see result 7). In parallel experiments, 5µl of the 23-kDa extract were incubated with 10µl of Ad2ts1 for 0-3 hours at 37°C. Complete digestion of pVII-VII took place within 30 minutes showing that the samples contained active protease.

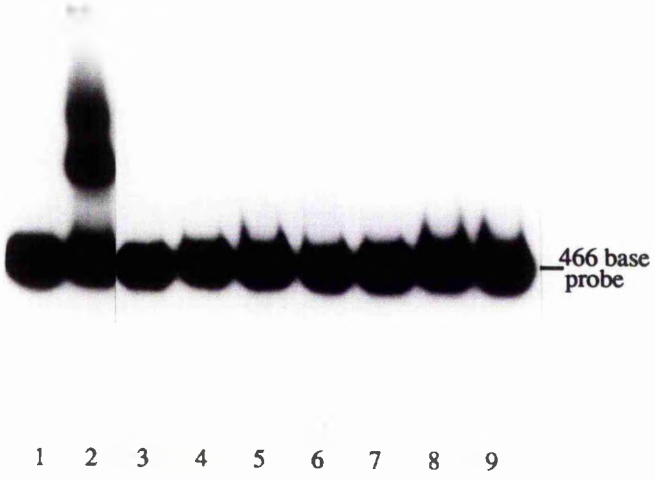


Fig. 4.6b DNA binding assays for the 23-kDa protease
The binding of the 23-kDa protein (source: 0.2M NaCl fractions eluted from the DNA-sepharose column; see fig.4.6b) to a ^{32}P labelled probe comprising bases 1-466 of the Ad2 genome was assessed by gel electrophoresis as described in method 2 and result 5. Assays were set up with 0.1ng of probe as described and the results are shown in the above autoradiogram. *Lanes 1* and *3*, probe alone; *lane* 2, probe with 0.1µg of pTP and *lanes 4-9*, probe with 0.1µg of partially purified 23-kDa protein. 50mM EDTA, 1mM $MgCl_2$, 1mM $MnCl_2$, 0.1mM $ZnCl_2$ and 2mM DTT were included in the assays shown in lanes 5, 6, 7, 8 and 9 respectively.

In summary, no conclusive evidence has been obtained to suggest that the Ad2 protease binds specifically to DNA. A systematic variation of the binding conditions, preferably using purified protein, however, would be required to eliminate such a possibility.

6 Expression of the pTP Using Recombinant Baculoviruses.

The gene for the Ad2 pTP was inserted into the vector pVL1393 and transfected into sf9 cells along with linearised AcRP23-lacZ DNA as described in method 4. The resulting recombinant virus was plaque purified and a 10^8 p.f.u. per ml stock was prepared. This was

used to infect sf9 cells in a 24 well plate and infected cells were harvested at 24, 48, 72 and 96 hours (chapter 3, method 3.8). A parallel infection was carried out using a recombinant pTP stock, prepared and supplied by L. Rogge. The cells from each well were harvested and analysed by SDS-PAGE as described in the methods section and the resulting Coomassie stained gel is shown in **fig.4.7a**. In each case good levels of expression were achieved at

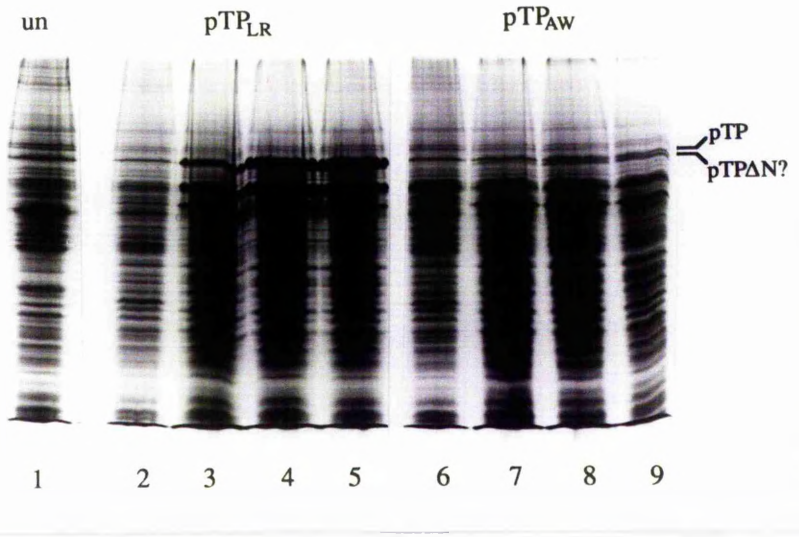


Fig. 4.7a Expression of the Ad2 pTP using recombinant baculoviruses
A Coomassie stained 12.5% polyacrylamide gel showing a time course infection of 24-96 hours for sf9 cells infected with 2 recombinant baculoviruses (for details see result 6). *Lane 1*, uninfected sf9 cells; *lanes 2-5*, cells infected with the recombinant baculovirus supplied by L.Rogge (LR) for 24, 48, 72 and 96 hours; *lanes 6-9*, cells infected with the recombinant baculovirus prepared in this project (method 4) for 24, 48, 72 and 96 hours.

48-96 hours post infection. The levels of expression of the pTP in the cells infected with the German stock, however, were at least 5 fold higher than in those infected with the recombinant pTP stock prepared as part of this project. Furthermore the pTP expressed using the German stock had a slightly slower mobility when analysed by SDS-PAGE.

To investigate this difference, the antigenicities of the 2 proteins were compared using 2 polyclonal antisera. The first was an antipeptide serum raised against a peptide corresponding to the N-terminal 15 amino acids of the Ad2 pTP which was prepared as described in method 3. The resulting antiserum reacted with the N-terminal pTP peptide; but not with an unrelated peptide, when assessed using an ELISA test and reacted well in Western Blots at a dilution of 1:500 (results not shown). The other antiserum used was raised in rabbits against purified pTP and was also used at a dilution of 1:500 (supplied by R.Hay and J.Bosher). **Fig.4.7b** shows

that whilst the N-terminal peptide antiserum reacts well with the pTP in extracts of cells infected with the German stock, it does not react with any proteins in cells infected with the stock prepared in this project (lanes 3 and 4). The pTP polyclonal antiserum, on the other

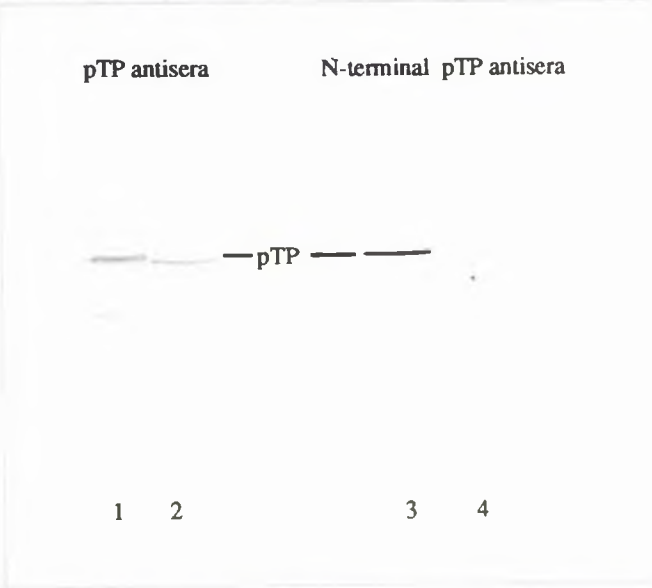


Fig. 4.7b Expression of the Ad2 pTP using recombinant baculoviruses
Western blots of 12.5% polyacrylamide gels probed with a polyclonal antiserum against the whole pTP and an antiserum raised against a peptide containing the N-terminal 15 amino acids of the Ad2 pTP. *Lanes 1* and *3*, sf9 cells infected with the recombinant baculovirus supplied by L. Rogge; *lanes 2* and *4*, sf9 cells infected with the recombinant baculovirus prepared in this project.

hand, reacted with proteins in cells infected with both stocks, although the mobility of the major band in the extracts prepared using the German stock was slightly slower (lanes 1 and 2). The other lower molecular weight bands reacting with the pTP polyclonal antiserum are, presumably, degradation products of the pTP.

The conclusion from this analysis is that the protein expressed using the recombinant pTP virus, prepared as part of this project, is not the complete pTP. Its failure to react with the N-terminal peptide antiserum suggests that the N-terminal 15 amino acids are missing. The apparent molecular weight of the protein, however, is only slightly less than that of the pTP expressed using the German stock and it does react well with the polyclonal antisera raised against the pTP. The most likely explanation for this is that an error in the cloning has eliminated the initiation codon for the pTP. As a result the protein is translated, using the AUG coding for Met_{19} as the initiation codon. In effect this means that the first 18 amino acids of the pTP are missing. To confirm this assumption it will be necessary to sequence the BamHI

insert containing the putative pTP cDNA prepared in this project.

7 Purification of the Baculovirus Expressed pTP

The pTP was purified to near homogeneity from a nuclear extract of sf9 cells, infected with the recombinant baculovirus stock provided by L. Rogge, using a combination of affinity and hydrophobic interaction chromatography.

The nuclear extract was prepared and applied to a single stranded DNA-Sepharose column as described in method 5. The pTP bound to the column and was eluted with a buffer made 0.6M with respect to NaCl. The fractions containing pTP were pooled and KCl was added to give a salt concentration of 1.2M before applying a tenth of the sample onto an FPLC phenyl Superose column. The pTP bound tightly to the column and eluted at low salt concentrations (see **fig.4.8a**). The use of a hydroxyapatite column instead of the phenyl Superose column, in a method described by Temperley and Hay (1992), was investigated, but the purity of the protein appeared less (results not shown).

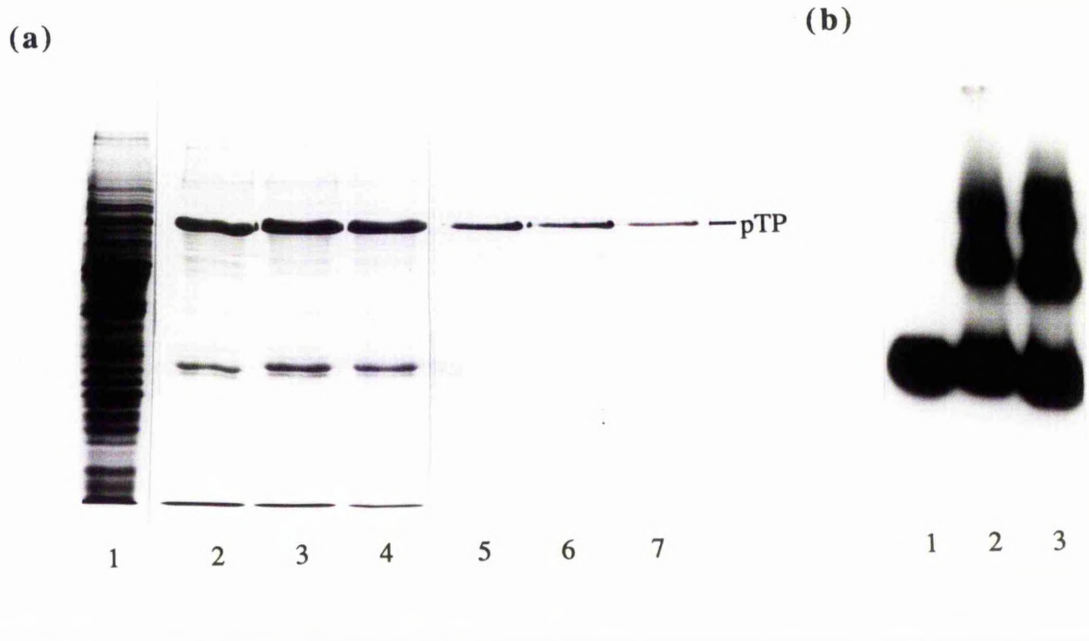


Fig. 4.8 Purification of the baculovirus expressed pTP
sf9 cells were infected for 72 hours with the pTP recombinant baculovirus supplied by L. Rogge, harvested and purified using DNA-Sepharose and phenyl-Superose columns as described in method 5 and result 7. The Coomassie stained 12.5% polyacrylamide gel in (a) shows *lane 1*, total cell extract; *lanes 2-4*, 0.6M NaCl fractions eluted from the single stranded DNA-Sepharose column; *lanes 5-7*, fractions eluted from the phenyl-Superose column. The autoradiogram in (b) shows a gel electrophoresis DNA binding assay carried out to test the binding of the purified pTP to base pairs 1-466 of the Ad2 genome. 0.1ng of a ^{32}P labelled DNA probe were incubated *lane 1*, alone; and with 0.1μg of pTP purified using *lane 2*, a hydroxyapatite column and *lane 3*, a phenyl-Superose column.

0.1μg of pTP purified using a ssDNA-Sepharose column and the phenyl Superose column or the hydroxyapatite column were tested for their ability to bind to the ^{32}P labelled 466 base pair probe corresponding to bases 1-466 of the Ad2 genome (see result 5). The resulting gel retention assay showed that the pTP prepared using the phenyl-Superose column bound more of the probe (**fig.4.8b**).

8 Coinfections Using Recombinant pTP and 23-kDa Baculovirus Stocks

Having shown that the 23-kDa protein is the Ad2 protease and that its activity is controlled in some way, the question as to whether another adenovirus protein is required for activity must be addressed. The coinfection of sf9 cells with recombinant pTP and 23-kDa baculovirus stocks (both 10^7 p.f.u. per ml) was carried out in an attempt to answer this.

The coinfections were carried out essentially as described for single infections in chapter 3, method 3.8. sf9 cells, in a 24 well plate, were infected with 200μl of various ratios of the pTP:23-kDa stocks ranging from 10:0 to 0:10. The infection time in each case was 72 hours and the cells were harvested, disrupted and analysed by SDS-PAGE as described in chapter 3. Four identical SDS gels were prepared and analysed as follows. One was stained with Coomassie blue (**fig.4.9a**), whilst the proteins from the other 3 were transferred onto nitrocellulose paper by Western blotting and probed with the polyclonal antisera raised against the 23-kDa N-terminal peptide (**fig.4.9b**), the pTP (**fig.4.9c**) and the pTP N-terminal peptide (**fig.4.9d**). The results show clearly that when the cells were coinfected with the pTP and 23-kDa stocks digestion of the pTP took place. The main products of the digestion appeared to be 2 bands with an apparent molecular weight of around 60-kDa (**fig.4.9c**) and a smaller protein of apparent molecular weight 15-kDa (**fig.4.9d**). The latter reacts with the N-terminal pTP antiserum and is the size that would be expected for the N-terminal fragment generated by the cleavage of the pTP to the iTP. Surprisingly the protein appears to be relatively stable in the insect cells and can be readily visualised by Coomassie staining (**fig.4.9a**). It is also noted that the antiserum raised against the purified pTP does not react with the 15-kDa fragment (**fig.4.9c**). The doublet present at 60-kDa presumably corresponds to the 2 iTP species and supports the proposal in chapter 1 that the Ad2 pTP is cleaved at 2 sites separated by only 8 amino acids to give the iTP.

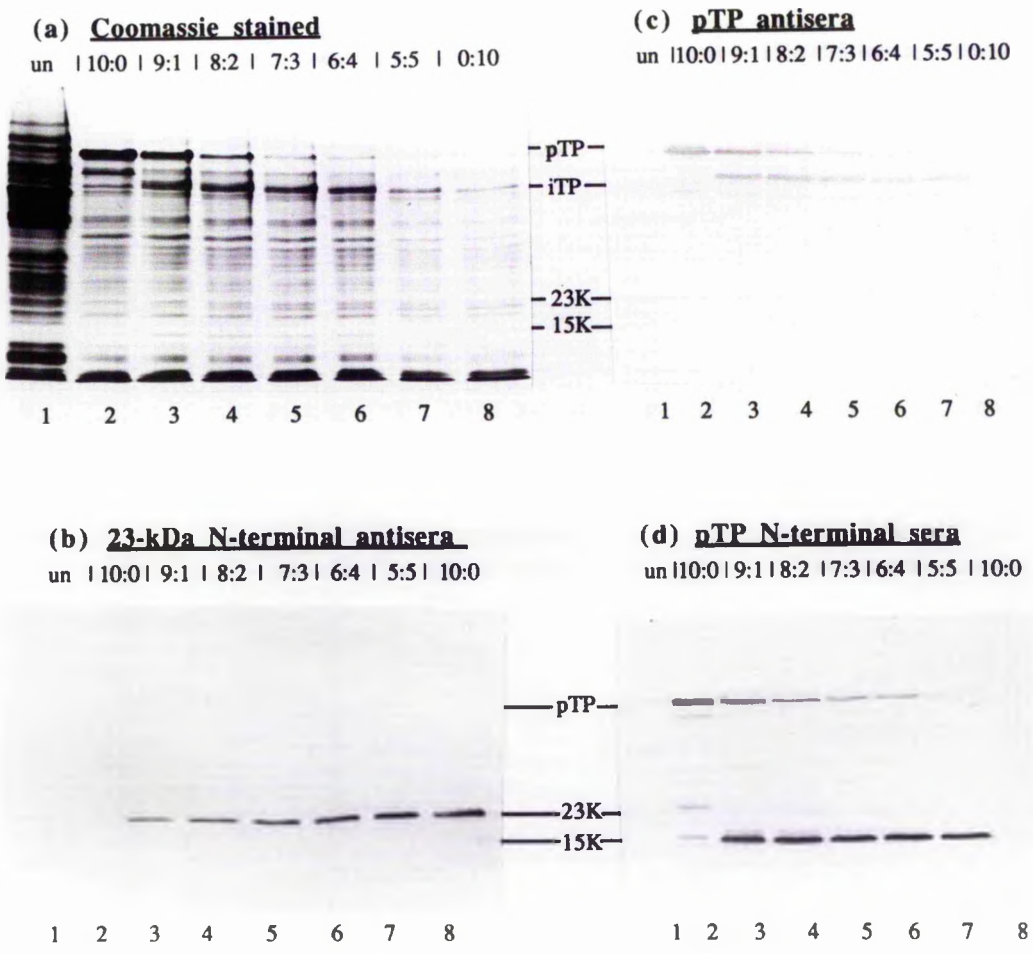


Fig. 4.9 Coinfections of insect cells with recombinant pTP and 23-kDa baculoviruses
The recombinant baculovirus expressing the full length pTP and the the pVL23K derived recombinant baculovirus were coinfected into sf9 cells. The ratio of the pTP:23-kDa virus stock was varied from 10:0 to 0:10 as indicated. Cells were harvested at 72 hours post infection, disrupted and samples were analysed on 4 identical 12.5% SDS-polyacrylamide gels. Gel (a) was stained with Coomassie blue, whilst gels (b) to (d) were electroblotted onto nitrocellulose paper and analysed by Western blotting. The antisera used to probe the filters are as indicated above (see result 8 for details).

A similar coinfection was set up using the recombinant pTP baculovirus stock prepared as part of this project. As discussed above it is believed that the stock expresses the pTP minus its N-terminal 18 amino acids. Cells coinfected with this pTP stock and the 23-kDa recombinant baculovirus were harvested and analysed by SDS-PAGE as described previously. **Fig.4.10** shows the Western Blot of the resulting cell extracts probed with the polyclonal antiserum raised against the pTP. As for the full length pTP, a doublet band at 60-kDa is present in the coinfected cells; but not in the uninfected, pTP or 23-kDa baculovirus infected cells. This suggests that the pTP lacking its N-terminal 18 amino acids is processed normally by the Ad2 protease.

Examination of **fig.4.9c** and **4.10** reveals that in each case a protein is present with an apparent molecular weight of 55-kDa. Unfortunately, there is also a band of similar mobility in cells infected with the recombinant pTP and 23-kDa baculoviruses alone, albeit in lesser amounts. Thus although it has been shown that the pTP is readily processed to the iTP in insect cells that have been coinfected with pTP and 23-kDa stocks, evidence that it is efficiently processed to the TP in sf9 cells is lacking.

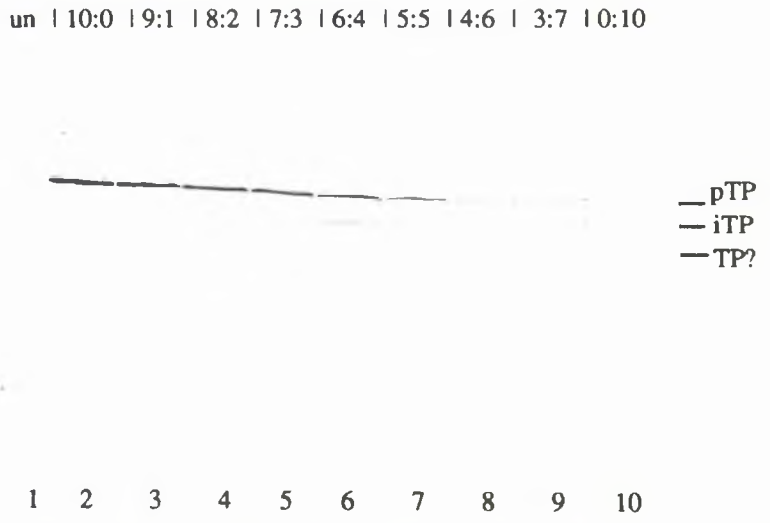


Fig. 4.10 Coinfections of sf9 cells with 23-kDa and pTPΔN baculovirus stocks
sf9 cells were infected with the recombinant pTP and 23-kDa baculovirus stocks prepared in this project. Cells were infected with ratios of the 2 virus stocks ranging from 10:0 to 0:10 as indicated and harvested at 72 hours post infection. The disrupted cells were analysed by SDS-PAGE and Western blotting as described in result 8. The Western blot of the samples probed with the polyclonal antiserum raised against the whole pTP is shown above.

9 Digestion of pTP *in vitro* by Baculovirus Expressed Ad2 Protease

The pTP was purified from infected sf9 cells using a single stranded DNA-Sepharose column and a phenyl Superose column as described above. The 23-kDa protease was partially purified from the cytoplasmic extract of infected sf9 cells using an anion exchange column (see chapter 3, method 3.9). 0.2μg of the purified pTP was incubated with 0.2μg of the baculovirus 23-kDa protein for 12 hours at 37°C in the presence and absence of 10μl of the resuspended Ad2 DNA pellets (see result 4). Two different Ad2 DNA extracts were used, one that was washed extensively with 4M NaCl and 4M urea and sonicated as described previously. The other extract was prepared by simply resuspending the DNA/protein pellet from the 2M NaCl, 2M urea treated Ad2 in 50mM Tris/HCl pH8 by vortexing. The second

extract contained a considerable amount of intrinsic protease activity.

The samples were analysed by SDS-PAGE and Western blotting using the polyclonal antiserum raised against the pTP (**fig.4.11**). The results show that incubation of the pTP with the partially purified baculovirus expressed 23-kDa protein gave virtually no digestion of the pTP after an overnight incubation (**fig.4.11a**, lane 2). Incubation with the washed Ad2 DNA extract gave partial digestion of the pTP to the iTP (lane 3), whilst incubation of the pTP with the baculovirus expressed 23-kDa protein and the washed DNA extract resulted in the near complete processing of the pTP to the iTP (lane 4). Indeed a band corresponding to the TP was also present. When an Ad2 DNA extract with more intrinsic protease activity was used the same effect was observed; but more processing took place and a strong band appeared at 55-kDa corresponding to the TP (**fig.4.11b**).

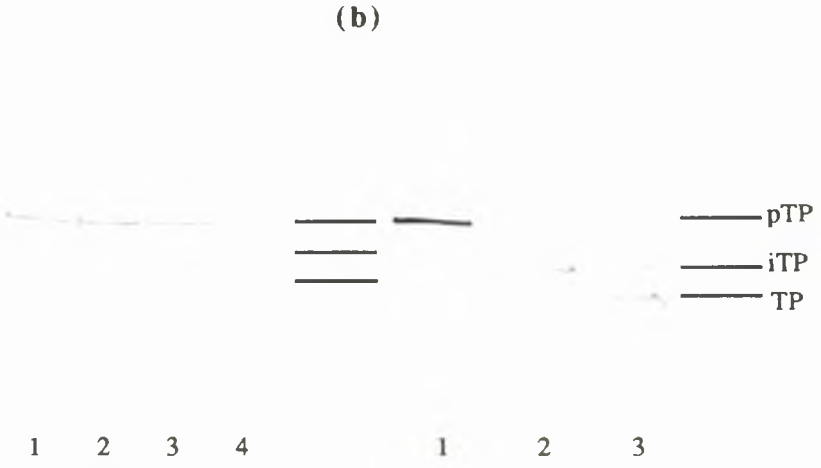


Fig. 4.11 Digestion of baculovirus expressed pTP *in vitro*
Full length pTP purified from infected sf9 cells was incubated *in vitro* with partially purified baculovirus expressed 23-kDa protease as described in result 9. The incubations were for 12 hours at 37°C and samples were analysed by SDS-PAGE and Western blotting. The polyclonal antiserum raised against the whole pTP (1:500 dilution) was used to probe the samples. Resuspended Ad2 DNA pellets were added to some of the assays. The DNA pellet used in (a) had been washed extensively and so contained less intrinsic activity than the DNA pellet used in (b). Samples were incubated as follows. **(a)** *Lane 1*, pTP alone; *lane 2*, pTP and 23-kDa protein; *lane 3*, pTP and Ad2 DNA extract and *lane 4*, pTP, 23-kDa protein and DNA extract. **(b)** *Lane 1*, pTP alone; *lane 2*, pTP and DNA extract; *lane 3*, pTP, 23-kDa protein and DNA extract.

10 *In Vitro* Replication Assays

Initiation reactions were set up as described in method 6. **Fig.4.12a**, lane 1 shows that the initiation reaction has taken place and that the ^{32}P dCMP is covalently linked to the 80-kDa

pTP. Incubation of the sample with disrupted, purified Ad2 for 1-12 hours (lanes 2-5) shows that the pTP is progressively digested to the 55-kDa TP via the 60-kDa intermediate. It is noted that even after a 12 hour incubation only 50% of the precursor has been completely processed. In parallel experiments in which ^{35}S Ad2ts1 extracts were incubated with disrupted purified virus, complete digestion of pVII to VII took place within 2 hours (results not shown). Although the substrate concentrations could not be determined in the assays, it would appear that the processing of the iTP to the TP is considerably slower than the processing of the other viral protein substrates.

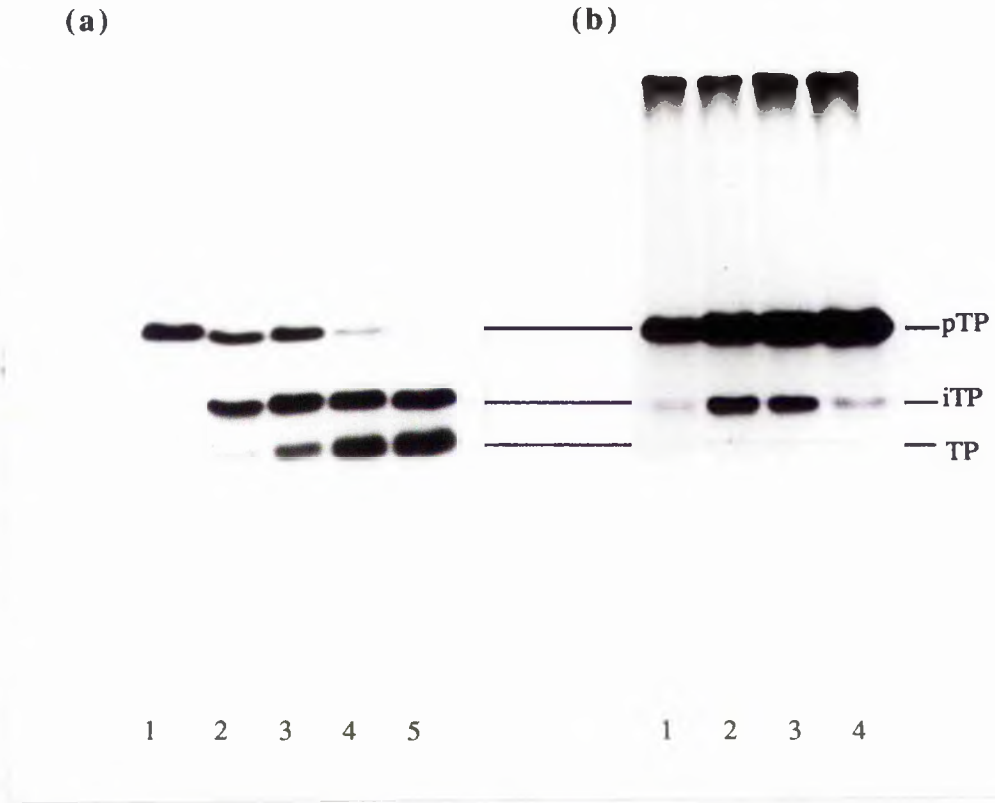


Fig. 4.12 ***In vitro*** **replication assays - initiation reactions.**
Autoradiograms of 10% polyacrylamide gels. **(a)** Initiation reactions were carried out as described in method 6 and the resulting samples were incubated with 20µl of disrupted, purified Ad2 at 37°C for *lane 1*, 0 hours; *lane 2*, 1 hour; *lane 3*, 2 hours; *lane 4*, 6 hours and *lane 5*, 12 hours. **(b)** Initiation reactions were carried out as described in method 6 in the presence of *lane 1*, 0mM DTT; *lane 2*, 0.5mM DTT; *lane 3*, 1mM DTT and *lane 4*, 5mM DTT.

Initiation reactions were set up as described in method 6 in the presence of 0, 0.5, 1 and 5mM DTT. The efficiency of the initiation reaction was shown to increase with DTT concentration. The amount of labelled pTP processed during the initiation reaction, however, was greatest at 0.5-1mM DTT (**fig.4.12b**) in agreement with the results obtained in chapter 2. The source of protease in these experiments were the Ad2 cores, which have been shown to contain the 23-kDa protein (Leith *et al.*,1989; Anderson,1990).

DISCUSSION

In the course of this chapter the control of the Ad2 protease and the processing of the pTP have been investigated. In the main, the results produced are of a preliminary nature and form the basis for future work to study the regulation of proteolysis and replication during adenovirus infection. Nevertheless, a number of significant observations have been made and these will be discussed in turn.

In chapter 2 the effects of a range of inhibitors on the Ad2 protease were considered using the ^{35}S and peptide assays. In the systems used, however, the inhibitors were not pre-incubated with the purified protease and so the possibility that they were acting on some other viral component could not be ruled out. **Table 4.1** summarises the results of a series of assays in which purified *E.Coli* expressed protease was incubated with a range of inhibitors, at their effective concentrations, prior to incubations with ^{35}S Ad2ts1 extracts. The results are virtually identical to those obtained in chapter 2 (**Table 2.6**) and confirm that the inhibition profile of the Ad2 protease fits better into the cysteine class than into the serine class. In view of this it was surprising to note that 2mM DTT inhibited the recombinant protease (**Table 4.1**), in contrast to its effect in the peptide assay where protease from disrupted, purified Ad2 was activated at least 5 fold by 0.5-2.5mM DTT (chapter 2). It was decided to investigate this anomaly further by studying the effect of a range of DTT concentrations in the ^{35}S Ad2ts1 assay using 2 sources of protease; the purified recombinant *E.Coli* expressed protein and purified disrupted Ad2. According to chapter 3, the protease in the virus is already activated; but the recombinant 23-kDa protein is inactive. **Fig.4.4** shows that whilst 1mM DTT severely inhibits the recombinant protease, the Ad2 protease is only affected by the thiol reagent at concentrations of 5-10mM. It is proposed, therefore, that whilst DTT has a positive effect on protease activity if the enzyme is present in an active form, it inhibits the activation step. This work provides further support for the proposal that the digestion by the recombinant Ad2 protease is a 2 step process. Indeed, these results are consistent with an observation made by Bhatti and Weber (1979b). These authors reported that the viral protease, but not the

nuclear Ad2 protease was activated by DTT. From this it would appear that most of the protease in the nuclei of infected cells is present in an inactive form, but that the enzyme packaged in the virion is fully active.

The next question to be addressed is by what mechanism is the Ad2 protease activated? Its high isoelectric point (predicted value, pH8.6) and tight association with the viral cores (Anderson,1990; chapter 2) have led to the proposal that the enzyme binds directly to DNA. Indeed it has been suggested that the purified *E.Coli* expressed 23-kDa protein binds to a packaging sequence located at base pairs 190-300 in the Ad2 genome (C.Anderson, personal communication). The possibility arises, therefore, that the binding of the protease to a specific DNA sequence could be sufficient to activate it. To test this hypothesis, the recombinant *E.Coli* and baculovirus proteases were tested for their ability to cleave the peptide substrate, MSGGAFSW, in the presence of Ad2 DNA, base pairs 1-600 of the Ad2 genome, single and double stranded calf thymus DNA and calf thymus DNA-Sepharose beads. The assays were carried out in the presence and absence of 1mM ATP; but no cleavage of the peptide could be detected even after incubations for 6 hours at 37°C. It was concluded, therefore, that DNA alone is not sufficient to activate the recombinant protease.

The results in **fig.4.5a** show that an Ad2 DNA pellet prepared by a series of 4M urea, 4M NaCl washes did activate the recombinant baculovirus protease. Surprisingly, there was still a small amount of intrinsic protease activity in the resuspended DNA pellet, emphasising just how tightly the protease is associated with the viral cores. Nevertheless, the addition of the partially purified baculovirus expressed 23-kDa protein to the DNA extract resulted in a 10 fold increase in the rate of cleavage of MSGGAFSW. The Coomassie stained gel in **fig.4.5c** reveals that the DNA pellet contains a number of proteins at relatively low concentrations. It seems likely that one of these, perhaps in association with the viral DNA, is sufficient to activate the protease.

The effect of DTT on the activity of the recombinant protease in the peptide assay was considered and the results are shown in **fig.4.5b**. Profiles 1 and 2 act as controls and show that the protease from disrupted Ad2 is activated 5 fold by 2mM DTT in the peptide assay. In stark contrast, the activity of the baculovirus recombinant protease was severely inhibited by

2mM DTT in the peptide assay (profiles 3 and 4). These findings are consistent with the results shown in **fig.4.4** where ^{35}S Ad2ts1 was used as the substrate. It is proposed, therefore, that the use of Ad2 DNA extracts along with the peptide substrate MSGGAFSW provides an assay system analogous to the ^{35}S Ad2ts1 assay.

Despite the fact that purified Ad2 DNA was not able to activate the recombinant protease, the binding of the enzyme to specific DNA sequences could not be ruled out and so the DNA binding properties of the baculovirus expressed 23-kDa protein were considered. The latter was shown to bind to a calf thymus ssDNA-Sepharose column and to elute at 0.2M NaCl (**fig.4.6a**). The affinity of the protein for non-specific DNA sequences is low; but this is not unusual for sequence specific DNA binding proteins (Sorger *et al.*,1989). A preliminary gel retention assay was set up to test the binding of the partially purified baculovirus expressed 23-kDa protein to base pairs 1-466 of the Ad2 genome, but no binding was detected (**fig.4.6b**). At this stage no evidence can be presented to suggest that the adenovirus protease binds specifically to DNA, but until a wider range of conditions are tested, the possibility cannot be ruled out.

In conclusion, it seems clear that the adenovirus protease is activated in some way, but the mode of activation remains elusive. The inhibitory and activatory effects of DTT may help to provide the answer and perhaps a thiol exchange mechanism is involved. Alternatively, a metal ion such as zinc might be required at low concentrations in the activation step and DTT could interfere with this. Recently it has been shown that *trans* cleavages by the poliovirus 2A protease are inhibited by 1,10 phenanthroline and one proposal is that the enzyme contains a metal binding site (Yu and Lloyd,1992). The activity of the recombinant baculovirus 23-kDa protease in the ^{35}S Ad2ts1 assay, however, was not found to be affected by the presence of 10mM 1,10 phenanthroline which is an effective chelator of zinc ions (**table 4.1**). It is clear that the protease is closely associated with a viral protein and/or viral DNA and this association alone could be sufficient to activate the enzyme. On the other hand, the enzyme may be activated first and then bind to the protein/DNA in such a way as to stabilise it in an active form. When removed from its association with the viral component the protease rapidly reverts to its inactive form, unless extracted in the presence of high salt (chapter 2). The apoenzyme

might be unable to associate with the protein/DNA unless it is activated first. Finally in this regard phosphorylation of the protease, as suggested by Chatterjee and Flint (1987), remains a possibility and the baculovirus system should provide the means to examine the 23-kDa protein for post-translational modifications in the future.

Processing of the Ad2 pre-terminal protein.

A recombinant pTP baculovirus was kindly provided by L. Rogge and coinfections of the 23-kDa and pTP viruses were carried out in sf9 cells. The results in **fig.4.9a-d** show that the pTP is processed by the 23-kDa protease in the insect cells, with the major digestion product being the iTP. The iTP appeared as a doublet when analysed by SDS-PAGE in agreement with the predictions made in chapter 1 that there are 2 potential cleavage sites, Phe_{175}-Gly_{176} and Arg_{183}-Gly_{184}, separated by only 8 amino acids in the Ad2 pTP. The 15-kDa N-terminal fragment generated by the processing of the pTP->iTP was stable in the insect cells, reacting strongly with the antiserum raised against the N-terminal 15 amino acids of the pTP. The possibility that this fragment has a functional role in the adenovirus infectious cycle is certainly worth considering in the future. It was also noted that the polyclonal antiserum raised against the full length pTP did not react with the 15-kDa N-terminal fragment. Presumably this is because the N-terminal domain was not surface exposed in the purified pTP used to immunise the rabbits.

In the course of this work an attempt was made to prepare a recombinant baculovirus expressing the full length pTP. As discussed in the results section, a virus was constructed that expressed a protein reacting with the polyclonal antiserum raised against the pTP. The mobility of this protein, however, was slightly faster than the full length pTP and it did not react with the N-terminal antipeptide serum. It was concluded that the pTP expressed by the recombinant baculovirus, prepared in this project, was probably being translated using the AUG codon for Met_{19} as its initiation codon. As a result the protein synthesised lacks its N-terminal 18 amino acids. It has been shown previously that the N-terminal 18 amino acids are essential for the initiation of replication (Pettit *et al*, 1989) and so the pTP lacking its N-terminus could act as a useful control in replication assays; but its identity must be confirmed by DNA sequencing of the putative pTP gene in the parent plasmid pVLpTP. Coinfections

were set up for the 23-kDa and 'defective' pTP baculovirus stocks as described previously. **Fig.4.10** reveals that the pTP lacking its N-terminus was digested by the protease to give the iTP in a similar fashion to the full length pTP (**fig.4.9c**). Thus despite the fact that the pTP requires its 18 N-terminal residues to function in replication, it would appear that the protein lacking amino acids 1-18 is processed normally by the protease.

One puzzling feature of the coinfections was that most of the processing of the pTP only went as far as the iTP and little, if any, TP was generated. Perhaps this is because, as was suggested in chapter 1, the putative cleavage site giving rise to the TP has a valine at P_1' and is a poor substrate. Additional evidence that the iTP is a poor substrate for the protease comes from the *in vitro* replication assays shown in **fig.4.12a**, where even after a 12 hour incubation with disrupted purified Ad2 more than 50% of the iTP remained unprocessed. Indeed, it is possible that the different rates of cleavage of the pTP -> iTP and iTP->TP could be significant, with the 2 cleavages being temporally separated in the infectious cycle.

The coinfection experiments also show that no other adenovirus protein or DNA sequence is required for the processing of the pTP->iTP by the 23-kDa protein. This means that the cofactor is probably cellular, although the possibility that the pTP, itself, or that a baculovirus component can act as the cofactor cannot be ruled out. The pTP was purified by a combination of affinity and hydrophobic interaction chromatography using DNA-Sepharose and phenyl-Superose columns. The resulting protein had DNA binding properties consistent with those described by Temperley and Hay (1992) using an alternative purification scheme. The exposure of the pTP to high salt (1.2M KCl) in the hydrophobic interaction step was deemed to be beneficial as it should serve to remove any fragments of DNA or other cellular components that might bind to it non-specifically.

Incubation of the peptide MSGGAFSW with purified pTP and the partially purified baculovirus expressed 23-kDa protease, in the presence or absence of purified Ad2 DNA, did not result in any cleavage of the peptide (results not shown). Furthermore, a 12 hour incubation of the purified pTP with the partially purified baculovirus expressed protease gave negligible processing of the pTP (**fig.4.11a**, lane 2). These results show that the pTP, on its own, is not the cofactor. The addition of a crude Ad2 DNA extract, however, resulted in the

complete processing to the iTP and a clear band at the mobility expected for the TP was also present (**fig.4.11a**, lane 4). The Ad2 DNA extract on its own gave partial processing of the pTP->iTP (lane 3). The enhanced cleavage of the iTP->TP in the presence of the Ad2 DNA extract could be explained by a conformational change in the pTP/iTP upon binding to the DNA. The results from the peptide assay (**fig.4.5a**), however, indicate that some other factor must be involved since the conformation of the peptide will not be changed by the presence of the DNA. The results from the coinfections show that although processing of the pTP does take place in insect cells it is probably not as efficient as in the adenovirus infected cells. This despite the fact that the level of expression of the 23-kDa protein in the insect cells is considerably higher. Thus it seems that the activation process can take place in baculovirus infected insect cells; but that it is much more efficient in the course of adenovirus infections. In the future it would be interesting to construct recombinant baculoviruses expressing other substrates of the Ad2 protease and to test for proteolysis of the substrates in coinfections with the recombinant 23-kDa virus. Fractionation of uninfected and wild type AcNPV infected sf9 cells may also help to identify the cofactor.

Finally, *in vitro* replication assays were set up using Ad2 cores as the source of template. The initiation reactions were carried out in the presence of 0-5mM DTT and the processing of the pTP to the iTP was found to optimal at 0.5-1mM DTT (**fig.4.12b**), consistent with the effects reported earlier for the cleavage of the peptide MSGGAFSW by protease from purified Ad2 (chapter 2). The source of protease in the initiation reactions was concluded to be the Ad2 cores and the stimulatory effect of DTT is consistent with the protease being present in an active form. This work shows that the initiation reaction provides an alternative assay system for the Ad2 protease and in the future the use of Ad2 complex (DNA+TP) or a linearised plasmid containing the origin of replication as the template would provide a defined system and eliminate the endogenous protease. It would be interesting to compare the efficiencies of pTP-DNA and TP-DNA as templates in such systems. The purification of the iTP, TP and 15-kDa N-terminal fragment of the pTP would enable the effects of each on initiation and elongation to be considered. It would also be possible to carry out N-terminal sequencing of the iTP and TP to confirm the cleavage sites predicted in chapter 1.

In the course of this chapter a number of systems have been set up to study proteolytic processing by the adenovirus protease, both *in vivo* and *in vitro*. In the future it should be possible to use them to answer the all important question of how the protease is activated. The processing of the pTP by the 23-kDa protein in insect cells also opens the way for the study of the effects of limited proteolysis of the pTP on adenovirus DNA replication.

CONCLUSIONS

The aim of this project was to further characterise the adenovirus protease. In the course of the work the substrate specificity of the enzyme has been defined and cleavage sites have been predicted in all the known substrates of the protease (chapter 1). It has been shown that the enzyme contains an essential thiol group and that it is probably a member of the cysteine protease class; although it differs from the classical papain-like enzymes (chapter 2). The L3 23-kDa protein has been purified from Ad2 and shown to be the protease and to be active as a monomer. Furthermore antipeptide sera raised against the N- and C-termini of the protease have been used to demonstrate that it is active in an intact form contrary to previous reports that it is proteolytically activated (chapter 3). The enzyme has been expressed in *E.Coli* and insect cells and the recombinant protease was able to cleave pVII-VII in a crude assay system containing cell extracts; but was inactive in the peptide assay developed as part of this project. A recombinant baculovirus expressing one of the substrates of the protease, the pre-terminal protein was constructed, although the expressed protein was shown to lack the 18 N-terminal amino acids. Coinfections of this recombinant baculovirus, or a virus expressing the full length pTP (gifted by L.Rogge), with the 23-kDa recombinant baculovirus were carried out in insect cells. The results of these experiments showed that the pTP was processed efficiently to the iTP; but not to the TP. One of the main conclusions drawn from the results in chapters 2-4 was that although the 23-kDa protein is the Ad2 protease some other factor is required for activity and it was suggested that the L3 23-kDa protein is synthesised as an apoenzyme. These results have already been discussed in chapters 1-4.

In the main introduction the infectious cycle of the adenovirus was described with particular attention being paid to the virus encoded proteins. The essential role that proteases play in the life cycles of other viruses was also discussed. In this section it is hoped to put the results of this project into context, with particular emphasis being put on future work.

<u>The Role of the Adenovirus Protease in the Lytic Cycle</u>

Six of the adenovirus structural proteins, namely pIIIa, pVI, pVII, pVIII, the L2 11-kDa protein and the pTP had previously been shown to be substrates of the Ad2 protease

(Anderson *et al.*,1973; Boudin *et al.*,1980; Weber and Anderson,1988). With the exception of the pTP, a product of the E2B region, all these proteins are encoded by the late transcription unit and are synthesised late in infection. This observation, taken along with the fact that virus particles are assembled in cells infected with Ad2ts1, at non-permissive temperatures, has led to the assumption that the adenovirus protease is only required for a maturation event late in the infectious cycle. Set against this, however, it has been shown that coinfection of sf9 cells with recombinant baculoviruses expressing the 23-kDa protein and the pTP leads to efficient processing of the pTP to the iTP in the insect cells. Thus it would appear that assembly of virus particles is not a prerequisite for protease activity. In view of this it is worth considering potential roles for the protease at earlier stages of the infectious cycle and in particular during uncoating, replication, the switch from early to late infection and virus assembly. In chapter 1 potential cleavage sites were identified in a number of viral proteins that had not previously been shown to be substrates of the protease (**table 1.5**). It is emphasised, however, that at this stage there is no evidence to suggest that any of these sites are cleaved *in vivo* and future investigations will be required.

Uncoating of Virions

In chapter 1 it was postulated that a conformational change in the virion coat, shown to take place in low pH, could expose the predicted cleavage sites in the hexon and penton proteins to the inside of the virion where the protease is located. Limited proteolysis of the coat proteins could be a key step in the uncoating of adenovirus particles, a process that is poorly understood. If the hexon and/or the penton are cleaved it is likely that only the correctly folded proteins in the mature virions would be processed at a precise point early in the infectious cycle and so evaluation of this hypothesis might prove difficult.

Switch from Early to Late Infection

The identification of potential cleavage sites in a number of the viral early proteins including the E1B 55-kDa protein leads to the suggestion that cleavage of one or more of these proteins could play a role in controlling the progress of the lytic cycle. Set against this, however, is the observation that infections of HeLa cells with Ad2ts1 at non-permissive temperatures appear to proceed normally until late stages in the infectious cycle (Weber,1976).

DNA Replication

The possible role of the Ad2 protease in DNA replication through the cleavage of the pTP was discussed in detail in chapter 4. Further studies are required to establish the structural and functional differences between the pTP, iTP and TP, with one possibility being that the cleavage of the pTP by the protease is required to separate the viral DNA from the nuclear matrix (Schaack *et al.*,1990). Also of interest in this regard is the observation that the sequences of the Ad2 polymerase protein (**table 1.5**) and nuclear factor I (results not shown) both contain potential cleavage sites. The site in the polymerase is conserved in all the serotypes in which the protein has been sequenced and contains an arginine at the P_1' position. To date no peptides with an arginine at P_1' have been tested, so it remains to be established whether or not a charged residue is permissible at this position. The locations of the 2 sites in NF1 (MEGG-I and LNGS-G centred at positions 320 and 427, respectively) are interesting in that cleavage at both of these sites would leave the DNA binding domain intact. Recent studies have shown that the N-terminal 222 amino acids of NFI are sufficient for the stimulation of Ad2 DNA replication (Bosher *et al.*,1991) and so the possibility exists that the Ad2 protease processes NFI to convert it from a cellular transcription to a viral replication factor. It has been suggested that the poliovirus 3C protease is responsible for the processing of transcription factor TFIIIC and it is possible that cellular proteins are processed by other viral proteases as part of the infectious cycle.

Alternatively the Ad2 protease might be required to switch off DNA replication, perhaps by breaking up the replication complex and detatching the viral DNA from the nuclear matrix.

Maturation of Virus Particles

Few studies have considered the conformational and functional changes of the individual structural proteins during maturation and the identification of the cleavage sites in all the established substrates of the protease should enable such work to be undertaken in the future.

Characterisation of the Ad2 Protease

Studies have been facilitated by the development of an assay for the Ad2 protease using

the synthetic peptide MSGGAFSW as a substrate. If kinetic analysis of potential inhibitors is to be carried out, however, it would be desirable to have a chromogenic substrate. With regards to inhibitors it was noted that peptide substrate analogues lacking the P_4 residue, FGGAKK and SGGAFSW, inhibited the cleavage of pVII to VII in the ^{35}S assay. Although these observations require confirmation, it is possible that they could form the basis for the design of a specific inhibitor of the Ad2 protease. A number of divalent ions including copper and zinc were shown to inhibit the protease at concentrations as low as 10μM. The fact that the inhibition could be completely reversed by the subsequent addition of EDTA or DTT to the protease suggests that the divalent ions are binding to a thiol group in the protease, perhaps at the active site. The observation that zinc is a potent inhibitor of a number of viral proteases including those from HIV and poliovirus has led to its classification as a general 'antiviral' agent and to the suggestion that zinc supplementation might be of some benefit to AIDS patients (Zhang *et al.*,1991). Recently phenylnorstatine derivatives of substrates have been designed that inhibit the HIV-1 protease with K_i's of 3-24μM (Raju and Deshpande,1991). The inhibition of the adenovirus protease by a number of transition elements and the elucidation of its substrate specificity means that a similar approach could be taken to design a specific inhibitor for the enzyme. The observation by Houde and Weber (1990b) that extracts from HeLa cells infected with adenovirus serotypes 1, 3, 4, 5, 6, 7 and 9 are able to efficiently process the Ad2 precursor protein pVII suggests that the substrate specificities of the proteases from other adenovirus serotypes are very similar, if not identical, to that described for Ad2. It is likely, therefore, that a specific inhibitor of the Ad2 protease would also inhibit the enzymes from other serotypes.

The mechanism of the adenovirus protease remains to be elucidated and work described in this project and by other workers shows that the sequence of the L3 23-kDa protein bears no resemblance to any other viral or cellular proteases. Its inhibitor profile suggests that it may be a member of the recently defined subclass of cysteine proteases that are related to trypsin; but elucidation of the crystal structure of the active enzyme would be required to confirm this.

Control of Protease Activity

The results in chapters 2-4 show conclusively that the activity of the Ad2 protease is

controlled in some way. This may be in order to prevent the premature cleavage of viral proteins by the enzyme during the infectious cycle. The HIV protease, the activity of which is not required until after assembly of the virus particles, is activated by dimerisation as discussed previously. The results in chapters 2 and 3 show that the adenovirus protease is active as a monomer and that it is not proteolytically activated. A requirement for high salt concentrations to stabilise the purified protein suggests that another cofactor is required to stabilise the enzyme in an active form. Other virus encoded proteases have been shown to require cofactors, particularly for *trans* cleavages. For example the poliovirus 2A protease and eIF3 act together to cleave the p220 subunit of eIF4 (Wyckoff *et al.*,1990) and the flavivirus NS3 protease requires another viral protein NS2B for the complete processing of its polyprotein (Chambers *et al.*,1991). In the case of the Ad2 protease the cofactor could be of viral or cellular origin, with the results from the coinfections of insect cells with pTP and 23-kDa recombinant baculoviruses suggesting that it is probably of cellular origin (chapter 4). A number of mechanisms could operate to control the enzyme. For example, a specific cellular or viral factor may be present that inhibits the protease early in infection or alternatively a covalent modification such as phosphorylation may be involved. As discussed in chapter 4 there is a suggestion that the 23-kDa binds specifically to DNA; but DNA was not found to activate the protease in any of the experiments discussed in this project. None of these control mechanisms on its own would account for all the observations made in this project.

In order to explain the results the following sequence of events is proposed. The Ad2 23-kDa protease is synthesised as an apoenzyme and is activated by another protein, perhaps a thioredoxin-like enzyme. Upon activation the protease binds to DNA which stabilises it in an active form. If the protein is removed from the DNA then it reverts to an inactive form unless purified in the presence of high salt which locks it in an active conformation. The involvement of a thioredoxin-like enzyme as a control element is an interesting possibility in view of its role in the binding of the T7 polymerase to DNA and the assembly of filamentous DNA bacteriophages (Gething and Sambrook,1992; Russel,1991). It might also help to explain the observations that DTT inhibits the recombinant 23-kDa protein; but stimulates the active enzyme in the purified virions. The binding of the active enzyme to DNA is also worth

considering since two other viral proteins, the poliovirus 3C protease and the Sindbis capsid protease have been shown to bind to RNA (Andino *et al.*,1990; Choi *et al.*,1991).

In the future it must be a priority to establish how the protease activity is controlled. Also of interest would be to determine the location of the 23-kDa protein within both Ad2 infected HeLa cells and recombinant baculovirus infected sf9 cells. The affinity purification of the antisera prepared in this project may be of use in immunofluoresence studies on Ad2 infected cells and cells infected with the Ad2ts1 mutant at permissive and non-permissive temperatures. At present it is thought that in Ad2ts1 the 23-kDa protease is not packaged within the virion (Anderson,1990). Presumably the proline to leucine mutation at position 139 in Ad2ts1 affects the structure of the enzyme at non-permissive temperatures. The location of this residue on the C-terminal side of the potential active site residues in a relatively variable region of the protein (see **fig.2.13**) means that it may affect the activation step.

In conclusion the adenovirus protease shares a number of features with other viral proteases including its unusual substrate requirements and inhibition profile; but this project suggests that it may also have some distinct attributes that merit consideration in the future.

REFERENCES

AKUSJARVI, G. & PERSSON, H. (1981). Gene and mRNA for precursor polypeptide VI from adenovirus type 2. *Journal of Virology* **38**, 469-482.

AKUSJARVI, G.; ALESTROM, P.; PETTERSSON, M.; LAGER, M; JORNVALL, H. & PETTERSSON, U. (1984). The gene for the adenovirus 2 hexon polypeptide. *Journal of Biological Chemistry* **259**, 13976-13979.

ALESTROM, P.; STENLUND, A.; LI, P. & PETTERSSON, U. (1982). A common sequence element in the inverted terminal repetitions of human and avian adenoviruses. *Gene* **18**, 193-197.

ALVEY, J.C.; WYCKOFF, E.E.; YU, S.F.; LLOYD, R. & EHRENFELD, E. (1991). *Cis-* and *trans-*cleavage activities of poliovirus 2A protease expressed in Escherichia coli. *Journal of Virology* **65**, 6077-6083.

ANDERSON, C.W. (1990). The proteinase polypeptide of adenovirus serotype 2 virions. *Virology* **177**, 259-272.

ANDERSON, C.W.; BAUM, P.R. & GESTELAND, R.F. (1973). Processing of adenovirus 2 induced proteins. *Journal of Virology* **12**, 241-252.

ANDERSON, C.W.; HARDY, M.M.; DUNN, J.J. & KLESSIG, D.F. (1983). Independent spontaneous mutants of adenovirus type 2-simian virus 40 hybrid AD2+ND3 that grow efficiently in monkey cells possess identical mutations in the adenovirus type 2 DNA-binding protein gene. *Journal of Virology* **48**, 31-39.

ANDERSON, C.W.; YOUNG, M.E.; FLINT, S.J. (1989). Characterisation of the adenovirus 2 virion protein, mu. *Virology* **172**, 506-512.

ANDERSON, K.P. & KLESSIG, D.F. (1984). Altered mRNA splicing in monkey cells abortively infected with human adenovirus may be responsible for inefficient synthesis of the virion fibre polypeptide. *Proceedings of the National Academy of Sciences (USA)* **81**, 4023-4027.

ANDINO, R.; RIECKOF, E. & BALTIMORE, D. (1990). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* **63**, 369-380.

ARGOS, P.; KAMER, P.; NICKLIN, M.J.H. & WIMMER, E. (1984). Similarity in gene organisation and homology between proteins of animal picornaviruses and plant comoviruses suggest common ancestry of these virus families. *Nucleic Acids Research* **12**, 7251-7267.

ARNOLD, E.; LUO, M.; VRIEND, G.; ROSSMANN, M.G.; PALMENBERG, A.C.; PARKS,G.D.; NICKLIN, M.J.H. & WIMMER, E. (1987). Implications of the picornavirus capsid structure for polyprotein processing. *Proceedings of the National Academy of Sciences (USA)* **84**, 21-25.

ATHERTON, E.; CAMERON, L.R. & SHEPPARD, R.C. (1988). Peptide synthesis. Part10. Use of pentafluorophenyl esters of fluorenylmethoxycarbonylamino acids in solid phase synthesis. *Tetrahedron* **44**, 843-857.

ATHERTON, E. & SHEPPARD, R.C. (1989). *Solid phase peptide synthesis, a practical approach.* Ed. by: D. Rickwood and B.D. Hames. IRL Press, Oxford University Press.

AULD, A.S. & VALLEE, B.L. (1987). Carboxypeptidase A. In: *Hydrolytic Enzymes* pp201-255. Ed. by: A. Neuberger and K. Brocklehurst. Amsterdam: Elsevier.

BABE, L.M.; PICHUANTES, S. & CRAIK, C.S. (1991). Inhibition of HIV protease activity by heterodimer formation. *Biochemistry* **30**, 106-111.

BARRETT, A.J. (1977). Introduction to the history and classification of tissue proteinases. In: *Proteinases*

in Mammalian Cells and Tissues pp1-55. Ed. by: A.J. Barrett. Amsterdam: North-Holland Publishing Company.

BARRETT, P.; CLARK, L. & HAY, R.T. (1987). A cellular protein binds to a conserved sequence in the adenovirus type 2 enhancer. *Nucleic Acids Research* **15**, 2719-2735.

BAUM, E.Z.; BEBERNITZ, G.A.; PALANT, O.; MUELLER, T. & PLOTCH, S.J. (1991). Purification, properties and mutagenesis of poliovirus 3C protease. *Virology* **185**, 140-150.

BAZAN, J.F. & FLETTERICK, R.J. (1988). Viral cysteine proteases are homologous to a trypsin-like family of serine proteases: structural and functional implications. *Proceedings of the National Academy of Sciences (USA)* **85**, 7872-7876.

BAZAN, J.F. & FLETTERICK, R.J. (1989). Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology* **171**, 637-639.

BERGER, A. & SCHECHTER, I. (1970). Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philosophical Transactions of the Royal Society of London (Series B)* **257**, 249-264.

BERK, A.J. (1986). Adenovirus promoters and E1a transactivation. *Annual Reviews of Genetics* **20**, 45-79.

BHATTI, A.R. & WEBER, J. (1978). Protease of adenovirus type 2. *In vitro* processing of core protein. *Biochemical and Biophysical Research Communications* **81**, 973-979.

BHATTI, A.R. & WEBER, J. (1979a). Protease of adenovirus type 2: partial characterisation. *Virology* **96**, 478-485.

BHATTI, A.R. & WEBER, J. (1979b). Protease of adenovirus type 2: subcellular localisation. *Journal of Biological Chemistry* **254**, 12265-12268.

BILLICH, A. & WINKLER, G. (1991). Analysis of subsite preferences of HIV-1 proteinase using MA/CA junction peptides substituted at the P3-P1' positions. *Archives of Biochemistry and Biophysics* **290**, 189-190.

BLAIR, G.E. & RUSSELL, W.C. (1978). Identification of a protein kinase activity associated with human adenoviruses. *Virology* **86**, 157-166.

BLAIR, W.S. & SEMLER, B.L. (1991). Role for the P4 amino acid residue in substrate utilisation by the poliovirus 3CD proteinase. *Journal of Virology* **65**, 6111-6123.

BOOY, F.P.; NEWCOMB, W.W.; TRUS, B.L.; BROWN, J.C.; BAKER, T.S. & STEVEN, A.C. (1991). Liquid crystalline, phage-like packaging of encapsidated DNA in herpes simplex virus. *Cell* **64**, 1007-1015.

BOS, J.C.; POLDER, L.J.; BERNARDS, R.; SCHRIER, P.I.; VAN DEN ELSEN, P.J., VAN DER EB, A.J. & VAN ORMONDT, H. (1981). The 2.2kb E1b mRNA of human Ad12 and Ad5 codes for 2 tumour antigens starting at different AUG triplets. *Cell* **27**, 121-131.

BOSHER, J.; LEITH, I.R.; TEMPERLEY, S.M.; WELLS, M. & HAY, R.T. (1991). The DNA-binding domain of nuclear factor I is sufficient to cooperate with the adenovirus type 2 DNA-binding protein in viral DNA replication. *Journal of General Virology* **72**, 2975-2980.

BOSHER, J.; ROBINSON, E.C. & HAY, R.T. (1990). Interactions between the adenovirus type 2 DNA polymerase and DNA binding domain of nuclear factor I. *New Biologist* **2**, 1083-1090.

BOUDIN, M.L.; MONCANY, M.; D'HALLUIN, J.C. & BOULANGER, P. (1979). Isolation and characterisation of adenovirus type 2 vertex capsomer (penton base). *Virology* **92**, 125-132.

BOUDIN, M.L.; D'HALLUIN, J.C.; COUSIN, C. & BOULANGER, P. (1980). Human adenovirus type 2

protein IIIa. *Virology* **101**, 144-156.

BOULANGER, P.A. & BLAIR, G.E. (1991). Expression and interactions of human adenovirus oncoproteins. *Biochemical Journal* **275**, 281-299.

BOULANGER, P.; LEMAY, P.; BLAIR, G.E. & RUSSELL, W.C. (1979). Characterisation of adenovirus protein IX. *Journal of General Virology* **44**, 783-800.

BOUTELJIE, J.; KARLSTROM, A.R.; HARTMANIS, M.G.N.; HOLMGREN, E.; SJORGREN, A. & LEVINE, R.L. (1990). Human immunodeficiency viral protease is catalytically active as a fusion protein: characterisation of the fusion and native enzymes produced in Escherichia coli. *Archives of Biochemistry and Biophysics* **283**, 141-149.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254.

BRAUN, D.K.; ROIZMAN, B. & PEREIRA, L. (1984). Characterisation of post-translational products of herpes simplex virus gene 35 products binding to the surface of full capsids but not empty capsids. *Journal of Virology* **49**, 142-153.

BRENNER, S. (1988). The molecular evolution of genes and proteins: a tale of two serines. *Nature* **334**, 528-530.

BRIDGE, E. & KETNER, G. (1989). Redundant control of adenovirus late gene expression by early region 4. *Journal of Virology* **63**, 631-638.

BROCKLEHURST, K.; WILLENBROCK, F. & SALITH, E. (1987). Cysteine proteinases. In: *Hydrolytic Enzymes* pp39-158. Ed. by: A. Neuberger and K. Brocklehurst. Amsterdam: Elsevier.

BROUGH, D.E.; RICE, S.A.; SELLS, S. & KLESSIG, D.F. (1985). Restricted changes in adenovirus DNA-binding protein that lead to extended host range or temperature sensitive phenotypes. *Journal of Virology* **55**, 206-212.

BROWN, D.T.; WESTPHAL, M.; BURLINGHAM, B.T.; WINTERHOF, U. & DOERFLER, W. (1975). Structure and composition of the adenovirus type 2 core. *Journal of Virology* **16**, 366-375.

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

BURSTEIN, H.; BIZUB, D. & SKALKA, M. (1991). Assembly and processing of avian retroviral *gag* polyproteins containing linked protease dimers. *Journal of Virology* **65**, 6165-6172.

BUTTERWORTH, B.E. & KORANT, B.D. (1974). Characterisation of the large picornaviral peptides produced in the presence of zinc ion. *Journal of Virology* **14**, 282-291.

CARRINGTON, J.C.; FREED, D.D. & OH, C.-S. (1990). Expression of potyviral polyproteins in transgenic plants reveals three proteolytic activities required for complete processing. *EMBO Journal* **9**, 1347-1353.

CEPKO, C.L. & SHARP, P.A. (1983). Analysis of Ad5 hexon and 100K ts mutants using conformation-specific monoclonal antibodies. *Virology* **129**, 137-154.

CHALLBERG, M.D.; DESIDERIO, S.V. & KELLY, T.J. (1980). Adenovirus DNA replication *in vitro*: characterisation of a protein covalently linked to nascent DNA strands. *Proceedings of the National Academy of Sciences (USA)* **77**, 5105-5109.

CHALLBERG, M.D. & KELLY, T.J. (1979). Adenovirus DNA replication *in vitro*. *Proceedings of the National Academy of Sciences (USA)* **76**, 655-659.

CHALLBERG, M.D. & KELLY, T.J. (1989). Animal virus DNA replication. *Annual Reviews of Biochemistry* **58**, 671-717.

CHAMBERS, T.J.; GRAKOUI, A. & RICE, C.M. (1991). Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 proteinase domain and NS2B are required for cleavages at dibasic sites. *Journal of Virology* **65**, 6042-6050.

CHAMBERS, T.J.; WEIR, R.C.; GRAKOUI, A.; McCOURT, D.W.; BAZAN, J.F.; FLETTERICK, R.J. & RICE, C.M. (1990). Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site specific cleavages in the viral polyprotein. *Proceedings of the National Academy of Sciences (USA)* **87**, 8898-8902.

CHANOCK, R.M. & McINTOSH, K. (1990). Parainfluenza viruses. In: *Fields Virology Volume 1. Second Edition* pp963-988. Ed. by: B.N. Fields and D.M. Knipe. Ravens Press Ltd., New York.

CHATTERJEE, P.K. & FLINT, S.J. (1987). Adenovirus type 2 endopeptidase: an unusual phosphoprotein enzyme matured by autocatalysis. *Proceedings of the National Academy of Sciences (USA)* **84**, 714-718.

CHATTERJEE, P.K.; YANG, U.C. & FLINT, S.J. (1986). Comparisons of the interactions of the adenovirus type 2 major core protein and its precursor with DNA. *Nucleic Acids Research* **14**, 2721-2735.

CHEAH, K-C.; LEONG, L.E-C. & PORTER, A.G. (1990). Site-directed mutagenesis suggests close functional relationship between a human rhinovirus 3C cysteine protease and cellular trypsin-like serine proteases. *The Journal of Biological Chemistry* **265**, 7180-7187.

CHEE-SHEUNG, C.C. & GINSBERG, H.S. (1982). Characterisation of a temperature sensitive fiber mutant of type 5 adenovirus and effect of the mutation on virion assembly. *Journal of Virology* **42**, 932-950.

CHEN, M.; MERMOD, N. & HORWITZ, M.S. (1990). Protein-protein interactions between adenovirus DNA polymerase and nuclear factor I mediate formation of the DNA replication preinitiation complex. *Journal of Biological Chemistry* **265**, 18634-18642.

CHOI, H-K.; TONG, L.; MINOR, W.; DUMAS, P.; BOEGE, U.; ROSSMANN, M.G. & WENGLER, G. (1991). Structure of sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organisation of the virion. *Nature* **354**, 37-43.

CLARK, M.E.; HAMMERLE, T.; WIMMER, E. & DASGUPTA, A. (1991). Poliovirus proteinase 3C converts an active form of transcription factor IIIc to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO Journal* **10**, 2941-2947.

CLEAT, P.H. & HAY, R.T. (1989). Cooperative interactions between NFI and the adenovirus DNA binding protein at the adenovirus origin of DNA replication. *EMBO Journal* **8**, 1841-1848.

CLEGHON, V.G. & KLESSIG, D.F. (1986). Association of the adenovirus DNA-binding protein with RNA both *in vitro* and *in vivo*. *Proceedings of the National Academy of Sciences (USA)* **83**, 8947-8951

COLBY, W.W. & 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.

CORDEN, J.; ENGELKING, H.M. & PEARSON, G.D. (1976). Chromatin-like organisation of the adenovirus chromosome. *Proceedings of the National Academy of Sciences (USA)* **73**, 401-404.

CORDINGLEY, M.G.; CALLAHAN, P.L.; SARDANA, V.V.; GARSKY, V.M. & COLONNO, R.J. (1990). Substrate requirements of human rhinovirus 3C protease for peptide cleavage *in vitro*. *Journal of Biological Chemistry* **265**, 9062-9065.

CORDINGLEY, M.G.; REGISTER, R.B.; CALLAHAN, P.L.; GARSKY, V.M. & COLONNO, R.J. (1989). Cleavage of small peptides *in vitro* by human rhinovirus 14 3C protease expressed in Escherichia

coli. *Journal of Virology* **63**, 5037-5045.

COULSON, A.F.W.; COLLINS, J.F. & LYALL, A. (1987). Protein and nucleic acid sequence database searching: a suitable case for parallel processing. *Computer Journal* **30**, 420-424.

CRAVEN, R.C; BENNETT, R.P. & WILLS, J.W. (1991). Role of the avian retroviral protease in the activation of reverse transcriptase during virion assembly. *Journal of Virology* **65**, 6205-6217.

CROSSLAND, L.D. & RASKAS, H.J. (1983). Identification of adenovirus genes that require template replication for expression. *Journal of Virology* **46**, 737-748.

CUPO, J.F.; REZANKA, L.J. & HARPST, J.A. (1987). Column purification of adenovirus cores. *Analytical Biochemistry* **164**, 267-270.

DALES, S. & CHARDONNET, Y. (1973). Early events in the interaction of adenovirus with HeLa cells. IV. Association with microtubules and the nuclear pore complex during vectorial movement in the inoculum. *Virology* **56**, 465-483.

DARKE, P.L.; NUTT, R.F.; BRADY, S.F.; GARSKY, V.M.; CICCARONE, T.M.; LEU, C-T.; LUMMA, P.K.; FREIDINGER, R.M.; VERBER, D.F. & SIGAL, I.S. (1988). HIV-1 protease specificity of peptide cleavage is sufficient for processing of *gag* and *pol* polyproteins. *Biochemical and Biophysical Research Communications* **156**, 297-303.

DE GROOT, R.J.; HARDY, W.R.; SHIRAKO, Y. & STRAUSS, J.H. (1990). Cleavage-site preferences of sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *EMBO Journal* **9**, 2631-2638.

DEJARLAIS, R.L.; SEIBEL, G.L.; KUNTZ, I.D.; FURTH, P.S.; ALVAREZ, J.C.; ORITZ DE MONTELLANO, P.R.; DECAMP, D.L.; BABE, L.M. & CRAIK, C.S. (1990). Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proceedings of the National Academy of Sciences (USA)* **87**, 6644-6648.

DEVANEY, M.A.; VAKHARIA, V.N.; LLOYD, R.E.; EHRENFELD, E. & GRUBMAN, M.J. (1988). Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *Journal of Virology* **62**, 4407- 4409.

DEVAUX, C. & BOULANGER, P. (1980). Reactive serine in human adenovirus hexon polypeptide. *Virology* **102**, 94-106.

DEVAUX, C.; CAILLET-BOUDIN, M-L.; JACROT, B. & BOULANGER, P. (1987). Crystallisation, enzymatic cleavage and polarity of the adenovirus type 2 fiber. *Virology* **161**, 121-128.

DEVAUX, C.; ADRIAN, M., BERTHET-COLOMINAS, C.; CUSACK, S. & JACROT, B. (1990). Structure of the adenovirus fibre I. *Journal of Molecular Biology* **215**, 567-588.

DEVEREUX, J.; HAEBERLI, P. & SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387-396.

D'HALLUIN, J.C.; MILLEVILLE, M.; BOULANGER, P.A. & MARTIN, G.R. (1978). Temperature sensitive mutant of adenovirus type 2 blocked in virion assembly: accumulation of light intermediate particles. *Journal of Virology* **26**, 344-356.

DINGWALL, C. & LASKEY, R.A. (1991). Nuclear targeting sequences - a consensus? *Trends in Biochemical Sciences* **16**, 478-481.

DOERFLER, W. & KLEINSCHMIDT, A. (1970). Denaturation pattern of the DNA of adenovirus type 2 as determined by electron microscopy. *Journal of Molecular Biology* **50**, 579-593.

DOUGHERTY, W.G. & CARRINGTON, J.C. (1988). Expression and function of potyviral gene products.

Annual Reviews in Phytopathology **26**, 123-143.

DOUGHERTY, W.G. & PARKS, T.D. (1989). Molecular genetic and biochemical evidence for the involvement of the heptapeptide cleavage sequence in determining the reaction profile of two tobacco etch virus cleavage sites in cell free assays. *Virology* **172**, 145-155.

DOUGHERTY, W.G. & PARKS, T.D. (1991). Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology* **183**, 449-456.

DUNN, B.M. (1989). Determination of protease mechanism. In: *Proteolytic Enzymes, a practical approach* pp57-81. Ed. by: R.J. Beynon and J.S. Bond. IRL Press, Oxford University Press.

ESCHE, H.; MATHEWS, M.B. & LEWIS, J.B. (1980). Proteins and messengers of the transforming region of the wild type and mutant adenoviruses. *Journal of Molecular Biology* **142**, 399-407.

EVERITT, E.; LUTTER, L. & PHILIPSON, L. (1975). Structural proteins of adenovirus. XII. Location and neighbour relationship among proteins of adenovirion type 2 as revealed by enzymatic iodination, immunoprecipitation and chemical cross-linking. *Virology* **67**, 197-208.

EVERITT, E.; SUNDQUIST, B.; PETTERSSON, U. & PHILIPSON, L. (1973). Structural proteins of adenoviruses. X. Isolation and topography of low molecular weight antigens from the virion of adenovirus type 2. *Virology* **62**, 130-147.

EVERITT, E.; PERSSON, M.J. & WOHLFART, C. (1988). pH-Dependent exposure of endoproteolytic cleavage sites of the adenovirus 2 hexon protein. *FEMS Microbiological letters* **49**, 229-233.

FALGOUT, B. & KETNER, G. (1988). Characterisation of adenovirus particles made by deletion mutants lacking the fiber gene. *Journal of Virology* **62**, 622-625.

FIELD, J.; GRONOSTAJSKI, M. & HURWITZ, J. (1984). Properties of the adenovirus DNA polymerase. *Journal of Biological Chemistry* **259**, 9487-9495.

FITZGERALD, P.M.D. & SPRINGER, J.P. (1991). Structure and function of retroviral proteases. *Annual Reviews of Biophysics and Biophysical Chemistry* **20**, 299-320.

FLINT, S.J.; BERGET, S.M. & SHARP, P.A. (1976). Characterisation of single stranded viral DNA sequences present during replication of adenovirus types 2 and 5. *Cell* **9**, 559-569.

FOUNDLING, S.I.; SALEMME, F.R.; KORANT, B.; WENDOLOSKI, J.J.; WEBER, P.C.; TREHARNE, A.C.; SCHADT, M.C.; JASKOLSKI, M.; MILLER, M.; WLODAWER, A.; STROP, P.; KOSTKA, V.; SEDLACEK, J. & OHLENDORF, D.H. (1989). Crystal structure of a retroviral proteinase from avian myeloblastosis associated virus. In: *Viral proteinases as targets for chemotherapy* pp181-190. Ed. by: H-G. Krausslich, S. Oroszlan and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

FRANKLIN, R.M.; PETTERSSON, U.; AKERVALL, K.; STRANDBERG, B. & PHILIPSON, L. (1971). Structural proteins of adenoviruses. V. On the size and structure of the adenovirus type 2 hexon. *Journal of Molecular Biology* **57**, 383-391.

FRASER, N.W.; NEVINS, J.R.; ZIFF, E.B. & DARNELL, J.E. (1979). The major late adenovirus type 2 transcription unit: termination is downstream from the last poly(A) site. *Journal of Molecular Biology* **129**, 643-656.

FRESHNEY, R.I. (1987). *Culture of animal cells. A manual of basic techniques. Second Edition.* Alan R. Liss, Inc., New York.

FURCINITTI, P.S.; VAN OOSTRUM, J. & BURNETT, R.M. (1989). Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography. *EMBO Journal* **8**, 3563-3570.

GARON, C.F.; BERRY, K.W.; HIERHOLZER, J.C. & ROSE, J.A. (1973). Mapping of base sequence heterologies between genomes from different adenovirus serotypes. *Virology* **54**, 414-426.

GETHING, M-J. & SAMBROOK, J. (1992). Protein folding in the cell. *Nature* **335**, 33-45.

GHOSH-CHOUDHURY, G.; HAJ-AHMAD, Y. & GRAHAM, F.L. (1987). Protein IX, a minor component of the human adenovirus capsid is essential for the packaging of full length genomes. *EMBO Journal* **6**, 1733-1739.

GINGERAS, T.R.; SCIAKY, D.; GELINAS, R.E.; BING-DONG, J.; YEN, C.E.; KELLY, M.M.; BULLOCK, P.A.; PARSONS, B.L.; O'NEILL, K.E. & ROBERTS, R.J. (1982). Nucleotide sequences from the adenovirus 2 genome. *Journal of Biological Chemistry* **257**, 13475-13480.

GINSBERG, H.S.; PEREIRA, H.G.; VALENTINE, R.C. & WILCOX, W.C. (1966). A proposed terminology for the adenovirus antigens and virion morphological subunits. *Virology* **28**, 782-783.

GINSBERG, H.S. (1979). Adenovirus structural proteins. In: *Comprehensive Virology, Volume 13* pp409-457. Ed. by: H. Fraenkel-Conrat and R.R. Wagner. Plenum Press, New York.

GOEBEL, S.J.; JOHNSON, G.P.; PERKUS, M.E.; DAVID, S.W.; WINSLOW, J.P. & PAOLETTI, E. (1990). The complete DNA sequence of vaccinia virus. *Virology* **179**, 247-266.

GOODING, L.R.; ELMORE, L.E.; TOLLEFSON, A.E.; BRODY, H.A. & WOLD, W.S.M. (1988). A 14700 MW protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell* **53**, 341-346.

GORBALENYA, A.E.; BLINOV, V.M. & DONCHENKO, A.P. (1986). Poliovirus encoded proteinase 3C: a possible evolutionary link between cellular serine and cysteine proteinase families. *FEBS Letters* **194**, 253-257.

GORBALENYA, A.E.; DONCHENKO, A.P.; KOONIN, E.V. & BLINOV, V.M. (1989). N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic Acids Research* **17**, 3889-3897.

GORBALENYA, A.E. & KOONIN, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Research* **17**, 8413-8440.

GORBALENYA, A.E.; KOONIN, E.V. & LAI, M.M. (1991). Identification of rubi- and aphthovirus proteases and the delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Letters* **288**, 201-205.

GRABLE, M. & HEARING, P. (1990). Adenovirus type 5 packaging domain is composed of a repeated element that is functionally redundant. *Journal of Virology* **64**, 2047-2056.

GRAHAM, F.L. (1984). Transformation by and oncogenicity of human adenoviruses. In: *The Adenoviruses* pp339-398. Ed. by: H.S. Ginsberg. Plenum Press, New York.

GRAVES, M.C.; LIM, J.J.; HEIMER, E.P. & KRAMER, R.A. (1988). An 11-kDa from of human immunodeficiency virus protease expressed in Escherichia coli is sufficient for enzymatic activity. *Proceedings of the National Academy of Sciences (USA)* **85**, 2449-2453.

GREEN, M.; LOWENSTEIN, P.M.; PUSZTAIR, R. & SYMINGTON, J.S. (1988). An adenovirus E1a protein domain activates transcription *in vivo* and *in vitro* in the absence of protein synthesis. *Cell* **53**, 921-926.

GREEN, M.; PARSONS, J.T.; PINA, M.; FUJINAGA, K.; CAFFIER, H. & LANDGRAF-LEURS, I. (1970). Transcription of adenovirus genes in productively infected and in transformed cells. *Cold Spring Harbor Symposium Quantitative Biology* **35**, 803-818.

GREEN, M.; PINA, M.; KIMES, R.; WENSINK, P.C.; MACHATTIE, L.A. & THOMAS, C.A. (1967). Adenovirus DNA. I. Molecular weight and conformation. *Proceedings of the National Academy of Sciences (USA)* **57**, 1302-1309.

GREEN, N.M.; WRIGLEY, N.G.; RUSSELL, W.C.; MARTIN, S.R. & MCLACHLAN, A.D. (1983). Evidence for a repeating cross-sheet structure in the adenovirus fibre. *EMBO Journal* **2**, 1357-1365.

GRUTTER, M. & FRANKLIN, R.M. (1974). Studies on the molecular weight of the adenovirus type 2 hexon and its subunit. *Journal of Molecular Biology* **89**, 163-171.

GUSTCHINA, A. & WEBER, I.T. (1990). Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: role of the flap. *FEBS Letters* **269**, 268-272.

GUSTCHINA, A. & WEBER, I.T. (1991). Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. *Proteins: Structure, Function and Genetics* **10**, 325-339.

GUYER, B.; O'DAY, D.M.; HIERHOLZER, J.C. & SCHAFFNER, W. (1975). Epidemic keratoconjunctivitis: a community outbreak of mixed adenovirus type 8 and type 19 infection. *Journal of Infectious Diseases* **132**, 142-150.

HAHN, C.S.; STRAUSS, E.G. & STRAUSS, J.H. (1985). Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid autoprotease. *Proceedings of the National Academy of Sciences (USA)* **82**, 4648-4652.

HAHN, C.S. & STRAUSS, J.H. (1990). Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. *Journal of Virology* **64**, 3069-3073.

HAMMARSKJOLD, M-L. & WINBERG, G. (1980). Encapsidation of adenovirus 16 DNA is directed by a small DNA sequence at the left end of its genome. *Cell* **20**, 787-795.

HAMMOND, G.; HANNAN, C.; YEH, T.; FISCHER, K.; MAUTHE, G. & 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.

HANDA, H.; KINGSTON, R.E. & SHARP, P.A. (1983). Inhibition of adenovirus early region IV transcription *in vitro* by a purified viral DNA binding protein. *Nature* **302**, 545-547.

HANNAN, C.; RAPTIS, L.H.; DERY, C.V. & WEBER, J. (1983). Biological and structural studies on adenovirus type 2 temperature-sensitive mutant defective for uncoating. *Intervirology* **19**, 213-223.

HARBER, J.J.; BRADLEY, J.; ANDERSON, C.W. & WIMMER, E. (1991). Catalysis of poliovirus VP0 maturation cleavage is not mediated by serine 10 of VP2. *Journal of Virology* **65**, 326-334.

HARRIS, M.P.G. & HAY, R.T. (1988). DNA sequences required for initiation of adenovirus type 4 DNA replication. *Journal of Molecular Biology* **201**, 57-67.

HARRISON, S.C. (1991). Viruses. *Current Opinion in Structural Biology* **1**, 288-295.

HASSON, T.B.; SOLOWAY, P.D.; ORNELLES, D.A.; DOERFLER, W. & SHENK, T. (1989). Adenovirus L1 52- and 55- kilodalton proteins are required for assembly of virions. *Journal of Virology* **63**, 3612-3621.

HAY, R.T. & RUSSELL, W.C. (1989). Recognition mechanisms in the synthesis of animal virus DNA. *Biochemical Journal* **258**, 3-16.

HAY, R.T.; STOW, N.D. & MCDOUGALL, I.M. (1984). Replication of adenovirus minichromosomes. *Journal of Molecular Biology* **175**, 493-510.

HAYES, B.W.; TELLING, G.C.; MYAT, M.; WILLIAMS, J.F. & 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.

HEINTZ, N. (1991). Transcriptional Regulation. *Trends in Biological Sciences* **16**, 393.

HELLEN, C.U.T.; FACKE, M.; KRAUSSLICH, H-G.; LEE, C-K. & WIMMER, E. (1991). Characterisation of poliovirus 2A proteinase by mutational analysis: residues required for autocatalytic activity are essential for induction of cleavage of eukaryotic initiation factor 4F polypeptide p220. *Journal of Virology* **65**, 4226-4231.

HELLEN, C.U.T.; KRAUSSLICH, H. & WIMMER, E. (1989). Proteolytic processing of polyproteins in the replication of RNA viruses. *Biochemistry* **28**, 9981-9890.

HELLEN, C.U.T.; LEE, C-K.; KRAUSSLICH, H-G.; FACHE, M. & WIMMER, E. (1990). Proteolytic processing of picornavirus polyproteins: mechanism, specificity of cleavage by polio 2A proteinase. Abstract W56-1 in VIIIth International Congress of Virology, Berlin.

HERRMANN, C.H.; SU, L-K. & HARLOW, E. (1991). Adenovirus E1A is associated with a serine/threonine protein kinase. *Journal of Virology* **65**, 5848-5859.

HIERHOLZER, J.C.; STONE, Y.O. & BRODERSON, J.R. (1991). Antigenic relationships among the 47 human adenoviruses determined in reference horse antisera. *Archives in Virology* **121**, 179-197.

HOEFFLER, W.K.; KOVELMAN, R. & ROEDER, R.G. (1988). Activation of transcription factor IIIC by the adenovirus E1a protein. *Cell* **53**, 907-920.

HONG, J.S. & ENGLER, J.A. (1991). The amino terminus of the adenovirus fiber protein encodes the nuclear localisation signal. *Virology* **185**, 758-767.

HORWITZ, M.S. (1990a). Adenoviridae and their replication. In: *Fields Virology Volume 2. Second Edition* pp1679-1722. Ed. by: B.N. Fields and D.M. Knipe. Raven Press Ltd., New York.

HORWITZ, M.S. (1990b). Adenoviruses. In: *Fields Virology Volume 2. Second Edition* pp1723-1742. Ed. by: B.N. Fields and D.M. Knipe. Raven Press Ltd., New York.

HOUDE, A. & WEBER, J.M. (1987). Sequence of the protease of human subgroup E adenovirus type 4. *Gene* **54**, 51-56.

HOUDE, A. & WEBER, J.M. (1990a). Adenovirus proteinases: comparison of amino acid sequences and expression of the cloned cDNA in Escherichia coli. *Gene* **88**, 269-273.

HOUDE, A. & WEBER, J.M. (1990b). Adenovirus type 2 precursor proteins are cleaved by proteinases of other adenoviruses. *Virology* **179**, 485-486.

HUANG, J.T. & SCHNEIDER, R.J. (1991). Adenovirus inhibition of cellular protein synthesis involves inactivation of cap binding protein. *Cell* **65**, 271-280.

HUANG, M-M. & HEARING, P. (1989). Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *Journal of Virology* **63**, 2605-2615.

HUEBNER, R.J.; ROWE, W.P. & CHANNOCK, R.M. (1958). New recognised respiratory tract viruses. *Annual Reviews of Microbiology* **12**, 49-76.

HYLAND, L.J.; TOMASZEK, T.A. & MEEK, T.D. (1991). Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **30**, 8454-8463.

ISHIBASHI, M. & YASUE, H. (1984). Adenoviruses of animals. In: *The Adenoviruses* pp497-562. Ed.

by: H.S. Ginsberg. Plenum Press, New York.

JACOBSON, M.F. & BALTIMORE, D. (1968). Polypeptide cleavages in the formation of poliovirus proteins. *Proceedings of the National Academy of Sciences (USA)* **61**, 77-84.

JOHNSON, J.E. (1991). Viruses getting their heads together. *Current Biology* **1**, 287-289.

KAISER, E.; COLESCOTT, R.L.; BOSSINGER, C.D. & COOK, P.I. (1970). Color test for detection of free terminal amino groups in solid phase synthesis of peptides. *Analytical Biochemistry* **34**, 595-598.

KARLSTROM, A.R. & LEVINE, R.L. (1991). Copper inhibits the protease from human immunodeficiency virus 1 by both cysteine-dependent and cysteine-independent mechanisms. *Proceedings of the National Academy of Sciences (USA)* **88**, 5552-5556.

KATOH, I.; YASUNAGA, T.; IKAWA, Y. & YOSHINAKA, Y. (1987). Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. *Nature* **329**, 654-656.

KAUFFMAN, R.S. & GINSBERG, H.S. (1976). Characterisation of a temperature-sensitive hexon transport mutant of type 5 adenovirus. *Journal of Virology* **19**, 643-658.

KAWAHARA, N.; YANG, X.Z; SAKAGUCH, T.; KIYOTANI, K.; NAGAI, Y. & YOSHIDA, T. (1992). Distribution and substrate specificity of intracellular proteolytic processing enzyme(s) for paramyxovirus fusion glycoproteins. *Journal of General Virology* **73**, 583-590.

KAY, J. & DUNN, B.M. (1990). Viral proteinases: weakness in strength. *Biochimica et biophysica acta* **1048**, 1-18.

KEAN, K.M.; TETERINA, N.L. & GIRARD, M. (1990). Cleavage specificity of the poliovirus 3C protease is not restricted to Gln-Gly at the 3C/3D junction. *Journal of General Virology* **71**, 2553-2563.

KEAN, K.M.; TETERINA, N.L.; MARC, D. & GIRARD, M. (1991). Analysis of putative active site residues of the poliovirus 3C protease. *Virology* **181**, 609-619.

KEDINGER, C.; BRISON, O.; PERRIN, F. & WILHELM, J. (1978). Structural analysis of replicative intermediates from adenovirus type 2 infected cell nuclei. *Journal of Virology* **26**, 364-379.

KELLY, T.J. (1984). Adenovirus DNA replication. In: *The Adenoviruses* pp271-307. Ed. by: H.S. Ginsberg. Plenum Press, New York.

KEMP, M.C.; HIERHOLZER, J.C.; CABRADILLA, C.P. & OBIJESKI, J.F. (1983). The changing etiology of epidemic keratoconjunctivitis: antigenic and restriction enzyme analysis of adenovirus types 19 and 37 isolated over a 10 year period. *Journal of Infectious Diseases* **148**, 29-33.

KENNY, M.K.; BALOGH, L.A. & HURWITZ, J. (1988). Initiation of adenovirus DNA replication: I. Mechanism of action of a host protein required for replication of adenovirus DNA templates devoid of terminal protein. *Journal of Biological Chemistry* **263**, 9809-9817.

KENNY, M.K. & HURWITZ, J. (1988). Initiation of adenovirus DNA replication: II. Structural requirements using synthetic oligonucleotide templates. *Journal of Biological Chemistry* **263**, 9801-9808.

KIMELMAN, D.; MILLER, J.S.; PORTER, D. & ROBERTS, B.E. (1985). E1a regions of human adenoviruses and of the highly oncogenic simian adenovirus 7 are closely related. *Journal of Virology* **53**, 399-409.

KINGSBURY, D.W. (1990). Orthomyxoviridae and their replication. In: *Fields Virology Volume 1. Second Edition* pp1075-1090. Ed. by: B.N. Fields and D.M. Knipe. Raven Press. Ltd., New York.

KITAMURA, N.; SEMLER, B.L.; ROTHBERG, P.G.; LARSEN, G.R.; ADLER, C.J.; DORNER, A.J.; EMINI, E.A.; HANECAK, R.; LEE, J.J.; VAN DER WERF, S.; ANDERSON, C.W. & WIMMER, E.

(1981). Primary structure, gene organisation and polypeptide expression of poliovirus RNA. *Nature* **291**, 547-553.

KITTS, P.A.; AYRES, M.D. & POSSEE, R.D. (1990). Linearisation of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Research* **18**, 5667-5672.

KLENK, H.D.; ROTT, R.; ORLICH, M. & BLODORN, J. (1975). Activation of influenza A viruses by trypsin treatment. *Virology* **68**, 426-439.

KONVALINKA, J.; STROP, P.; VELEK, J.; CERNA, V.; KOSTKA, V.; PHYLIP, L.H.; RICHARDS, A.D.; DUNN, B.M. & KAY, J. (1990). Sub-site preference of the aspartic proteinase from the human immunodeficiency virus, HIV-1. *FEBS Letters* **268**, 35-38.

KOTLER, M.; DANHO, W.; KATZ, R.A.; LEIS, J. & SKALKA, A.M. (1989). Avian retroviral protease and cellular aspartic proteases are distinguished by activities on peptide substrates. *Journal of Biological Chemistry* **264**, 3428-3435.

KOTLER, M.; KATZ, R.A.; DANHO, W.; LEIS, J. & SKALKA, A.M. (1988). Synthetic peptides as substrates and inhibitors of a retroviral protease. *Proceedings of the National Academy of Sciences (USA)* **85**, 4185-4189.

KRAUSSLICH, H-G. (1991). Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proceedings of the National Academy of Sciences (USA)* **88**, 3213-3217.

KRAUSSLICH, H-G.; NICKLIN, M.G.H.; TOYODA, H.; ETCHISON, D. & WIMMER, E. (1987). Poliovirus proteinase 2A induces cleavage of eukaryotic initiation factor 4F polypeptide p220. *Journal of Virology* **61**, 2711-2718.

KRAUSSLICH, H-G. & WIMMER, E. (1988). Viral proteinases. *Annual Reviews of Biochemistry* **57**, 701-754.

KRING, S.C. & SPINDLER, K.R. (1990). Sequence of mouse adenovirus type 1 DNA encoding the amino terminus of protein IVa2. *Nucleic Acids Research* **18**, 4003.

LAWSON, M.A. & SEMLER, B.L. (1991). Poliovirus thiol proteinase 3C can utilise a serine nucleophile within the putative catalytic triad. *Proceedings of the National Academy of Sciences (USA)* **88**, 9919-9923.

LAZAROWITZ, S.G.; GOLDBERG, A.R. & CHOPPIN, P.W. (1973). Proteolytic cleavage by plasmin of the HA polypeptide of influenza virus: host cell activation of serum plasminogen. *Virology* **56**, 172-180.

LEE, C-K. & WIMMER, E. (1988). Proteolytic processing of poliovirus polyprotein: elimination of 2Apro-mediated alternative cleavage of polypeptide 3CD by *in vitro* mutagenesis. *Virology* **166**, 405-414.

LEIS, J.; BIZUB, D.; WEBER, I.T.; CAMERON, C.; KATZ, R.; WLODAWER, A. & SKALKA, A.M. (1989). Structure-function analysis of the retroviral aspartic proteinase. In: *Viral proteinases as Targets for Chemotherapy* pp235-243. Ed by: H-G. Krausslich, S. Oroszlan and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

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

LICHY, J.H.; FIELD, J.; HORWITZ, M.S. & HURWITZ, J. (1982). Separation of adenovirus terminal protein precursor from its associated DNA polymerase: role of both proteins in the initiation of adenovirus DNA replication. *Proceedings of the National Academy of Sciences (USA)* **79**, 5225-5229.

LINDENBAUM, J.O.; FIELD, J. & HURWITZ, J. (1986). The adenovirus DNA-binding protein and adenovirus polymerase interact to catalyse elongation of primed DNA templates. *Journal of Biological Chemistry* **261**, 10218-10227.

LILLIE, J.W. & GREEN, M.R. (1989). Transcription activation by the adenovirus E1a protein. *Nature* **338**, 39-44.

LINNE, T. & PHILIPSON, L. (1980). Further characterisation of the phosphate moiety of the adenovirus type 2 DNA binding protein. *European Journal of Biochemistry* **103**, 259-270.

LOEB, D.D.; SWANSTROM, R.; EVERITT, L.; MANCHESTER, M.; STAMPER, S.E. & HUTCHISON, C.A. (1989). Complete mutagenesis of the HIV-1 protease. *Nature* **340**, 397-400.

LOPEZ-OTIN, C.; SIMON-MATEO, C.; MARTINEZ, L. & VINUELA, E. (1989). Gly-Gly-X, a novel consensus sequence for the proteolytic processing of viral and cellular proteins. *Journal of Biological Chemistry* **264**, 9107-9110.

LOUIS, J.M.; MCDONALD, R.A.; NASHED, N.T.; WONDRAK, E.M.; JERINA, D.M.; OROSZLAN, S. & MORA, P.T. (1991). Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in Escherichia coli. *European Journal of Biochemistry* **199**, 361-369.

LUCAS, J.J. & GINSBERG, H.S. (1971). Synthesis of virus-specific RNA in KB cells infected with type 2 adenovirus. *Journal of Virology* **8**, 203-213.

LUNBLAD, R.L. & STEIN, W.H. (1969). On the reaction of diazoacetyl compounds with pepsin. *Journal of Biological Chemistry* **244**, 154-160.

LYONS, R.H.; FERGUSON, B.Q. & ROSENBERG, M. (1987). Pentapeptide nuclear localisation signal in adenovirus E1a. *Molecular Cell Biology* **7**, 2451-2456.

McFERRAN, J.B.; ROWLEY, H.M.; McNULTY, M.S. & MONTGOMERY, L.J. (1977). Serological studies on flocks showing depressed egg produciton. *Avian pathology* **6**, 405-413.

MAIZEL, J.V.; WHITE, D.O. & SCHARFF, M.D. (1968a). The polypeptides of adenovirus. I. Evidence for multiple protein components of the virion and a comparison of types 2, 7A and 12. *Virology* **36**, 115-125.

MAIZEL, J.V.; WHITE, D.O. & SCHARFF, M.D. (1968b). The polypeptides of adenovirus. II. Soluble proteins, cores, top components and structure of the virion. *Virology* **36**, 126-136.

MANIATIS, T.; FRITSCH, E.F. & SAMBROOK, J. (1982). *Molecular Cloning. A laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.

MARGOLIN, N.; HEATH, W.; OSBORNE, E.; LAI, M. & VLAHOS, C. (1990). Substitutions at the P2' site of *gag* p17-p24 affect cleavage efficiency by HIV-1 protease. *Biochemical and Biophysical Research Communications* **167**, 554-560.

MATHEWS, M.B. & SHENK, T. (1991). Adenovirus virus-associated RNA and translational control. *Journal of Virology* **65**, 5657-5662.

MERCER, A.A.; FRASER, K.M.; STOCKWELL, P.A. & ROBINSON, A.J. (1989). A homologue of retroviral pseudoproteases in the parapoxvirus, orf virus. *Virology* **172**, 665-668.

MILES, B.D.; LUFTIG, R.B.; WEATHERBEE, J.A.; WEIHING, R.R. & WEBER, J. (1980). Quantitation of the interaction between adenovirus types 2 and 5 and microtubules inside infected cells. *Virology* **105**, 265-269.

MILLER, L.K. (1988). Baculoviruses for foreign gene expression in insect cells. In: *Vectors a survey of molecular cloning vectors and their uses* pp457-465. Ed. by: R.L. Rodriguez and D.T. Denhardt. Butterworths (Boston).

MILLER, M.; JACKOLSKI, M.; RAO, M.; LEIS, J. & WLODAWER, A. (1989a). Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337**, 576-579.

MILLER, M.; SCHENIDER, J.; SATHYANARAYANA, B.K.; TOTH, M.V.; MARSHALL, G.R.; CLAWSON, L.; SELK, L.; KENT, S.B.H. & WLODAWER, A. (1989b). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3A° resolution. *Science* **246**, 1149-1152.

MIRZA, M.A. & WEBER, J. (1982). Structure of adenovirus chromatin. *Biochimica et biophysica acta* **696**, 76-86.

MORAN, E. & MATHEWS, M.B. (1987). Multiple functional domains in the adenovirus E1a gene. *Cell* **48**, 177--178.

MORIN, N. & BOULANGER, P. (1986). Hexon trimerisation occurring in an assembly defective, 100K temperature-sensitive mutant of adenovirus 2. *Virology* **152**, 11-31.

MUL, Y.M. & VAN DER VLIET, P.C. (1992). Nuclear factor I enhances adenovirus DNA replication by increasing the stability of a preinitiation complex. *EMBO Journal* **11**, 751-760.

MUL, Y.M. VERRIJZER, C.P. & VAN DER VLIET, P.C. (1990). Transcription factors NFI and NFIII/oct-1 function independently employing different mechanisms to enhance adenovirus DNA replication. *Journal of Virology* **64**, 5510-5518.

NASSAL, M.; GALLE, P.R. & SCHALLER, H. (1989). Protease like sequence in hepatitis B virus core antigen is not required for e antigen generation and may not be part of an aspartic-acid type protease. *Journal of Virology* **63**, 2598-2604.

NAVIA, M.A.; FITZGERALD, P.M.D; McKEEVER, B.M.; LEU, C-T.; HEIMBACH, J.C.; HERBER, W.K.; SIGAL, I.S.; DARKE, P.L. & SPRINGER, J.P. (1989). Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337**, 615-620.

NEILL, S.D.; HEMSTROM, C.; VIRTANEN, A. & NEVINS, J. (1990). An adenovirus E4 gene product trans-activates E2 transcription and stimulates stable E2F binding through a direct association with E2F. *Proceedings of the National Academy of Sciences (USA)* **87**, 2008-2012.

NERMUT, M.V. (1979). Structural elements in adenovirus cores. Evidence for a core shell and linear structures in 'relaxed cores'. *Archives of Virology* **64**, 175-196.

NERMUT, M.V. (1984). The architecture of adenoviruses. In: *Adenoviruses* pp5-34. Ed. by: H.S. Ginsberg. Plenum Press, New York.

NEURATH, H. (1989). The diversity of proteolytic enzymes. In: *Proteolytic enzymes, a practical approach* pp1-13. Ed. by: R.J. Beynon and J.S. Bond. IRL Press, Oxford University Press.

NEVINS, J.R. (1987). Regulation of early adenovirus gene expression. *Microbiological Reviews* **51**, 419-430.

NEVINS, J.R. (1991). Transcriptional activation by viral regulatory proteins. *Trends in Biochemical Sciences* **16**, 435-439.

NICKLIN, M.J.H.; TOYODA, H.; MURRAY, M.G. & WIMMER, E. (1986). Proteolytic processing in the replication of polio and related viruses. *Biotechnology* **4**, 33-42.

NORRBY, E. (1969). The relationship between soluble antigens and the virion of adenovirus type 3. III. Immunologic characteristics. *Virology* **37**, 565-576.

NORRBY, E.; BARTH, A.; BOULANGER, T.; DREIZIN, R.; GINSBERG, H.S.; KALTER, S.S.; KAWAMURU, H.; ROWE, W.P.; RUSSELL, W.C.; SCHLESINGER, W. & WIGAND, R. (1976). Adenoviridae. *Intervirology* **7**, 117-125.

OROSZLAN, S. (1989). Biosynthesis and proteolytic processing of retroviral proteins: an overview. In:

Viral Proteinases as Targets for Chemotherapy pp87-100. Ed by: H-G. Krausslich; S. Oroszlan and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

ORR, D.C.; LONG, A.C.; KAY, J.; DUNN, B.M. & CAMERON, J. (1989). Hydrolysis of a series of synthetic peptide substrates by the human rhinovirus 14 3C proteinase, cloned and expressed in Escherichia coli. *Journal of General Virology* **70**, 2931-2942.

OSWALD, M. & VON DER HELM, K. (1991). Fibronectin is a non-viral substrate for the HIV proteinase. *FEBS Letters* **292**, 298-300.

PALLAI, P.V.; BURKHARDT, F.; SKOOG, M.; SCHREINER, K.; BAX, P.; COHEN, K.A.; HANSEN, G.; PALLADINO, D.E.H.; HARRIS, K.S.; NICKLIN, M.J. & WIMMER, E. (1989). Cleavage of synthetic peptides by purified poliovirus 3C proteinase. *Journal of Biological Chemistry* **264**, 9738-9741.

PALMENBERG, A.C. & PARKS, G.D. (1989). Cleavage specificity of EMC 3C proteinase. In: *Viral Proteinases as Targets for Chemotherapy* pp33-37. Ed. by: H-G. Krausslich, S. Oroszlan, and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

PARKS, T.D. & DOUGHERTY, W.G. (1991). Substrate recognition of the NIa proteinase of two potyviruses involves multiple domains: characterisation using genetically engineered hybrid proteinase molecules. *Virology* **182**, 17-27.

PASTAN, I.; SETH, P.; FITZGERALD, D. & WILLIGHAM, M. (1987). Adenovirus entry into cells: some new observations on an old problem. In: *Concepts in Viral Pathogenesis* pp141-146. Ed. by: A. Notkins and M.B.A. Oldstone. Springer-Verlag, New York.

PEARL, L.H. & TAYLOR, W.R. (1987). A structural model for retroviral proteases. *Nature* **329**, 351-354.

PEARSON, G.D. & HANAWALT, P.C. (1971). Isolation of DNA replication complexes from uninfected and adenovirus infected Hela cells. *Journal of Molecular Biology* **62**, 65-80.

PEREIRA, H.G. & RUSSELL, W.C. (1969). Crystallographic data for the hexon of adenovirus type 5. *Nature* **222**, 1165-1166.

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

PERRICAUDET, M.; AKUSJARVI, A.; VIRTANEN, A. & PETTERSSON, U. (1979). Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. *Nature* **281**, 694-696.

PETTIT, S.C.; HOROWITZ, M.S. & ENGLER, J.A. (1989). Mutations of the precursor to the terminal protein of adenovirus serotypes 2 and 5. *Journal of Virology* **63**, 5244-5250.

PHILIPSON, L. (1983). Structure and assembly of adenoviruses. *Current Topics in Microbiology and Immunology* **109**, 1-52.

PINA, M. & GREEN, M. (1969). Biochemical studies on adenovirus multiplication. XIV. Macromolecule and enzyme synthesis in cells replicating oncogenic and non-oncogenic human adenoviruses. *Virology* **38**, 573-586.

POLGAR, L. (1987). Structure and function of serine proteases. In: *Hydrolytic Enzymes* pp159-200. Ed. by: A. Neuberger and K. Brocklehurst. Amsterdam: Elsevier.

POORMAN, R.A.; TOMASSELLI, A.G.; HEINRIKSON, R.L. & KEZDY, F.J. (1991). A cumulative specificity model for proteases from human immunodeficiency virus types 1 and 2, inferred from statistical analysis of an extended substrate data base. *Journal of Biological Chemistry* **266**, 14554-14561.

PORATH, J.; CARLSSON, J.; OLSSON, I. & BELFRAGE, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599.

POSSEE, R.D. & HOWARD, S.C. (1987). Analysis of the polyhedrin gene promoter of Autographa californica nuclear polyhedrosis virus. *Nucleic Acids Research* **15**, 10233-10248.

PRAGE, L.; PETTERSSON, U.; HOGLUND, S.; LONDBERG-HOLM, K. & PHILIPSON, L. (1970). Structural proteins of adenoviruses. IV. Sequential degradation of the adenovirus type 2 virion. *Virology* **42**, 341-358.

PRESTON, V.G.; RIXON, F.J.; McDOUGALL, I.M.; McGREGOR, M. & KOBAISI, A.L. (1992). Processing of the Herpes Simplex virus assembly protein ICP35 near its carboxy terminal end requires the product of the whole of the UL26 reading frame. *Virology* **186**, 87-98.

PREUGSCHAT, F.; LENCHES, E.M. & STRAUSS, J.H. (1991). Flavivirus enzyme-substrate interactions studied with chimeric proteinases: identification of an intragenic locus important for substrate recognition. *Journal of Virology* **65**, 4749-4758.

RABILLOUD, T.; CARPENTLER, G. & TARROUX, P. (1988). Improvement and simplification of low-background silver staining of proteins by using sodium dithionite. *Electrophoresis* **9**, 288-291.

RAJU, B. & DESHPANDE, M.S. (1991). Substrate analog inhibitors of HIV-1 protease containing phenylnorstatine as a transition state element. *Biochemical and Biophysical Research Communications* **180**, 181-186.

RAMAGE, R. & GREEN, J. (1987). Ng-2,2,5,7,8-Pentamethylchroman-6-sulphonyl-L-arginine: a new acid labile derivative for peptide synthesis. *Tetrahedron Letters* **28**, 2287-2290.

RAWLE, F.C.; TOLLEFSON, A.E.; WOLD, W.S.M. & GOODING, C.R. (1989). Mouse anti-adenovirus cytotoxic T-lymphocytes. Inhibition of lysis by E3 gp19K, but not E3 14.7K protein. *Journal of Immunology* **143**, 2031-2037.

RIECHMANN, J.L.; LAIN, S. & GARCIA, J.A. (1992). Highlights and prospects of potyvirus molecular biology. *Journal of General Virology* **73**, 1-16.

ROBERTS, M.M.; WHITE, J.L.; GRUTTER, M.G. & BURNETT, R.M. (1986). Three dimensional structure of the adenovirus major coat protein hexon. *Science* **232**, 1148-1151.

ROBERTS, N.A.; MARTIN, J.A.; KINCHINGTON, D.; BROADHURST, A.V.; CRAIG, J.C.; DUNCAN, I.B.; GALPIN, S.A.; HANDA, B.K.; KAY, J.; KROHN, A.; LAMBERT, R.W.; MERRETT, J.H.; MILLS, J.S.; PARKES, K.E.B.; REDSHAW, S.; RITCHIE, A.J.; TAYLOR, D.L.; THOMAS, G.J. & MACHIN, P.J. (1990). Rational design of peptide-based HIV proteinase inhibitors. *Science* **248**, 358-361.

ROBERTS, R.J.; AKUSJARVI, G.; ALESTROM, P.; GELINAS, R.E.; GINGERAS, T.R. & SCIAKY, D. (1986). A consensus sequence for the adenovirus-2 genome. In: *Adenovirus DNA. The Viral Genome and its Expression* pp1-51. Ed. by: W. Doerfler. Martinus Nijhoff Publishing, Boston.

ROSENBERG, A.H.; LADE, B.N.; CHUI, D-S.; LIN, S-W.; DUNN, J.J. & STUDIER, F.W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125-135.

ROSENFELD, P.J.; O'NEILL, E.A.; WIDES, R.J. & KELLY, T.J. (1987). Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Molecular and Cell Biology* **7**, 875-886.

ROTH, M. (1971). Fluorescence reaction of amino acids. *Analytical Chemistry* **43**, 880-882.

ROWE, W.P.; HUEBNER, R.J.; GILMORE, L.K.; PARROTT, R.H. & WARD, T.G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine* **84**, 570-573.

RUIGROK, R.W.H.; BARGE, A.; ALBIGES-RIZO, C. & DAYAN, S. (1990). Structure of adenovirus

fibre. *Journal of Molecular Biology* **215**, 589-596.

RULEY, H.E. (1983). Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. *Nature* **304**, 602-606.

RUOSLAHTI, E. & PIERSCHBACHER, M.D. (1986). Arg-Gly-Asp: a versatile cell recognition signal. *Cell* **44**, 517-518.

RUSSEL, M. (1991). Filamentous phage assembly. *Molecular Microbiology* **5**, 1607-1613.

RUSSELL, W.C. & BLAIR, G.E. (1977). Polypeptide phosphorylation in adenovirus-infected cells. *Journal of General Virology* **34**, 19-35.

RUSSELL, W.C.; LAVER, W.G. & SANDERSON, P.J. (1968). Internal components of adenovirus. *Nature* **219**, 1127-1130.

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

RUSSELL, W.C.; WEBSTER, A.; LEITH, I.R. & KEMP, G.D. (1989). Phosphorylation of adenovirus DNA-binding protein. *Journal of General Virology* **70**, 3249-3259.

RYAN, M.D.; KING, A.M.Q. & THOMAS, G.P. (1991). Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *Journal of General Virology* **72**, 2727-2732.

SALAS, M. (1991). Protein-priming of DNA replication. *Annual Reviews of Biochemistry* **60**, 39-71.

SALVESON, G. & NAGASE, H. (1989). Inhibition of proteolytic enzymes. In: *Proteolytic Enzymes, a practical approach* pp83-104. Ed by: R.J. Beynon and J.S. Bond. IRL Press, Oxford University Press.

SANTORO, C.; MERMOD, N.; ANDREWS, P.C. & TJIAN, R. (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* **334**, 218-224.

SARNOW, P.; HEARING, P.; ANDERSON, C.W.; HALBERT, D.N.; SHENK, T. & LEVINE, A.J. (1984). Adenovirus early region 1B 58000 dalton tumor antigen is physically associated with an early region 4 25000 dalton protein in productively infected cells. *Journal of Virology* **49**, 692-700.

SAROV, I. & JOLIK, W.K. (1972). Studies on the nature and location of the capsid polypeptides of vaccinia virions. *Virology* **50**, 579-592.

SASSENFELD, H.M. (1990). Engineering proteins for purification. *Trends in Biotechnology* **8**, 88-93.

SCHAACK, J.; HO, W.Y-W.; FREIMUTH, P. & SHENK, T. (1990). Adenovirus terminal protein mediates both nuclear matrix association and efficient transcription of adenovirus DNA. *Genes and Development* **4**, 1197-1208.

SCHEIDTMANN, K.H. (1989). Immunological detection of proteins of known sequence. In: *Protein Structure, a practical approach* pp93-115. Ed. by: T.E. Creighton. IRL Press, Oxford University Press.

SCHLESINGER, S. & SCHLESINGER, M.J. (1990). Replication of Togaviridae and Flaviviridae. In: *Fields Virology Volume 1. Second Edition* pp697-712. Ed. by: B.N. Fields and D.M. Knipe. Raven Press Ltd., New York.

SCHNEIDER, J. & KENT, S.B.H. (1988). Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. *Cell* **54**, 363-368.

SCHNEIDER, R.J.; WEINBERGER, C. & SHENK, T. (1984). The adenovirus VA1 RNA functions during

initiation of translation in virus-infected cells. *Cell* **37**, 291-298.

SCUPHAM, R.K.; JONES, K.J.; SAGIK, J.B. & BOSE, H.R. (1977). Virus-directed post-translational cleavage in Sindbis virus infected cells. *Journal of Virology* **22**, 568-571.

SEE, Y.P. & JACKOWSKI, G. (1989). Estimating molecular weights of polypeptides by SDS gel electrophoresis. In: *Protein Structure, a practical approach* pp1-21. Ed by: T.E. Creighton. IRL Press, Oxford University Press.

SEIBERG, M.; ALONI, Y. & LEVINE, A.J. (1989). The adenovirus type 2 DNA-binding protein interacts with the major late promoter attenuated RNA. *Journal of Virology* **63**, 1134-1141.

SHARP, P.A. (1984). Adenovirus transcription. In: *The Adenoviruses* pp173-204. Ed. by: H.S. Ginsberg. Plenum Press, New York.

SHAW, A.R. & ZIFF, E.B. (1980). Transcripts from the adenovirus major late promoter yield a family of 3' coterminal RNAs during early infection and five families at late times. *Cell* **22**, 905-916.

SHIRAKO, Y. & STRAUSS, J.H. (1990). Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus non-structural polyprotein P123. *Virology* **177**, 54-64.

SHOEMAN, R.L.; KESSELMEIER, C.; MOTHES, E.; HINER, B. & TRAUB, P. (1991). Non-viral cellular substrates for human immunodeficiency virus type 1 protease. *FEBS Letters* **278**, 199-203.

SHU, L.; HORWITZ, M.S. & ENGLER, J.A. (1987). Expression of enzymatically active adenovirus DNA polymersase from cloned DNA requires sequences upstream of the main open reading frame. *Virology* **161**, 520-526.

SKERN, T.; SOMMERGRUBER, W.; AUER, H.; VOLKMANN, P.; ZORN, M.; LIEBIG, H-D.; FESSEL, F.; BLAAS, D. & KUECHLER, E. (1991). Substrate requirements of human rhinoviral 2A proteinase. *Virology* **181**, 46-54.

SLABAUGH, M.B. & ROSEMAN, N.A. (1989). Retroviral protease-like gene in the vaccinia virus genome. *Proceedings of the National Academy of Sciences (USA)* **86**, 4152-4155.

SMART, J.E. & STILLMAN, W. (1982). Adenovirus terminal protein precursor. *Journal of Biological Chemistry* **257**, 13499-13506.

SMITH, D.B. & JOHNSON, K.S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathionine s-transferase. *Gene* **67**, 31-40.

SMITH, M.C.; FURMAN, T.C.; INGOLIA, T.D. & PIDGEON, C. (1988). Chelating peptide-immobilised metal affinity chromatography. *Journal of Biological Chemistry* **263**, 7211-7215.

SOMMERGRUBER, W.; ZORN, M.; BLAAS, D.; FESSL, F.; VOLKMANN, P.; MAURER-FOGY, I.; PALLAI, P.; MERLUZZI, V.; MATTEO, M.; SKERN, T. & KEUCHLER, E. (1989). Polypeptide 2A of human rhinovirus type 2: identification as a protease and characterisation by mutational analysis. *Virology* **169**, 68-77.

SONENBERG, N. (1990). Poliovirus translation. *Current Topics in Microbiology and Immunology* **161**, 23-47.

SORGER, P.K.; AMMERER, G. & SHORE, D. (1989). Identification and purification of sequence specific DNA- binding proteins. In: *Protein Function, a practical approach* pp199-223. Ed by: T.E. Creighton. IRL Press, Oxford University Press.

STADEN, R. (1986). The current status and portability of our sequence handling software. *Nucleic Acids Research* **14**, 217-231.

STEWART, L. & VOGT, V.M. (1991). Trans-acting viral protease is necessary and sufficient for activation of avian leukosis virus reverse transcriptase. *Journal of Virology* **65**, 6218-6231.

STEWART, P.L.; BURNETT, R.M.; CYRKLAFF, M. & FULLER, S.D. (1991). Image reconstruction reveals the complex molecular organisation of the adenovirus. *Cell* **67**, 145-154.

STILLMAN, B.W. (1989). Initiation of eukaryotic DNA replication *in vitro*. *Annual Reviews of Cell Biology* **5**, 197-245.

STILLMAN, B.W.; LEWIS, J.B.; CHOW, L.T.; MATHEWS, M.B. & SMART, J.E. (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. *Cell* **23**, 497-508.

STRAUS, S.E. (1984). Adenovirus infections in humans. In: The Adenoviruses pp451-496. Ed by: H.S. Ginsberg. Plenum Press, New York.

STRAUSS, E.G.; RICE, C.M. & STRAUSS, J.H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133**, 92-110.

STRAUSS, E.G.; SCHMALJOHN, A.L.; GRIFFIN, D.E. & STRAUSS, J.H. (1987). Structure-function relationships in the glycoproteins of alphaviruses. In: *Positive-strand RNA Viruses* pp365-378. Ed. by: M.A. Brinton & R. Rueckert. Alan R.Liss Inc., New York.

STRAUSS, J.H. & STRAUSS, E.G. (1990). Alphavirus proteinases. *Seminars in Virology* **1**, 347-356.

STRAUSS, J.H.; STRAUSS, E.G.; HAHN, C.S.; HAHN, Y.S.; GALLER, R.; HARDY, W.R. & RICE, C.M. (1987). Replication of alphaviruses and flaviviruses: proteolytic processing of polyproteins. In: *Positive-strand RNA Viruses* pp209-225. Ed. by: M.A. Brinton and R. Rueckert. Alan R. Liss Inc., New York.

STUDIER, F.W.; ROSENBERG, A.H.; DUNN, J.J. & DUBENDORFF, J.W. (1990). Use of T7 RNA polymerase to direct the expression of cloned genes. *Methods in Enzymology* **185**, 60-89.

STUIVER, M.H. & VAN DER VLIET, P.C. (1990). Adenovirus DNA binding protein forms a multimeric protein complex with double-stranded DNA and enhances binding of nuclear factor I. *Journal of Virology* **65**, 379-389.

STUNNENBERG, H.G.; LANGE, H.; PHILIPSON, L.; VAN MILTENBURG, R.T. AND VAN DER VLIET, P.C. (1988). High expression of functional adenovirus DNA polymerase and precursor terminal protein using recombinant vaccinia virus. *Nucleic Acids Research* **16**, 2431-2444.

SUMMERS, M.D. & SMITH, G.E. (1987). *A manual of methods for baculovirus vectors and insect cell culture procedures*. Texas Agricultural Experimental Station, College Station, Texas. Bulletin 1555.

SUNG, M.T.; CAO, T.M.; LISCHWE, M.A. & COLEMAN, R.T. (1983). Molecular processing of adenovirus proteins. *Journal of Biological Chemistry* **258**, 8266-8272.

SUSSENBACH, J.S. (1984). The structure of the adenovirus genome. In: *The Adenoviruses* pp35-124. Ed. by: H.S. Ginsberg. Plenum Press, New York.

TAMANOI, F. & STILLMAN, B.W. (1982). Function of adenovirus terminal protein in the initiation of DNA replication. *Proceedings of the National Academy of Sciences (USA)* **79**, 2221-2225.

TASHIRO, M.; CIBOROWSKI, P.; REINACHER, M.; PULVERER, G.; KLENK, H-D. & ROTT, R. (1987). Synergistic role of staphylococcal proteases in the induction of influenza virus pathogenicity. *Virology* **157**, 421-430.

TEMPERLEY, S.M. & HAY, R.T. (1991). Replication of adenovirus type 4 DNA by a purified fraction from infected cells. *Nucleic Acids Research* **19**, 3243-3249.

TEMPERLEY, S.M. & HAY, R.T. (1992). Recognition of the adenovirus type 2 origin of DNA replication by virally encoded DNA polymerase and preterminal proteins. *EMBO Journal* **11**, 761-768.

THIMMAPPAYA, B.; WEINBERGER, C.; SCHNEIDER, R.J. & SHENK, T. (1982). Adenovirus VA1 RNA is required for efficient translation of viral mRNAs at late times after infection. *Cell* **31**, 543-551.

TOLLEFSON, A.E.; KRAJCSI, P.; YEI, S.; CARLIN, C.R. & WOLD, W.S.M. (1990). A 10400 molecular weight membrane protein is coded by region E3 of adenovirus. *Journal of Virology* **64**, 794-801.

TOMASSELLI, A.G.; HOWE, W.J.; HUI, J.O.; SAWYER, T.K.; REARDON, I.M.; DECAMP, D.L.; CRAIK, C.S. & HEINRIKSON, R.L. (1991). Calcium-free calmodulin is a substrate of proteases from human immunodeficiency viruses 1 and 2. *Proteins: Structure, Function and Genetics* **10**, 1-9.

TOMASSELLI, A.G.; HUI, J.O.; SAWYER, T.K.; STAPLES, D.J.; BANNOW, C.A.; REARDON, I.M.; CHAUDHARY, V.K.; FRYLING, C.M.; PASTAN, I.; FITZGERALD, D.J. & HEINRIKSON, R.L. (1990). Proteases from human immunodeficiency virus and avian myeloblastosis virus show distinct specificities in hydrolysis of multidomain protein substrates. *Journal of Virology* **64**, 3157-3161.

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

TOOGOOD, C.I.A.; MURALI, R.; BURNETT, R.M. & HAY, R.T. (1989). The adenovirus type 40 hexon: sequence, predicted structure and relationship to other hexons. *Journal of General Virology* **70**, 3203-3214.

TOZSER, J.; GUSTCHINA, A.; WEBER, I.T.; BLAHA, I.; WONDRAK, E.M. & OROSZLAN, S. (1991). Studies on the role of the S4 substrate binding site of HIV proteases. *FEBS Letters* **279**, 356-360.

TREMBLAY, M.L.; DERY, C.V.; TALBOT, B.G. & WEBER, J. (1983). *In vitro* cleavage specificity of the adenovirus type 2 proteinase. *Biochimica et biophysica acta* **743**, 239-245.

TRENTIN, J.J.; YABE, Y. & TAYLOR, G. (1962). The quest for human cancer viruses. *Science* **137**, 835-841.

TSUZUKI, J. & LUFTIG, R.B. (1983). The adenovirus type 5 capsid protein IIIa is phosphorylated during an early stage of infection of Hela cells. *Virology* **129**, 529-533.

VALENTINE, R.C. & PEREIRA, H.G. (1965). Antigens and structure of the adenovirus. *Journal of Molecular Biology* **13**, 13-20.

VAN DEN ELSEN, P.; HOUWELING, A. & VAN DER EB, A.J. (1983). Expression of region E1B of human adenovirus in the absence of region E1A is not sufficient for complete transformation. *Virology* **128**, 377-390.

VAN OOSTRUM, J. & BURNETT, R.M. (1985). Molecular composition of the adenovirus type 2 virion. *Journal of Virology* **56**, 439-448.

VANSLYKE, J.K.; FRANKE, C.A. & HRUBY, D.E. (1991a). Proteolytic maturation of vaccinia virus core proteins: identification of a conserved motif at the N-terminus of the 4b and 25K virion proteins. *Journal of General Virology* **72**, 411-416.

VANSLYKE, J.K.; WHITEHEAD, S.S.; WILSON, E.M. & HRUBY, D.E. (1991b). The multistep proteolytic maturation pathway utilised by vaccinia virus P4a protein: a degenerate conserved cleavage motif within core proteins. *Virology* **183**, 467-478.

VARGA, M.J.; BERGMAN, T. & 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. & FLINT, S.J. (1983). The structure of nucleoprotein cores releases by

adenovirions. *Nucleic Acids Research* **11**, 441-460.

VERCHOT, J.; KOONIN, E.V. & CARRINGTON, J.C. (1991). The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* **185**, 527-535.

VERRIJZER, C.P.; KAL, A.J. & VAN DER VLIET, P.C. (1990). The DNA binding domain (POU domain) of transcription factor oct-1 suffices for stimulation of DNA replication. *EMBO Journal* **9**, 1883-1888.

VIRTANEN, A.; GILARDI, P.; NASLUND, A.; LEMOULLEC, J-M.; PETTERSSON, U. & PERRICAUDET, M. (1984). mRNAs from human adenovirus 2 early region 4. *Journal of Virology* **51**, 822-831.

VIRTANEN, A. & PETTERSSON, U. (1985). Organisation of early region 1B of human adenovirus type 2: identification of four differentially spliced mRNAs. *Journal of Virology* **54**, 383-391.

VON DER HELM, K. (1977). Cleavage of Rous sarcoma viral polypeptide precursor into internal structural proteins *in vitro* involves viral protein p15. *Proceedings of the National Academy of Sciences (USA)* **74**, 911-915.

VOS, H.L.; VAN DER LEE, F.M. & SUSSENBACH, J.S. (1988). The binding of *in vitro* synthesised adenovirus DNA binding protein to single stranded DNA is stimulated by zinc ions. *FEBS Letters* **239**, 251-254.

VOS, P.; VERVER, J.; JAEGLE, M.; WELLINK, J.; VAN KAMMEN, A. & GOLDBACH, R. (1988). Two viral proteins involved in the proteolytic processing of cowpea mosaic virus polyproteins. *Nucleic Acids Research* **16**, 1967-1985.

WATSON, C.J. & HAY, R.T. (1990). Expression of adenovirus type 2 DNA polymerase in insect cells infected with a recombinant baculovirus. *Nucleic Acids Research* **18**, 1167-1173.

WEBER, I.T. (1989). Structural alignment of retroviral protease sequences. *Gene* **85**, 565-566.

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

WEBER, J.M. & ANDERSON, C.W. (1988). Identification of the gene coding for the precursor of adenovirus core protein X. *Journal of Virology* **62**, 1741-1745.

WELLINK, J. & VAN KAMMEN, A. (1988). Proteases involved in the processing of viral polyproteins. *Archives of Virology* **98**, 1-26.

WELLINK, J.; VERVER, J.; VERKERK, R.; SCHEPER, G.; VAN BOKHOVEN, H. & VAN KAMMEN, A. (1989). Processing of the cowpea mosaic virus polyproteins. In: *Viral Proteinases as Targets for Chemotherapy* pp39-46. Ed. by: H-G. Krausslich, S. Oroszlan and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

WHITE, E.; DENTON, A. & STILLMAN, B. (1988). Role of the adenovirus E1b 19000 dalton tumour antigen in regulating early gene expression. *Journal of Virology* **62**, 3445-3454.

WHITTAKER, J.R. & PEREZ-VILLASENOR, J. (1968). Chemical modification of papain. I. Reaction with the chloromethyl ketones of phenylalanine and lysine and with phenylmethylsulphonyl fluoride. *Archives of Biochemistry and Biophysics* **124**, 70-78.

WHYTE, P.; BUCHKOVICH, K.J.; HOROWITZ, J.M.; FRIEND, S.H.; RAYBUCK, M.; WEINBERG, R.A. & HARLOW, E. (1988). Association between an oncogene and anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* **334**, 124-129.

WILLENBROCK, F. & BROCKLEHURST, K. (1985). A general framework of cysteine proteinase

mechanism deduced from studies on enzymes with structurally different analogous catalytic-site residues Asp-158 and -161 (papain and actinidin), Gly-196 (cathepsin B) and Asn-165 (cathepsin H). *Biochemical Journal* **227**, 521-528.

WIMMER, E. (1989). Proteolytic processing in viral replication. In: *Viral Proteinases as Targets for Chemotherapy* pp1-15. Ed. by: H-G. Krausslich, S. Oroszlan, and E. Wimmer. Cold Spring Harbor Laboratory Press, New York.

WINTER, N. & D'HALLUIN, J-C. (1991). Regulation of biosynthesis of subgroup C adenovirus protein IVa2. *Journal of Virology* **65**, 5250-5259.

WISKERCHEN, M.; BELZER, S.K. & COLLET, M.S. (1991). Pestivirus gene expression: the first protein product of bovine viral diarrhoea virus large open reading frame, p20, possesses proteolytic activity. *Journal of Virology* **65**, 4508-4514.

WISKERCHEN, M. & COLLET, M.S. (1991). Pestivirus gene expression: protein p80 of bovine viral diarrhoea virus is a proteinase involved in polyprotein processing. *Virology* **184**, 341-350.

WLODAWER, A.; MILLER, M.; JASKOLSKI, M.; SATHYANARAYANA, B.K.; BALDWIN, E.; WEBER, I.T.; SELK, L.M.; CLAWSON, L.; SCHNEIDER, J. & KENT, S.B.H. (1989). Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **245**, 616-621.

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

WYCKOFF, E.E.; HERSHEY, J.W.B. & EHRENFELD, E. (1990). Eukaryotic initiation factor 3 is required for poliovirus 2A protease induced cleavage of the p220 component of eukaryotic initiation factor 4F. *Proceedings of the National Academy of Sciences (USA)* **87**, 9529-9533.

YEH-KAI, L.; AKUSJARVI, G.; ALESTROM, P.; PETTERSSON, U.; TREMBLAY, M. & WEBER, J.M. (1983). Genetic identification of an endoproteinase encoded by the adenovirus genome. *Journal of Molecular Biology* **167**, 217-222.

YU, S.F. & LLOYD, R.E. (1991). Identification of essential amino acid residues in the functional activity of poliovirus 2A protease. *Virology* **182**, 615-625.

YU, S.F. & LLOYD, R.E. (1992). Characterisation of the roles of conserved cysteine and histidine residues in poliovirus 2A protease. *Virology* **186**, 725-735.

ZHANG, Z-Y.; REARDON, I.M.; HUI, J.O.; O'CONNELL, K.L.; POORMAN, R.A.; TOMASSELLI, A.G. & HEINRIKSON, R.L. (1991). Zinc inhibition of renin and the protease from human immunodeficiency virus type 1. *Biochemistry* **30**, 8717-8721.

ZHAO, L-J. & PADMANABHAN, R. (1988). Nuclear transport of adenovirus DNA polymerase is facilitated by interaction with preterminal protein. *Cell* **55**, 1005-1015.