The Activation of Adenovirus Type 2 Protease and the Function of the Activating Peptide.

By Brett David Pollard

The University of St. Andrews,

St. Andrews

A thesis presented for the degree of Doctor of Philosophy at the University of St. Andrews, March 2001

DECLARATION

I, Brett David Pollard, hereby certify that this thesis, which is approximately 35000 words in length, has been written by me, that it is a record of work carried out by me and that it has not been submitted in any previous application for higher degree.

Date 21/5/01.. Signature of candidate

I was admitted as research student in September 1997 and as a candidate for the degree of Doctor of Philosophy in September 1997: the higher study for which this is a record was carried out in the faculty of Sciences of the University of St. Andrews between 1997 and 2000.

Date 21/5/01.. Signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 21/05/01 Signature of supervisor ....

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.

Date 21/5/01.. Signature of candidate .

Acknowledgements

First and foremost I would like to thank Dr. Graham Kemp for his endless patience and guidance throughout my Ph.D. Thanks also to Prof. Ron Hay and Rick Randall for their advice and ideas, and to Paul Talbot, Ellis Jaffray and Alex Houston for there technical assistance.

Throughout my education many people have proved sources of inspiration and support, my father, Carl and Joanne, Les and Anne Sykes, Nigel Owen, Dr. Barry Fletcher, Paul Richards, and Nicholas Walker to name but a few. It has been a privileged to have had the friendship and support of these people, and many others in St. Andrews, and further afield.

I would like to thank my mother for the sacrifices she made to ensure I received the best possible education, without which I would not have been where I am now. Her constant and unconditional encouragement, love and support are duly appreciated.

This work was supported by the BBSRC.

Abstract

Some 42 serotypes of Adenovirus are infectious to Humans. To be able to prevent infection using vaccination would be difficult to say the least, since one might have to administer as many as 42 vaccines, one for each serotype. However, if one could produce an antiviral agent specific to an event within the virus's life cycle to prevent the infectiousness one agent could potentially be effective to all serotypes. One such target is the 23 kD protease enzyme responsible primarily for maturation cleavages of structural proteins comprising the viral capsid and the cytoskeletal proteins to escape the cell. The work described here is aimed at understanding more about the role of the peptide responsible for the activation of the protease and determining whether the unusual activation mechanism of this protease is a suitable target for the therapeutic inhibition.

Another aim of the study was to attempt to develop an *in vivo* protease inhibitor assay to complement *in vitro* activation studies which required the development of a single step purification protocol for the protease. However, attempts to stably integrate the protease and pVI in HeLa and 293 cell lines were not successful due to the toxicity of the basal level of expression and the cocktail of antibiotics required for the selection for integration. Use of transient transfections illustrated that the protease could cleave cytokeratin *in vivo* in the absence of activating peptide, and suggested that the protease was still toxic to the cells. Studies revealed that the protease from the human serotype 2 could be activated by the peptide from the avian Adenovirus responsible for Egg Drop Syndrome (EDS virus). Since the sequence of the EDS peptide and protease are the most distinct in evolutionary terms from the human type 2, such a cross-activation suggests that any interference with the activation system could be effective across the serotype range.

Results obtained here show that the length of the peptide is critical, and contrary to previous reports no evidence was obtained to support the suggestion that the peptide VEGGS is capable of activating the protease. Mutation studies on the contiguous stretch of 4 basic residues within the activating peptide sequence suggested that this does not have a critical role in activation. However, parallel studies showed that this motif could act as a nuclear localisation signal as it is capable of targeting a chimeric protein to the nucleus in a model system. Finally a rapid assay was developed for assessing the binding affinity of peptides to the protease, and results obtained with this suggested variants which are worthy of further study in the search for a possible inhibitor.

List of Abbreviations

A	Adenine
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
bp	Base pairs
BSA	Bovine serum albumen
C	Cytosine
CE	Capillary Electrophoresis
CDK	Cyclin-dependant kinase
cDNA	Complementary DNA
CM	Carboxymethyl
DAPI	4, 6 diamidino-2-phenylindole
ddH_2O	deionised distilled water
D-MEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
E. coli	*Escherichia coli*
ECL	Enhanced chemiluminesence
EDTA	Ethylenediaminetetracetic acid
E1-E4	Early genes or proteins 1 to 4
FCS	Foetal calf serum
FITC	Fluorescein isothiocyanate
G	Guanine
GST	Glutathione S-transferase
GTP	Guanosine triphosphate
h	Hour(s)
HCl	Hydrochloric acid
HIV	Human immunodeficiency virus
hpi	Hours(s) post infection
HPLC	High performance liquid chromatography
hpt	Hour(s) post transfection
HSA	Human serum albumin
Ig	Immunoglobulin
IPTG	Isopropyl-β-D-thiogalactopyranose
KCl	Potassium chloride
kD	Kilo Dalton
LB	Luria-Bertani broth
L1-L5	Late genes or proteins 1 to 5
MHC	Major histocompatibility complex
min	Minute(s)

MLP	Major late promoter
MLTU	Major late transcription unit
mRNA	Messenger ribonucleic acid
MW	Molecular weight
NES	Nuclear export sequence
NF	Nuclear factor
NLS	Nuclear localisation signal
NPC	Nuclear pore complex
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pfu	Plaque forming units
pH	pondus hydrogen ($-\log_{10}[H^+]$)
Pi	Pi orbital electrons
pRB	Retinoblastoma susceptibility gene product
pTP	Pre-terminal protein
PVDF	Polyvinylidene difluoride
RNA	Ribonucleic acid
SBTI	Soya bean trypsin inhibitor
SDS	Sodium dodecyl sulphate
Sf 9	*Spodoptera frugiperda*
ssDNA	Single stranded DNA
SV40	Simian virus type 40
T	Thymine
TBE	Tris-borate-EDTA buffer
TEMED	N. N, N, N' tetramethylethylenediamine
TF	Transcription factor
TFA	Trifluoracetic acid
TLCK	Tosyl-lysine-chloromethyl ketone
TP	Terminal protein
TPCK	Tosyl-phenylalanine-chloromethyl ketone
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
ts	Temperature sensitive
UV	Ultra-violet
VA RNA	Viral associated RNA
v/v	volume per volume ratio
Wt	Wild-type
w/v	Weight per volume ratio

Abbreviations for Amino Acids and Side-chains

Alanine	ala	A	$-CH_3$

Arginine	arg	R	$-(CH_2)_3-NH-C=NH-NH_2$
Aspargine	asn	N	$-CH_2-CONH_2$
Aspartic acid	asp	D	$-CH_2-COOH$
Cysteine	cys	C	$-CH_2-SH$
Glutamine	gln	Q	$-CH_2-CH_2-CONH_2$
Glutamic acid	glu	E	$-CH_2-CH_2-COOH$
Glycine	gly	G	$-H$
Histidine	his	H	$-CH_2$-imidazole
Isoleucine	ile	I	$-CH(CH_3)-CH_2-CH_3$
Leucine	leu	L	$-CH_2-CH(CH_3)_2$
Lysine lys	K		$-(CH_2)_4-NH_2$
Methionine	met	M	$-CH_2-CH_2-S-CH_3$
Phenylalanine	phe	F	$-CH_2$-phi
Proline	pro	P	-[N]-$(CH_2)_3$-[CH]
Serine ser	S		$-CH_2-OH$
Threonine	thr	T	$-CH-(CH_3)-OH$
Tryptophan	try	W	$-CH_2$-indole
Tyrosine	tyr	Y	$-CH_2$-phi-OH
Valine val	V		$-CH-(CH_3)_2$

Genetic Code

```
TTT phe F TCT ser S TAT tyr Y TGT cys C
TTC phe F TCC ser S TAC tyr Y TGC cys C
TTA leu L TCA ser S TAA OCH Z TGA OPA Z
TTG leu L TCG ser S TAG AMB Z TGG try W

CTT leu L CCT pro P CAT his H CGT arg R
CTC leu L CCC pro P CAC his H CGC arg R
CTA leu L CCA pro P CAA gln Q CGA arg R
CTG leu L CCG pro P CAG gln Q CGG arg R

ATT ile I ACT thr T AAT asn N AGT ser S
ATC ile I ACC thr T AAC asn N AGC ser S
ATA ile I ACA thr T AAA lys K AGA arg R
ATG met M ACG thr T AAG lys K AGG arg R

GTT val V GCT ala A GAT asp D GGT gly G
GTC val V GCC ala A GAC asp D GGC gly G
GTA val V GCA ala A GAA glu E GGA gly G
GTG val V GCG ala A GAG glu E GGG gly G
```

INDEX

1. Introduction

1.1 Viruses

Viruses are unique in that they are neither living nor dead. Much of a viruses "life" actually consists of metabolic inactivity. Whilst out-with a host they are deficient in any endogenous energy producing mechanisms to support energy dependent processes associated with life. To become active they must hijack a cell and its metabolic pathways to awaken from a dormant state.

Initially before the advent of the electron microscope, it was thought that viruses were able to pass through ultrafine cell surface filters. We now know that viral entry to a cell is much more complex and viruses adopt a variety of strategies to facilitate entry into a cell. Strategies include specific receptor recognition as in HIV where Gp120 on HIV env and CD4 receptors on T-cells interact and which allow the fusion between the cell membrane and the virus envelope. The Adenovirus like many other viruses makes contact with a cell surface receptor via the capsid protein - fibre (protein IV) (Defer *et al.*, 1990). Once inside the cell the protective coat proteins break apart and the genetic material (double stranded DNA in the case of Adenovirus) is released (for a review see Greber *et al.*, 1993).

1.2 Adenovirus.

The Adenovirus was first characterised by Rowe *et al* (1953) as the causative agent of acute respiratory disease. Rowe excised adenoid tissue from children in an attempt to identify different tissue types to support polio virus infection. After a period of 4 weeks incubation progressive tissue degeneration was observed as a result of a previously unidentified virus already in the adenoid cells (latent infection). Later observations showed that secretions from acute respiratory infection in young military recruits induced cytopathic changes in

human cells (Hilleman and Werner, 1954). These observations were attributed to what we now know as the Adenovirus.

1.3 Classification

The *Adenoviridae* family consists over 100 serotypes of Adenoviruses belonging to two genera - *Aviadenovirus* and *Mastadenovirus*. The two genera are identified via the presence of a genus-specific antigen associated with their ability to cross-react immunologically. *Mastadenovirus* group consists of bovine, caprine, equine, human, murine, ovine, porcine and simian viruses, and *Aviadenovirus* those infecting duck, fowl, goose, pheasant and turkey.

Further sub-divisions of viruses are characterised by neutralization studies using serotype specific antisera. The antisera neutralizing one serotype should not neutralize those viruses of other serotypes.

Most Adenoviruses haemagglutinate monkey or rat cells and this characteristic has been used to subgroup human Adenoviruses. The subgroups also share clinical symptoms and the ability to induce tumours in hamsters (Russell, 1995).

Then in 1997 (Harrach *et al.*, 1997) a third genus in the *Adenoviridae* was identified and termed the *Atadenovirus*. This genus was described as an intermediate between the *Aviadenovirus* and *Mastadenovirus* with a very high AT content. This third genus was found to cluster together independently of their host species e.g. BAV-4 and 7 infecting cattle, O287 infecting sheep and EDS infecting fowl resulting in Egg Drop Syndrome.

1.4 Viral Infection

Viral infection results in two possible cellular fates being observed. The viral DNA may partially or completely integrate into the host genome lying dormant; or the cells may play host to a lytic infection, whereby the virus

immediately subverts the cellular transcription/translation machinery to produce thousands of progeny.

Integration of the viral genome has been observed to induce malignant transformation in new born rodent cells infected with Adenovirus subgroups A and B (Horwitz, 1990 a&b), and sparked a great deal of research into identifying whether certain serotypes were capable of oncogenic transformations in humans. No documented case to date has suggested that Adenovirus infection is a causative agent of human tumours. However, oncogenicity is thought to be a result of constitutive E1a gene expression. E1a products interact with retinoblastoma susceptible gene product (pRB) inhibiting the tumour suppresser function of pRB. E1b (52 kD and 19 kD) expression products interact with p53 anti-oncogene protein and the cytoskeleton respectively and are generally thought not to be sufficient to transform a cell (Boulanger and Blair, 1991).

Lytic infections follow a complicated pathway. Suffice it to say here that the viral genome is divided temporally into early, immediate early and late genes. Early genes become active 0.5-7 hpi, broadly producing transcription activators to regulate viral genes and proteins to switch off non-essential cellular genes. During virion synthesis normal cellular activities are down regulated to facilitate efficient translation and transcription of the viral genome and conserve energy to translocate late gene product across the nuclear membrane into the nucleus where they are processed to form mature virions.

1.5 Virion Structure

Adenoviruses are non-enveloped, icosahedral in geometry and contain a double stranded genome of between 34000-36000 base pairs (Akusjarvi *et al.*, 1984). A variety of biochemical and biophysical techniques have been applied

to elucidate the virion structure. Initial structural information came from purified virus extracts separated using SDS-polyacrylamide gel electrophoresis which revealed twelve bands corresponding to distinct polypeptides (Anderson *et al.*, 1973). Further analysis suggested that the virus had an external capsid component and an internal core structure composed of viral DNA and a number of proteins associated with DNA packaging and protection (fig 1.01).

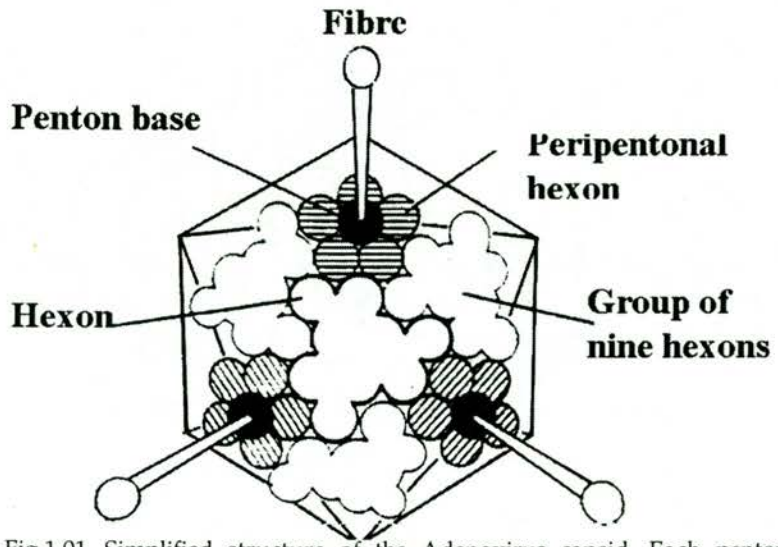


Fig.1.01. Simplified structure of the Adenovirus capsid. Each penton base is surrounded by 5 hexons (peripentonal hexons). The group of nine hexons are held together with the "cement" protein IX. (van Oostrum *et al.*, 1987)

1.5.1 Capsid/Structural proteins.

Initially three distinct groups of protein were identified as major capsomeric units, namely hexon, penton base and fibre projection (Norrby 1969, Horwitz *et al.*, 1969 and Prage *et al.*, 1970). Since then a number of other polypeptides have been associated with the capsid: VI; VIII; IX and IIIa.

The major capsid constituent is hexon protein accounting for 240 of the 252 capsomeric units. Hexon monomer (II) is composed of 967 amino acids with an apparent molecular weight of 108 kD. Hexon is perhaps the best characterised of the viral structural proteins, crystallisation of hexon and X-

ray diffraction work revealed two very similar 8 stranded β-barrel domains (P1 and P2) in each monomer and four loops termed l_1 through to l_4 which project away from the β-barrel. Each hexon capsomere is a trimer of three identical monomers associating via the P1 and P2 β-barrel domains. Each monomer's loops 1, 2 and 4 intertwine to form three towers which give the characteristic triangular top. The pseudo-hexagonal base of the hexon allows close packaging of the surface molecules fulfilling the role of protection. The triangular top on the surface of the hexon contains most of the antigenic determinants (van Oostrum, 1985; Roberts *et al.*, 1986). This structure is depicted in figure 1.02.

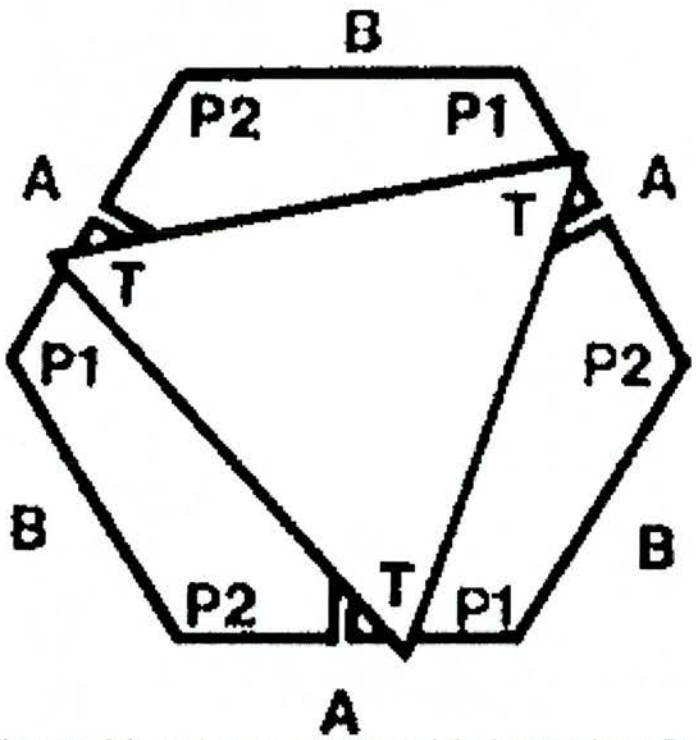


Fig. 1.02. Schematic representation of the hexon trimer. P1 and P2 (8 beta barrels) interact to form a stable trimer. The superimposed triangle represents the tower region formed by loops 1,2 and 4. (Adapted from Roberts *et al.*, 1986)

Under mild dissociation conditions such as addition of 0.5% sodium deoxycholate (Russell *et al.*, 1971) hexons are released as planar groups of nine (GON). The nine hexons appear to be held together by polypeptide IX on the outer surface of the virion with a stoichiometery of twelve per facet. Deletion

mutants of polypeptide IX produce heat labile virions indicating that IX was not necessary for assembly, rather acting like cement (Furcinitti *et al.*, 1989; Colby and Shenk, 1981). This accounts for 180 of the 240 hexons units. The remaining 60 hexons are termed peripentonal hexons and are found at the margins as links between the facets and pentons. Figure 1.03 is a schematic representation of one facet. Each point of the triangle is a penton base surrounded by 5 peripentonal hexons. The two slightly darker hexons trimers at the boundary of each facet represent the overlap region with other facets.

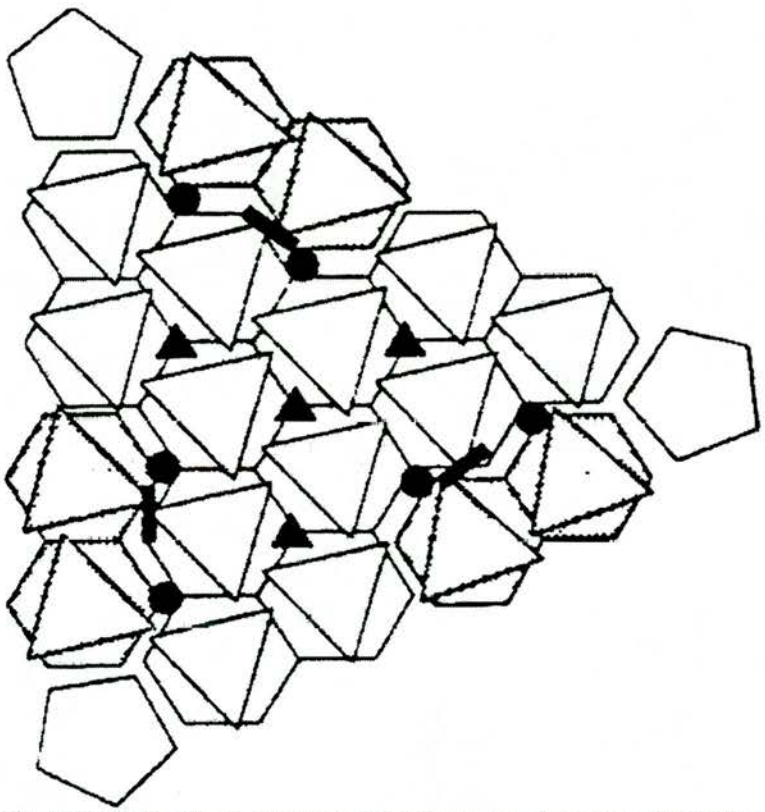


Fig. 1.03. Virion Facet. Schematic representation of one facet from the Adenovirus capsid. Hexons are represented by hexagonal bases with a triangular top super imposed, pentons as pentagons. The structures integrity is maintained by three polypeptides, polypeptide VIII (rectangle), polypeptide IX (triangle) and polypeptide VI(circle) (figure adapted from Stewart *et al.*, 1991).

1.5.2 Penton

Penton is a complex oligomeric protein consisting of a long fibrous projection terminating in a knob linked non-covalently to penton base. The penton is situated at each of the 12 vertices of the 20 faced icosahedral virion (Boudin *et*

al., 1979; Devaux *et al.*, 1990; Valentine and Pereira, 1965; van Oostrum and Burnett, 1985).

The penton base is composed of 5 subunits of polypeptide III (molecular weight 63 kD) and the name is derived from the fact that pentons are surrounded by hexons on five sides. The penton-hexon contact is in sharp contrast to the tight hexon-hexon interaction, characterised by narrow connections. The penton base has few antigenic sites due to concealment by the fibre (Stewart and Burnett, 1991).

Fibre has molecular weight of approximately 62 kD. The length of the fibre projection is characteristic of Adenovirus subtypes. In Adenovirus 2 the fibre is 581 amino acids long with glycosolations and repeating units of 15 amino acid residues between residues 43 and 400. Antibodies to the fibre prevent entry of virions into red blood cells and the attachment of fibre to HeLa and KB cell lines. It has been suggested that the fibre interacts with CAR (coxsackie-adenovirus-receptor) receptors (Bergelson *et al.*, 1997, 1998; Tomko *et al.*, 2000). Using a competitive binding assays the head domain of the fibre was determined to attach to the cell surface of the HeLa cell membrane (Louis *et al.*, 1994). The attachment marks the beginning of the uncoating process, virion identified from within cells lack fibre and peripentonal proteins (Greber *et al.*, 1993).

1.5.3 Minor capsid polypeptides

Polypeptide IIIa is released with mild treatments which exclusively dissociated vertex components of the capsid. IIIa is located between peripentonal hexons and GON-hexons. This idea was challenged by van Oostrum and Burnett (1985) who proposed that IIIa was an anchor between vertices and viral core, probably interacting with polypeptide VIII (Anderson *et al.*, 1973; Everitt *et al.*, 1973 and Everitt *et al.*, 1975).

pVI is processed by the Ad2 protease to VI via N and C-terminal cleavage sites (Matthews and Russell, 1994 and 1995). Successive cleavages of pVI improved VI-hexon binding capabilities. Two important regions of amino acid residues (48-74 and 233-239) are highly conserved throughout Mastadenovirus pVI proteins. These regions act in *cis* to facilitate strong association between VI and hexon. Polypeptide VI is thought to be on the internal surface of the capsid (hence not labelled by enzymatic iodination (Everitt, 1975)) and in contact with the DNA as a dimer of trimers (van Oostrum and Burnett, 1985, Philipson, 1984).

Polypeptide VIII (15 kD), is found as a monomer, in association with hexon (Everitt *et al.*, 1973). Three potential protease cleavage sites have been identified within pVIII (Webster *et al*, 1989 a and b). Hong and Boulanger (1995) reported VIII to associate closely with penton base and fibre at the vertices of the virion. A pVIII temperature sensitive mutant H5sub304 (Liu *et al*, 1985) produces thermolabile virions not unlike those of IX suggesting a possible function as a scaffolding protein. The exact function of VIII is not know but Stewart *et al* (1993) believe it is located on the capsid surface.

Protein IX with molecular weight of 15 kD is 139 amino acids long and does not undergo proteolytic maturation (cf. V, VII, etc). Temperature sensitive mutants produce thermolabile virions, although not in the classical sense. IX is responsible for stabilising the group of nine hexons.

1.5.4 Adenovirus Core Proteins.

The nucleoprotein core of Adenoviruses serotype 2 contains a linear double stranded DNA genome of approximately 36 kb, (Akusjarvi *et al.*, 1984)) complexed with viral encoded proteins. Each 5′ end is covalently attached via 5′ phosphodiester bonds to a 55 kD protein, the terminal protein (Smart and Stillman 1982).

Unlike the papoviruses Adenoviruses encode their own basic core proteins. Polypeptides V, VII and X are highly basic molecules with molecular weights 41.6 kD and 19.4 kD and 11 kD respectively. These polypeptides, are produced as precursors and undergo Adenovirus protease mediated cleavage maturation (with the exception of V). In fact primary studies into the protease's kinetic profile was carried out using pVII as substrate (Bhatti and Weber 1979 and Webster *et al.*, 1993) before moving to use synthetic peptides. Polypeptide VII is the most tightly DNA associated core protein (Brown *et al.*, 1975) and is thought to be a viral version of histone proteins. Polypeptide V is also capable of binding DNA although it has also been shown to associate with penton base (Everitt, 1975). V may act as a link between capsid and core components to stabilise the virion.

Polypeptide X (pMu) with molecular weight of 11k is 79 amino acids in length and arginine rich (21%) (Weber and Anderson, 1988). pMu contains two protease cleavage sites, cleavage at residue 31 relinquished the carboxy termini which is postulated to be the X (mu) component. X is observed in the chromatin core interacting with the viral DNA structure (Chatterjee *et al.*, 1986). These proteins are responsible for the packaging DNA in the virion in a histone like manner.

1.6 Viral Life Cycle.

Viral life cycle is conventionally divided into two phases separated by the onset of viral DNA replication. The first stage comprises viral attachment to cell surface receptors, penetration, the transcription and translation of early gene products which mediate viral gene expression, DNA replication and evasion of host cell defences such as apoptosis and cell cycle modulation. The genetic organisation of Adenovirus type 2 is depicted in figure 1.04.

Early genes are conventionally denoted E and are expressed almost exclusively until 5-6 hpi at a multiplicity of infection of 10 pfu per cell. DNA replication is first detected 6 hours post infection (hpi) in cultured monolayers and this marks the progression to late events. However, early genes expression is not restricted exclusively to this time frame. It has been observed that early products expressed later in infection tend to be smaller versions of those seen early in infection (Shenk, 1996).

Late gene expression is responsible for the accumulation of mainly structural proteins associated with the assembly of viral progeny and preparation for their release from the cell. Somewhere in the region of 10^4 virion particles are produced per cell per infection (Greene and Daesch, 1961). Because of the viral proteins produced to regulate cell cycle activity, cells infected do not usually divide. As a result of this the DNA and protein content of infected cells tends to be twice that of normal uninfected cells.

In between early and late gene expression is the expression of intermediate early genes IVa2 and IX.

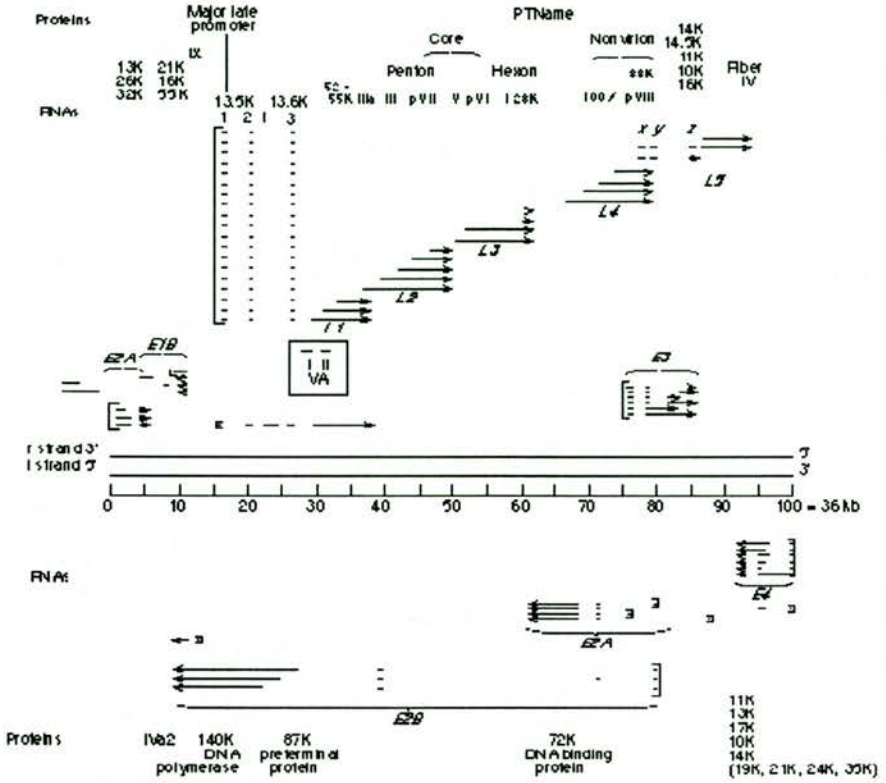


Fig 1.04 A Schematic Representation of the Genetic Organisation of the Adenovirus Genome. E1a and E1b are translated first and initially control the expression of host and viral cell proteins. Late genes are all transcribed from the MLP at approximately 15 map units. Late products are mainly viral capsid proteins and the viral encoded 23kD protease responsible for the capsid maturation. A summary of cellular and viral protein targets of the protease in shown in table 1 (Shenk, 1996).

1.6.1 Cell Attachment and Adsorption

Adenovirus attachment is facilitated by the C-terminal distal knob of fibre interacting with a cell surface receptor (Lonberg-Holm and Philipson, 1969; Hennache and Boulanger, 1977). It has been suggested that the fibre interacts with CAR like receptors (Bergelson *et al.*, 1997, 1998; Tomko *et al.*, 2000), although other receptors may be involved in the viruses entry to a cell as the virus does show a degree of specificity for certain cell types.

Adsorption is facilitated through penton interaction with intergrins αvβ3 and αvβ5 (Wickham *et al.*, 1993; Chiu *et al.*, 2000). Penton base alone is not sufficient to facilitate entry to a cell because antibodies to fibre and the addition of soluble fibre to cultured cells block the adsorption.

Once formed, the Adenovirus-Receptor complex diffuses into clathrin coated pits (Chardonnet and Dales, 1970 a&b; Fitzgerald *et al.*, 1983; Vagra, 1991), approximately 80-85% of bound virus enters the cell (Greber *et al.*, 1993) and virus appears in endosomes after 10 minutes. An interaction between penton base and intergrin is thought to be responsible for the release of virus into the cytosol after approximately 5 minutes (Blumenthal, 1986; Seth, 1984; Svensson, 1985).

Transportation to the nucleus is thought to occur via microtubles (Dales and Chardonnet, 1973) the interaction is thought to be mediated by hexon. Adenovirus appears at the nuclear pore after 40 min, after 120 min DNA is found free from hexon protein. The mechanism responsible for nuclear localisation of DNA is as yet uncharacterised.

The uncoating process starts with the loss of the vertex protein fibre (IV) and penton (III pentamer). Fibre and IIIa are lost more readily than III. VI and VIII are thought to bridge the DNA core to the capsid and their dissociation and degradation permit the eventual release of DNA. The final step is the loss of IX and at this stage the DNA is free of hexons.

It was shown that the released DNA associates with the nuclear matrix through its terminal protein (Bodnar *et al.*, 1989; Fredman and Engler, 1993; Schaack *et al.*, 1990). Mutants of TP which fail to associate tightly with the nuclear matrix do not initiate efficient transcription.

1.6.2 Early Gene Activation

Early gene expression has three main functions, firstly to prepare the host cell to enter s-phase and thus synthesise DNA ; secondly to protect the host cell from antiviral measures; and thirdly to produce the viral gene products necessary for the replication of viral DNA.

Once the viral chromosome reaches the nucleus the E1a unit is the first to be transcribed under the control of a constitutively active promoter (Hearing and Shenk, 1983. 1986). In the first stages only two forms of mRNA are expressed, E1a 12s and 13s. These products are a result of differential splicing and produce proteins identical up to the end of the CR2 domain. The primary E1a translation products are extensively phosphorylated (Harlow 1985), however the function of the phosphorylation remains unknown.

E1a products are composed of three conserved regions, CR1-3, separated by less highly conserved regions (Dyson *et al.*. 1992; Moran and Mathews, 1987; Shenk and Flint, 1991). The conserved regions do not directly interact with DNA (Chatterjee *et al.*, 1988; Ferguson *et al.*, 1985), interaction is mediated by association with other proteins to control gene expression.

E1a is able to activate expression of other Adenovirus transcription units by increasing the rate of transcription of viral genes (Nevins, 1981). E1a proteins can act as *trans* activators through their ability to interact with auxiliary factors which mediate the basal level of transcription, with activating proteins that bind up and down stream promoters and enhancer elements, and with regulatory subunits that influence the activity of DNA binding factors.

E1a induces the expression of VA RNA genes (Hoeffler *et al.*, 1988; Hoeffler and Roeder, 1985; Yoshinaga *et al.*, 1986) which are transcribed by RNA polymerase III. VA genes are responsible for antagonising interferon response to prevent the cell reaching an antiviral state.

As infection progresses the rate of transcription of early genes is reduced partly due to cell death but also in a controlled manner by the virus. E1a products repress a multitude of enhancers, one of which is its own enhancer (Borrelli *et al.*, 1984; Velcch and Ziff, 1985). E2 DNA binding protein acts to inhibit expression from the E4 promoter by an as yet uncharacterised

mechanism (Handa *et al.*, 1983; Nevins and Winkler, 1980). E1a activation of the E4 gene and the subsequent accumulation of E4-14kD protein antagonises the AP1 transcription factors induction by E1a.

Early gene products are translated on polysomes along with cellular mRNA's. At a later stage the viral strategy changes to promote preferential viral mRNA translation. E1a products are not the only early gene products to activate transcription. E4 17 kD and E2 DNA binding protein also facilitate activation.

1.6.3 Prevention of Apoptosis and Activation of Host Cell

In order to induce a cellular environment favourable to viral DNA synthesis the Adenovirus expresses E1a and E1b proteins which interfere with normal cell cycle control. E1a is capable of stimulating quiescent cells into s-phase of cell cycle.

E1a interacts with the Retinoblastoma tumour supressor protein, pRB, which regulates the transcription factor E2F (Whyte *et al.*, 1988). E2F binding sites are present in E1a and E2 promoters so when E1a competes E2F out of its interaction with pRB it stimulates its own transcription. E1a is also able to interact with other members of the pRB family such as p107 and p130. The interaction with p107 is mediated through the CR1 domain, where as p300 is mediated via half the CR1 domain and the poorly conserved amino-terminal Wong and Ziff 1994). The interaction with p300 inhibits cell differentiation by preventing the p300 bridging enhancer and initiation sites. pRB and p300 are two ways of dealing with the same event and thus allow successful course of infection in a variety of cell lines.

E1a is also capable of down regulating LMP2 which is responsible for the surface presentation of antigens. This is facilitated through E1a competing with IRF-1 to bind Stat1 and thus preventing the activation of transcription.

E1b products target other proteins important in cell cycle control, namely the p53 tumour supressor protein, which regulates the G_1S progression of the cell cycle. High levels of p53 are associated with DNA damage and prevent the cell progressing until repair mechanisms have been successfully initiated. Continued high level expression of p53 activates WAF1 and GADD transcription and thus apoptosis. E1b-55kD can bind to p53 and block its transcriptional activation function (Sarnow *et al.*, 1982a). It has also been shown that p53 can be sequestered outside the nucleus by Ad 5 E1b (Zantema *et al.*, 1985).

E1a 's ability to increase DNA synthesis also has the side effect of increasing levels of p53 expression and its stability in the nucleus. Two E1b 's can block the apotosis induced by E1a, E1b 55kD and E1b 19kD. E1b 55kD is able to bind p53 modify its function. It is not exactly clear how E1b 19 kD modifies p53's function although it is believed that E1b is able to block p53 mediated transcription(Shen and Shenk, 1994).

1.6.4 Viral DNA Replication

With an infection multiplicity of 10 pfu per cell DNA replication in the closely related Adenovirus type 2 and 5 begins after 5 hours. The initiation of replication occurs as the three E2 products which are required for DNA replication, accumulate (Hay and Russell, 1989; Challberg and Kelly, 1989; Stillman, 1989).

DNA replication occurs in two stages with the inverted terminal repeats acting as the origin of replication (fig 1.05). The first stage of synthesis can occur at either end of the DNA molecule, but only one strand acts as a template, and parent and synthesised daughter strand pair up. The second strand is then synthesised. The displaced parent strands circularises and

anneals through the inverted terminal repeats and is recognised by the same replication machinery as the first stage.

Replication is regulated by 3 *cis*-acting elements located within the inverted terminal repeats which occupy the first 51 bases of the chromosome. The sequences are termed domain A, B and C. Bases 1-18 represent domain A which codes for the minimal origin of replication, cellular *ORP-A* binds to the first 12 bases (Rosenfeld *et al.*, 1987). 9-18 is the binding site for two E2 products, preterminal protein (pTP) and the viral DNA polymerase (Challberg and Rawlins, 1984; Lally *et al.*, 1984; Tamanoi and Stillman, 1983). Preterminal protein is covalently attached to the 5′ end of the DNA molecule and has a protease cleavage site which is hydrolysed at the latter stages of infection. Terminal protein remains attached to the DNA (Temperley and Hay, 1992) after hydrolysis. Figure 1.06 illustrates the proposed interactions of the proteins involved in the initiation of DNA replication. The viral polymerase is capable of 5′-3′ polymerase activity and 3′-5′ exonuclease activity. 3′-5′ exonuclease activity is believed to function to proof read the replication process.

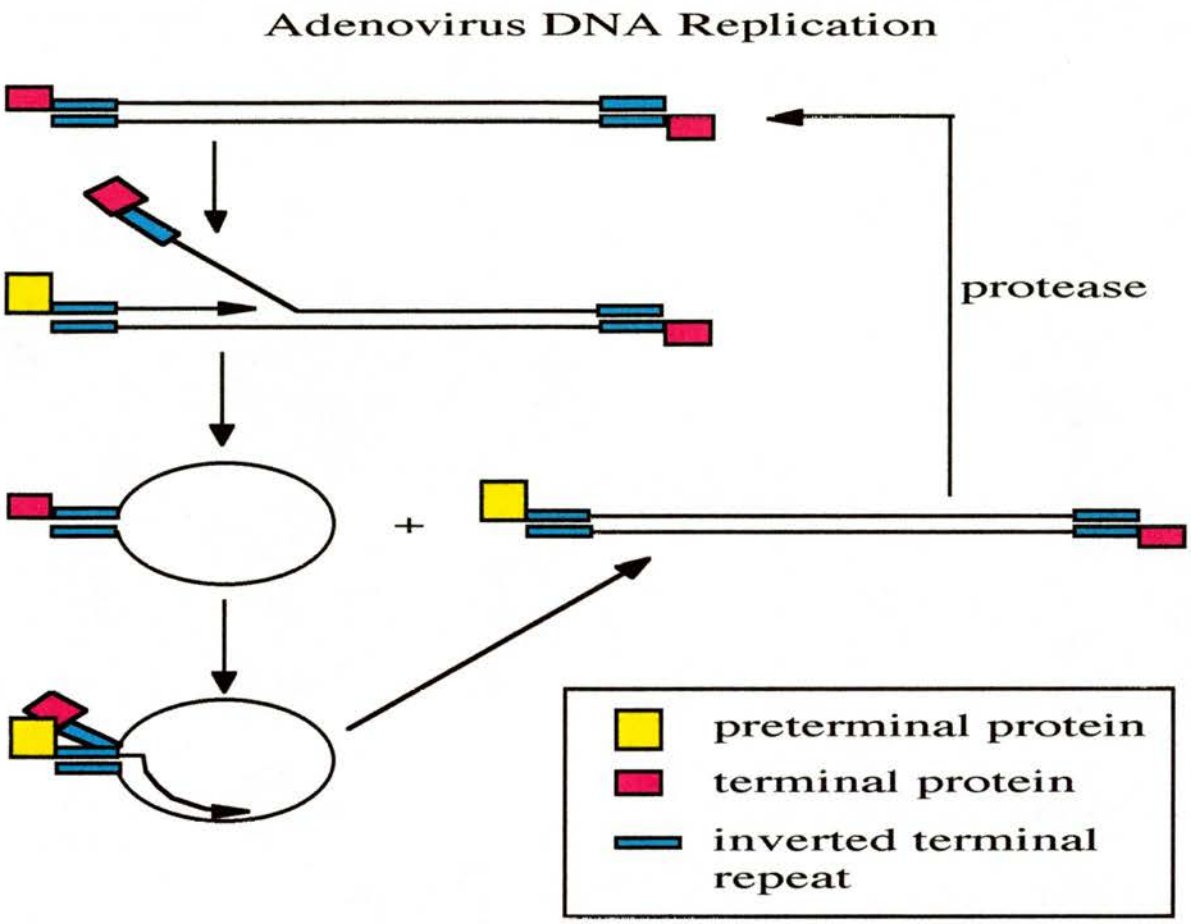


Fig 1.05 Adenovirus DNA Replication. Replication is initiated by the terminal protein acting as the primer for replication.. Elongation requires the E2 coded polymerase and ssDNA binding protein. A second round of replication then occurs using the displaced stand as a template (adaptation from Ron Hay).

Domain B which occupies bases 19-31 and domain C (40-51) are not absolute requirements for successful DNA replication, but serve to enhance the efficiency of the initiation. Two nuclear factors (transcription factors) bind these domains NFI and NFIII respectively. NFI interacts with the preterminal protein-polymerase complex stabilising it at the origin of replication (Mul *et al.*, 1990; Mul and van der Vliet, 1992). E2 single stranded DNA binding protein can also activate the binding of NF1.

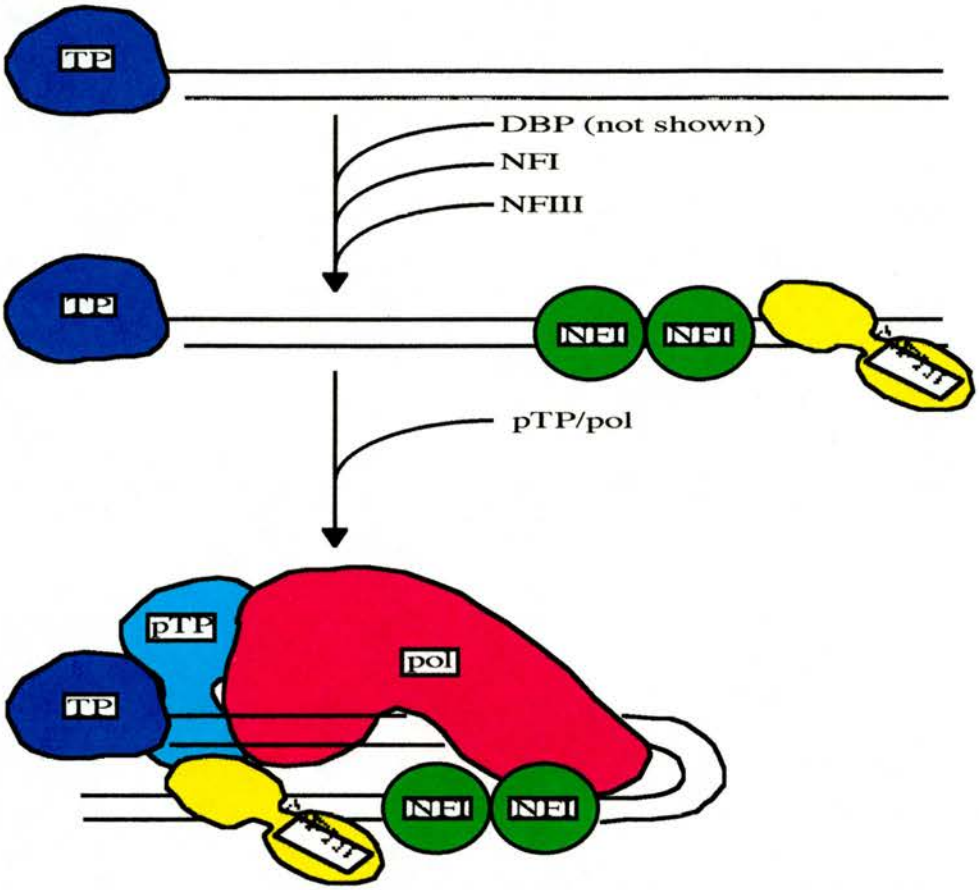


Fig 1.06 Factors Involved with Initiation of DNA Synthesis in Adenovirus Replication. DNA replication is initiated by the binding of the Adenovirus DNA polymerase pTP heterodimer to TP which acts as a primer at the origin of replication. Initiation of elongation is enhanced by NF I, NFIII and DBP (Bosher *et al* ., 1990; Webster *et al.*, 1997).

The mechanism by which NFIII (Oct1) activates initiation is less well characterised. Its thought to be important in bending the DNA molecule perhaps leading to a more favourable conformation for replication to begin (Raychaudhuri *et al.*, 1990).

Preterminal protein acts as a primer for DNA replication and preserves the integrity of the viral chromosome terminus during multiple rounds of replication. The interactions is maintained through an ester bond between βOH of serine and dCMP of the viral chromosome. Bond formation is thought to occur only after pTP-polymerase complex has bound the origin of replication. DNA elongation requires the polymerase, ssDNA binding protein

and cellular factor NFII. E4 transcription products are also thought to be required for efficient DNA replication (Bridge *et al.*, 1993; Halbert *et al.*, 1985). In the absence of E4 ORF 3 and 6, ORF4 is thought to be able to modulate expression through its affect on protein phosphatase 2A, and consequently AP1 transcriptional regulation of E2 expression (Medghalchi *et al.*, 1997).

1.6.5 Activation of Late Genes

Efficient expression of late genes only occurs once the onset of viral DNA replication has occurred. The product of transcription is a single 29 000 nucleotide RNA unit (Evans *et al.*, 1977) which is processed by differential poly A sites and differential splicing (Nevins and Darnell, 1978). 18 distinct mRNA have been identified covering 5 families of products, these are termed L1 to L5 and are based upon common Poly A sites. All products are under the control of the major late promoter (MLP). There are two factors involved in transcription, a *cis* acting change in the chromosome and induction of at least one new virus coded *trans* acting factor.

The time-dependent cis-acting modification of the Adenovirus genome was illustrated by the fact the early products from one virus could not switch on late transcription of another virus causing an infection in the same cell (Thomas and Mathews, 1980). One reason for this may be a change in the constituents of the viral chromatin during the process of viral DNA replication as unreplicated chromatin appears to associate with different proteins to those of replicated chromatin (Chatterjee *et al.*, 1986). The most probable reason for this is that DNA replication opens up the DNA molecule allowing transcription factors access to promoters with the displacement of histone like proteins (pVII).

A second explanation could be due to the delay of the second viruses core's arrival and the decondensing of DNA. Also chromosomal domains could

become available in a defined order, the MLP being the last in the order possibly due to the processing of transcription machinery (Toth *et al.*, 1992). Finally viral chromosomes may establish compartmentalised environments in the nucleus which may take a period of time to recruit necessary cell factors for the expression of late genes. Discrete viral centres have been observed in infected nuclei using electron microscopy. DNA foot printing has identified a transcription factor, either USF or MLTF present and bound to DNA after DNA replication and not before (Carthew *et al.*, 1985). USF activates transcription but if a point mutation is introduced its no longer able to bind and have the same transcription rate. So in summary it is likely that the *cis* acting modifications to the viral chromosome involves changes to allow the access of transcription factors to the template.

The second component that contributes to activation of the major late promoter is a virus encoded transcription factor. In addition to the upstream region containing the USF/MLTF, CAAT and TATA binding motifs contribute to regulation along with the regions at +85 and +120 downstream of the of the major late promoter (MLP) (Mansour *et al.*, 1986; Leong *et al.*, 1990). These downstream domains are able to bind two factors which can co-operate with USF and MLTF leading to optimal transcription activation.

IVa2 is a delayed early gene that has a repressor site. The theory is that early gene products switch on IVa2 by relieving its repression. IVa2 is then able to bind downstream of the MLP and contribute to its induction. This is a sequential event which therefore would contribute to timing which together with the DNA replication control mechanism would be a very efficient way to co-ordinate mRNA production.

The accumulation of viral late genes mRNA in the cytoplasm is promoted by E1b 55kD and E4 34kD at the expense of the accumulation of cellular mRNA.

E1b 55 and E4-34kD form a complex (Sarnow *et al.*, 1984) localised within and surrounding intracellular inclusions believed to be viral transcription and replication centres. It is believed that the E1b E4 complex relocalises a cellular factor for nuclear transport of mRNA from the site of synthesis to the nuclear pore (Ornelles and Shenk, 1991). This explains why there is a block on the cellular mRNA export and the activation of viral expression as transport factors are removed from cellular sites of transcription and processing to viral sites.

Viral mRNAs in addition to being preferentially transported are also preferentially translated (Babich *et al.*, 1983). Host cell mRNA is not degraded, during the above process, β-tubulin is the only cellular mRNA to be exported from the nucleus, however, the mRNA is not translated.

Selective translation of viral mRNA occurs in a number of ways. Firstly cellular protein kinase r (PKR) is activated by double stranded RNA which accumulates during Adenovirus infection (Maran and Mathews, 1988; O'Malley *et al.*, 1986). PKR phosphorylates eIF-2a which inactivates translation of host cell mRNA. If cells are deficient in PKR, host cell translation shut off does not occur. VA RNA of Adenovirus has been shown to be an agonist of PKR, and co-purifies with viral mRNA's (Mathews, 1980). If they reside in the same locality as the viral mRNA in the cytoplasmic compartments they will inactivate PKR and allow eIF2a to translate viral mRNA i.e. functional compartmentalisation.

eIF-4F inactivation contributes to the selective translation of Adenovirus infected cells (Huang and Schneider, 1991). eIF-4F binds cap domains of mRNA to facilitate scanning by the 40s ribosome to AUG sites. eIF-4F is activated by phosphorylation and late during infection is dephosphorylated about the time that cellular mRNA translation ceases. The dephosphorylation

results in a loss of eIF-4F activity. The tripartite leader sequence is thought to negate the need for 40s helicase activity and allow translation of viral mRNA's after host cell translation shut off (Huang and Schneider, 1991). Cellular proteins tend to have more secondary structure at 5' end which the 40s helicase activity is needed to scan through to find the start codon.

Finally L4-100kD protein encoded for in the L4 region of the MLP selectively activates late protein synthesis. Viruses expressing defective L4 100 kD protein fail to efficiently translate late viral mRNA's, but are able to block host cell translation. Its thought that since L4 100kD can bind to mRNA it may function at the polysome to facilitate viral translation (Adam and Dreyfuss, 1987).

1.6.6 Virus Assembly and Release from Host Cell

After DNA replication and translation of a large amount of Adenovirus structural proteins initiation of capsid assembly begins. Hexon trimers rapidly assemble and L4 100kD is thought to act as a scaffold via an as yet uncharacterised process. Mutant viruses defective in L4 can block assembly of capsomeres (Roberts *et al.*, 1984; Leibowitz and Horwitz, 1975). Penton capsomeres form the pentomeric base and trimeric fibre filaments assemble (Horwitz *et al.*, 1969). These processes all occur in the cytoplasm, but once proteins are in the nucleus virion assembly occurs. Understandably mutation of any of these building blocks, scaffold proteins or the enzyme responsible for the processing of precursor proteins, inevitably leads to incomplete virion assembly or in the case of the viral protease, a non infectious particle.

Combining studies of kinetics and mutants has shed light on the process of virion assembly. Assembly begins with the formation of an empty capsid (Philipson, 1984; Sundquist *et al.*, 1973) and then the entry of the viral DNA molecule. The DNA capsid interaction is mediated by *cis* acting elements

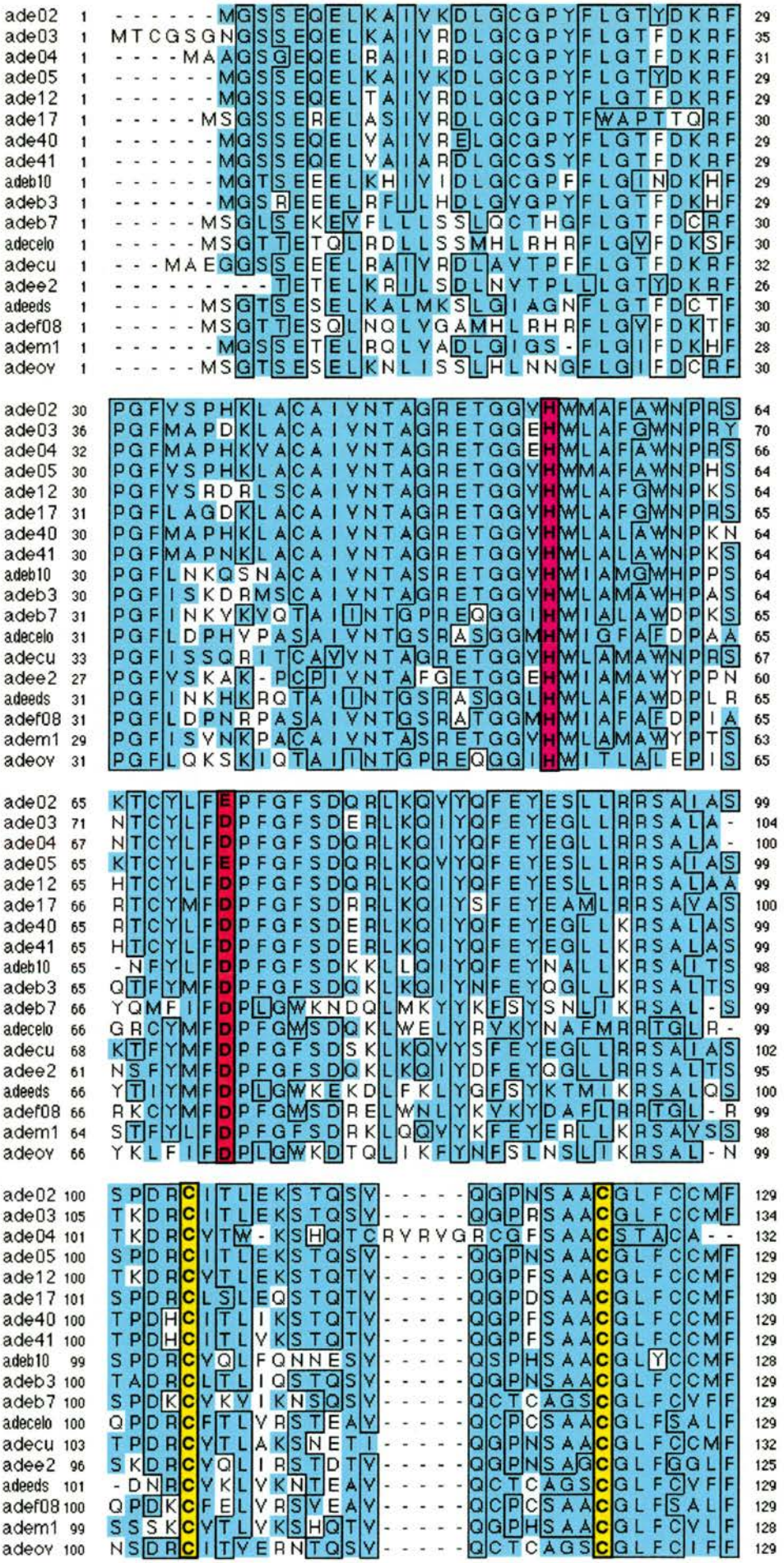
found 260 base pairs down from the left end of the viral chromosome (Grable and Hearing, 1992; Hammarskjold and Winberg, 1980; Hearing *et al.*, 1987; Tibbetts, 1977). Encapsidation starts at the left end of the chromosome and proceeds to the right. The proteins interacting with the packaging sequence remains as yet unidentified. The mechanism of the entry of viral DNA into the capsid is somewhat disputed, some groups believe that viral DNA replication and encapsidation are linked (Weber *et al.*, 1985), others protest that replication and encapsidation occur in separate nuclear compartments (Hasson *et al.*, 1992). L1 52/55kD protein is a combination of 2 differentially phosphorylated proteins of 48 kD, and facilitates encapsidation by acting as a scaffold protein (Hasson *et al.*, 1992; Hasson *et al.*, 1989). Once assembly of the virion particle is complete, for the particle to become infectious the virally coded protease must cleave at least 4 virion constituents (pVI, pVII, pVIII and pTP).

After assembly is complete virus particles are then ready to escape from the cell and spread infection. The first important step in viral escape is the cleavage of vimentin independent of viral DNA expression (Beltz and Flint, 1979) during the adsorption process vimentin degradation is thought to be in response to E1b 19 kD (White and Cipriani, 1989 & 1990) occurring within 30 minutes of viral infection.

Secondly the viral coded protease has been shown to cleave cytokeratin K 18 (Chen *et al.*, 1993) creating headless proteins incapable of maintaining the structural integrity of the host cell. This loss of integrity makes the cell more susceptible to lysis, thus facilitating the release of the viral progeny.

1.7 Adenovirus Protease

The Adenovirus protease is unique in the fact that unlike most other proteolytic enzymes it requires the addition of an eleven amino-acid peptide



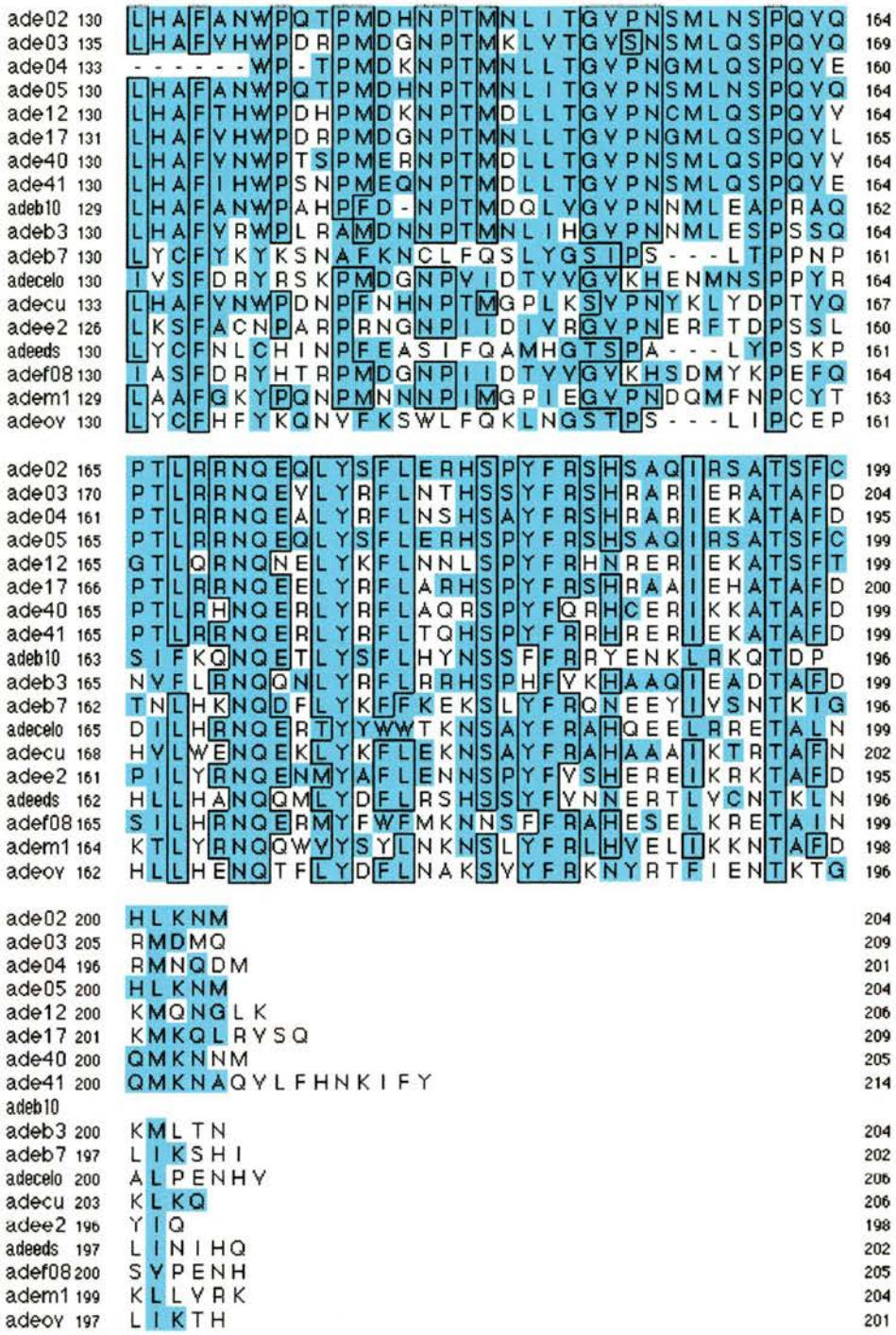


Fig 1.07 Alignment of Multiple Adenovirus Protease Sequences. Highlighted in yellow are the conserved cysteine residues responsible for the interaction with activating peptide (C104), and the catalytic mechanism (C122). Histidine 54 (magenta) which acts as a proton donor and acceptor during cleavage and aspartate/glutamate 71 (red) are also indicated. Cyan denotes homologous amino acid substitutions (identified using GUS method SeqVu version 1.01) and boxed regions represent conserved amino acid residues. The numbering system refers to the positions of residues as compared to Adenovirus type 2. ade denotes Adenovirus, suffix of b denotes bovine; e, equine; c, canine; f, fowl; m, murine; ov ovine and celo CELO. The final number refers to the serotype.

(GVQSLKRRRCF) rather than proteolytic processing to achieve maximum activity. The primary sequence of proteases sequenced to date also show a high degree of homology, there are 4 highly conserved residues illustrated in fig 1.07, namely H54, E/D 71, C104 and C122.

The protease is thought to be active at three times during the viral life cycle; first during the delivery of the genome to the cell (Cotten and Weber 1995); then at 14 hpi onwards to process pro-proteins of viral structural and non structural proteins (Anderson, 1973, Webster *et al* 1989); and finally at the end of infection (36 hpi) to cleave cellular proteins namely cytokeratin K18 and K7 to facilitate host cell lysis (Chen *et al* 1993). It is the spatial and temporal control of the protease which is of great interest in the search for an anti-viral agent. Premature activation of the protease or prevention of the delivery of the protease to the nucleus (Rancourt *et al.*, 1995) would produce non-viable progeny. Table 1.01 is a summary of the precursor polypeptides and proteins which are cleaved by the protease during the course of viral infection.

1.7.1 Characterisation of the Protease to Date

Proteolytic processing of Adenovirus proteins was first described by Anderson (1973). The processing appeared to be a prerequisite for the production of mature and infectious virions. Although the need for proteolytic maturation was evident, the origin of the protease was unknown. Use of Ad2ts1 virus defective in precursor polypeptide processing at the non permissive temperature confirmed the protease to be viral in origin (Bhatti and Weber, 1978 Weber, 1976 Yeh-Kai *et al*, 1983). In Ad2ts1 infections precursor proteins were found to assemble in the place of mature proteins found in the wild type. The resulting viral particles were shown to be non-infectious. The putative gene of this defect was mapped to co-ordinates 57-69 of the Adenovirus 2 genome (Hassell and Weber 1978).

Mechanistic classification of the protease proved to be very difficult. Inhibition profiles suggested that the protease was a chymotrypisn like, non-metallo, neutral serine protease (Bhatti and Weber, 1979). The same authors also reported that the protease was found in the cytoplasm and nucleus.

The isolation of a temperature sensitive mutant Ad2ts1, defective in protease activity at the nonpermissive temperature, and mapping of the mutation lead to the suggestion that the protease was coded for in the L3 region of the viral genome (Yeh-Kai *et al.*, 1983).

Both Sung *et al* (1983) and Tremblay *et al* (1983) reported that the protease was specific for G-A bonds. Tremblay *et al* (1983) studying the similarities of pVI and pVII proposed that protease was specific for glycine alanine (G-A) bonds (fig 1.08), but further work suggested that alanine glycine bonds were also cleaved. Figure 1.08 shows the cleavage sites in pVI and pVII and how they relate to the cleavage pocket. The authors also noted that a range of small peptides with G-A bonds were cleaved by the protease and concluded that there was a need for secondary structure interaction between the enzyme and substrate. At this time the group concluded that the protease was a serine class protease since activity was abolished by diisopropyl fluorophosphate - a serine protease inhibitor.

S K M F G G ↓A K K R pVI

S N M S G G ↓A F S W pVII

P_6 P_5 P_4 P_3 P_2 P_1 $P_{1'}$ $P_{2'}$ $P_{3'}$ $P_{4'}$

Fig. 1.08 Consensus Cleavage Sites. Shows the consensus cleavage sites in pVI and pVII, ↓ represents cleavage site. The relative positions of the residues around the cleavage site are displayed.

Weber and Anderson (1988) reported that a glycine methionine (G-M) bond was cleaved in the maturation of 11 k protein (mu) whilst Webster *et al* (1989) with digestion assays of octapeptides based on the predicted cleavage sites in

pVII suggested that all cleaved peptides conformed to one of two consensus sequences (M,L)XGX↓G or (M,L)XGG↓X. This was extended by Anderson (1990) to include Isoluecine (I) in the P_4 position. Substrate specificity required that the side chain of the amino acid P_1' point away from the scissile bond. Bulky side chains were unable to occupy P_1' unless the side chain pointed away from the scissile bond such as that of phenylalanine in pVII (Webster *et al*, 1989a). pTP is processed to TP, a 55kD protein, by the Adenovirus protease during infection. However, a synthetic substrate peptide containing the putative cleavage site was not cleaved *in vitro*. This was thought to be because the valine side chain did not attain the correct conformation to permit bond cleavage. These observations suggested that although primary structure considerations were important, secondary and tertiary considerations were equally important in the cleavage of potential target sequences. This was consistent with the observations of Tremblay *et al* and Sung *et al* (both 1983).

Definitive enzyme classification of the protease was still proving problematic. Sequence alignments suggested His 54, Asp 102 and Ser 160 were contenders for the active site triad (Houde and Weber 1990) inferring a serine class protease. However the characteristic GDSGG motive of serine protease was not observed in the protease's amino acid sequence alignments. Webster *et al* (1989), studying a panel of serine and cysteine protease inhibitors determined that the protease belong to cysteine class and as C104 and C122 were the only truly conserved cysteines, that they were the likely candidates for the catalysis reaction.

Serine classification was finally laid to rest when serine 95, 160 and 82 were mutated to alanines and protease retained its activity (Grierson *et al*, 1994; Rancourt *et al.*, 1994). Mutations directed at C104 and C122 reduced activity to

less than 5% of that achieved by the wild type supporting the cysteine centred theory. Further refinement of 12 Adenovirus protease sequences identified C104 and C122 as the only conserved cysteine residues. C104 and C122 were likely to be the active site nucleophile and the site of interaction of the protease with activating peptide.

Chatterjee and Flint (1987) proposed that protease was produced as a zymogen and activity was only achieved after autolytic proteolysis of an A-G bond at amino acids 46-47 and resulted in a 19kD phosphorylated product. However, with the identification of an activating peptide the autoproteolytic processing theory was put to rest. (Mangel *et al.*, 1993 and Webster *et al.*, 1993).

The protease has since been implicated in the release of viral progeny when it was shown to cleave cytokeratin K7 and K18 (Chen *et al.*, 1993 Zhang and Schneider, 1994), which promoted host cell lysis by disrupting the mechanical integrity of infected cells. K18 was shown to be cleaved at the amino terminal head producing headless intermediate filaments unable to maintain the structural integrity of the cell. Sequence data of K18 showed two potential consensus cleavage sites, one conforming to XGXG, the other XGGX. Cleavage was only observed at the XGXG site (70-74) not the XGGX (66-70) which was difficult to explain on two accounts, firstly why the selectivity and secondly, how was the cleavage controlled temporally and spatially.

The viral protease was then implicated in the entry of virus to host cells (Cotten and Weber, 1995 Greber *et al.*, 1996). In this role the protease would have to be active in the absence of the activating peptide. This is not unique as it had been reported that the protease processed pTP when co-expressed with protease in baculovirus expression system (Webster *et al*, 1993). Protease was

also reported to be active when it entered the reduced environment of endosome. Cotten and Weber (1995) suggested that the protease may be responsible for the observed degradation of hexon and VI during the entry to the cell of a virus. The consensus cleavage site of the protease would be violated in that the cleavage observed occurs at PEGR-G (1995 Greber *et al.*, 1996), however, violations have been seen in other serotypes such as Ovine 287 strain (Vrati *et al.*, 1996) where the protease is capable of processing at MRAT-G and NTGW-G sites.

Jones *et al* (1996) using in vivo pTP digestion assays in sf9 cells with wild type and site directed mutants (C104A, C122A and H54S) found that the mutant proteases were unable to process pTP activity. H54 and C122 were implicated in the catalytic mechanism, retention of activity with C104A suggested that C104 had a role in the protease activation mechanism.

C104 was finally removed from contention of having a role in the active site by a series of tryptophan fluorescence studies. Conformational changes induced by the binding of the activating peptide changed the fluorescent character of the protease. No change in fluorescence state was observed with the mutant C104S protease with increasing concentration of activating peptide. When C122S was incubated with activating peptide an increase in fluorescence was observed. This result could be explained if C10 of the activating peptide binding interacted with C104 of the protease causing a conformational change. If this were the case C122 would be the active site nucleophile (Jones *et al.*, 1996). Previously reported inactivity of C104A mutant was therefore not due to the mutation of an active site residue but rather the result of the activating peptide being unable to form a disulphide bond with the protease.

Protein	Structure/functional	References
pVI	Structural	Webster *et al* 1993
pVII	Structural	Webster *et al* 1993
pVIII	Structural	Hannan *et al* 1983
pTP	Functional	Smart and Stillman 1982
pIIIa	Structural	Webster *et al* 1993
11k (pMu)	Functional	Anderson and Weber (1998)
L1-52K	Structural/functional	Hasson *et al* (1992)
Cytokeratin 7	Structural	Chen *et al* 1993
CytoKeratin 18	Structural	Chen *et al* 1993
Actin	Structural	Chen *et al* 1993

Table 1.01 Summary of the Viral and Cellular Proteins Which are Targets for the Protease.

With the solution of the three dimensional structure (Ding *et al.*, 1996), histidine 54, aspartate/glutamine 71 and cysteine 122 were confirmed as the active site triad. The active site lies within a 25 Å long bent groove approximately 8 Å wide with C122 and H54 at the centre. Cysteine 122, histidine 54 and glutamate 71 are believed to form a charge relay system analogous to that of papain with cysteine 122 and histidine 54 acting as the nucleophilic anion pair (Ding *et al.*, 1996).

Further work (Diouri *et al.*, 1996) on the preference of the active site using flourogenic substrates and the processing of pVI to iVI to VI showed that the XGXG motif was more rapidly cleaved than XGGX. This added further support to the conclusions of Webster *et al.*, (1989) who with work on MSGAGFSW and MSGGAFSW suggested the same phenomenon. Diouri suggested that preference was a physiological role as the first cleavage of pVI to iVI is at XGXG site and in infection would occur in the absence of activating peptide. iVI to VI cleavage occurs later at XGGX site. This is also observed

with the processing of pTP protein during early and late infection. XGXG site is thought to be the preferential site of processing for the protease in the absence of activating peptide. Protease mediated cleavage of cytokeratin in the absence of activating peptide may represent one aspect of the control of the Adenovirus life cycle.

1.7.2 Identification of the Viral Cofactor and its Properties.

Recombinant protease from pT7AD23K5 vector (Anderson 1990) was shown to cleave ^{35}S-methione labelled precursor proteins. The activity of the protease was greatest with extracts prepared from young fully assembled virions which suggested that there may be some factor involved in the control of the protease.

If the protease was purified to homogeneity from insect cells and incubated with synthetic substrate MSGGAFSW (mimicking pVII consensus cleavage site) cleavage was not observed (Webster *et al.,* 1993) unless protease was incubated with crude sources of Ad2ts1. When Ad2ts1 was incubated with synthetic peptide substrate cleavage was only observed at the permissive temperature. It was obvious from these observations that there was some factor within the Ad2ts1 preparation which was capable of activating the protease.

Eventually an 11 residue peptide (GVQSLKRRRCF) was identified which was capable of restoring the activity of the recombinant protease in the absence of Ad2ts1 extract (Webster *et al.,* 1993; Mangel *et al.,* 1993). This activating peptide as originating from the C-terminus of pVI, at a consensus cleavage site of the protease.

The proposed mechanism of protease activation is that the oxidised dimeric form of activating peptide is reduced by the protease during heterodimer formation. The activating peptide binds the protease via a disulphide bond

formed between C104 of the protease and C10 of the activating peptide. The interaction between pVI and the protease is only thought to occur once the virion has assembled.

Sequence elements from both ends of the activating peptide important in the activation of protease (Cabrita *et al.*, 1997). Truncated activating peptides drastically reduced the activity of the protease. Removal of N-terminal glycine lead to a significant reduction in substrate cleavage, and if the valine was removed as well, substrate cleavage was almost abolished. An 11 residue peptide with serine at position 10 did not form a heterodimer with the protease. Therefore, not only were glycine 1 and valine 2 important but cysteine 10 also confirming the conjectures made by Webster *et al* (1993).

Is the activating peptide exclusively responsible for activation of the protease? Other groups have reported that DNA, polyanions, (Mangel *et al.*, 1993) more recently, the peptide sequence VEGGS (Diouri *et al.*, 1996a) as having activating properties. That DNA is a cofactor is disputed, Mangel's group reported loss of activity if DNase 1 was added to activity assays. If Ad2 DNA was then added activity was restored. Kemp and co-workers have been unable to repeat these results i.e. both with DNA and VEGGS (unpublished).

1.7.3 Classification and Mechanism of the Viral Protease

Proteolytic enzymes are broadly classified into 4 distinct classes according to their essential catalytic residues at the active site (Bond and Butler, 1987). Serine, cysteine (thiol), aspartate (acidic) and metalloproteases form the major categories. Generally metalloproteases contain metal ions such as zinc at their active site which is thought to enhance the nucleophilic properties of water, as well as polarising the scissile peptide bond prior to nucleophilic attack.

Aspartate proteases contain an aspartyl diad in the active site and are thought to act via an acid-base catalysis mechanism. Although recently, water has been implicated at the active site.

Serine and cysteine proteases are similar in their catalytic mechanism, although the side chain responsible for the nucleophilic attack is different. In the case of serine, the side chain is unusually active and nucleophilically attacks the carbonyl carbon of the scissile peptide bond. The attack involves an intermediate step - acylation of the serine or cyteine side chain.

Usually enzymes are classified on the basis of inhibition study profiles. However, the distinction of serine and cysteine proteases is sometimes difficult to make as there is often considerable overlap in the inhibitors. This was exemplified in the classification to the Adenovirus protease. Initially the protease was classified as a non-metallo serine protease (Bhatti and Weber, 1979). Subsequent studies revealed that it was a cysteine protease (Webster *et al.*, 1989b) and since then a number of independent studies have verified this fact (Cotten and Weber, 1995; Greber *et al.*, 1996; McGrath *et al.*, 1994; Rancourt *et al.*, 1994; Tihanyi 1993; Ding *et al.*, 1996).

The cleavage mechanism of a cysteine protease is outlined below. Basically the mechanism is initiated by a nucleophilic attack on the carbonyl carbon of the peptide bond followed by the acylation of the thiol group and the release of the N-terminal peptide group (fig 1.09 a and b).

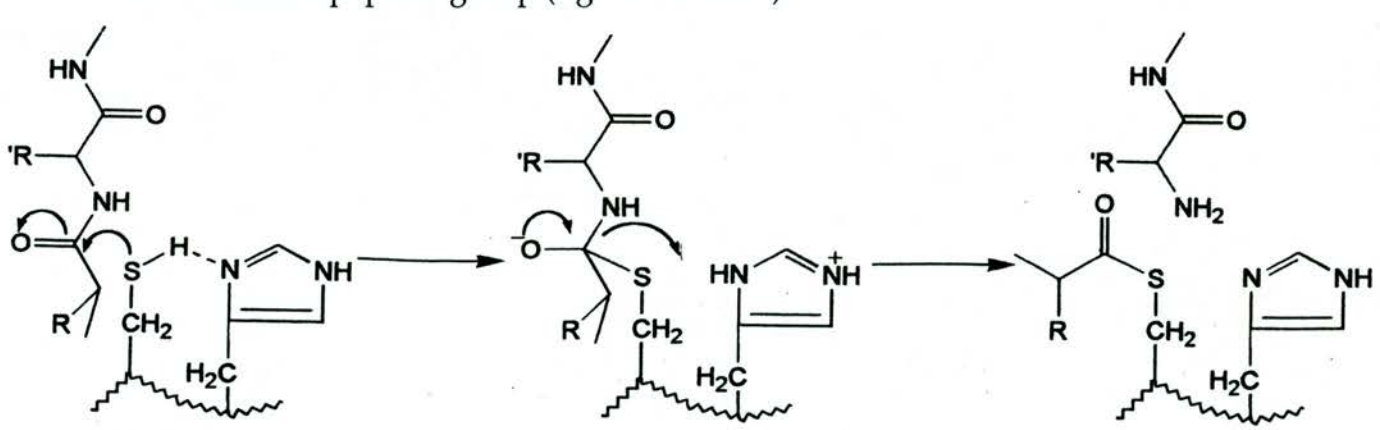


Fig. 1.09a

The next step is the hydrolysis of the thiol acyl group to release the C-terminal peptide.

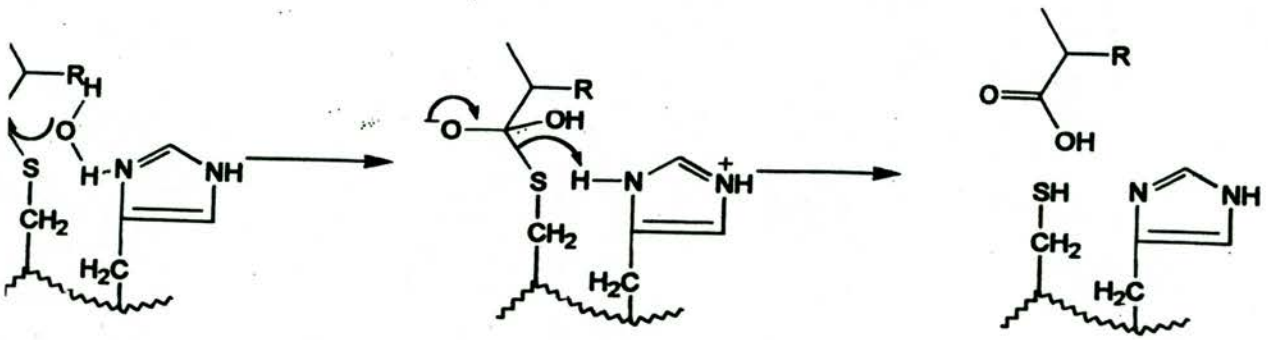


Fig 1.09b

The reactive site is composed of a reactive cysteine side chain coupled to a histidine which functions as a proton acceptor and donor. The curly arrows in the figure depict electron movements associated with these events. Thus another useful way of categorising proteases is to compare the primary sequence of the protease to other know proteases looking for sequence homologies. From the sequence alignment (fig 1.07) of the Adenovirus proteases sequenced to date we can see the conservation of histidine 54, glutamic/aspartic acid 71 and cysteine 104 and 122. The number system refers to that of Adenovirus type 2 sequence.

1.8 Nuclear Localisation Signals

In all Eukaryotic cells a nuclear membrane surrounds the genetic material, separating the translation of mRNA to proteins from transcription of the genetic material to produce the mRNA. This compartmentalisation demands that transport proceeds efficiently and with proper discrimination between compartments.

In common with all eukaryotes, embedded within the nuclear membrane is a structure called the nuclear pore complex (NPC) which mediates transport

events. The nuclear pore is an aqueous channel which allows the diffusion of ions and small molecules, and selective entry of large macromolecules. These large molecules have been shown to bear a short signal which facilitates entry.

1.8.1 Nuclear envelope

Basically the nuclear envelope acts as a molecular sieve. Proteins with molecular weight of less than 50 kD injected into cells were shown to form an equilibrium across the membrane at a rate inversely proportional to their size (Peters, 1986). Larger molecules were found to be excluded from the nucleus. This was thought to be due to the size of the nuclear pore, quoted to be 4.5-6 nm. Exogenous nuclear proteins, no matter how large were able to enter the nucleus showing that they must have a specific mechanism of entry (Bonner, 1975; De Robertis *et al.*, 1978). Recent evidence illustrated that small nuclear proteins also enter the nucleus by an energy dependent mechanism (Breeuwer and Goldfarb, 1990). No macromolecule with a nuclear physiological relevance has been identified to enter the nucleus by simple diffusion.

1.8.2 What is the Signal for Nuclear Localisation?

The first defined nuclear localisation signal was identified in SV 40 large T antigen. Gene fusion experiments showed that the sequence of PKKKRKV was necessary and sufficient to confer nuclear localisation. Since then a whole menagerie of sequences have been identified with several common factors. The sequence must be accessible and is usually short, less than 10 amino acids, consisting of basic residues. The sequence may occur anywhere in the primary structure, and there may be more than one. Bipartite sequences may be separated by 10 amino acids (Dingwall and Laskey, 1991) but as with all

sequences, they must be exposed on the surface of the protein (Roberts *et al.*, 1987).

1.8.3 How is the Signal Recognised ?

Initial research illustrated that nuclear transport conformed to saturable kinetics (Adam and Gerace, 1991) implicating the role of a receptor in the transport process. Since then a whole catalogue of cytosolic receptors have been recognised as identifying and binding a range of nuclear localisation signals (NLS) presented (Nakielny and Dreyfuss, 1999). These receptors are believed to dock at the nuclear pore complex. For a review of the docking process and proteins involved see Jans *et al* (2000).

1.8.4 Nuclear Pore Complex

The NPC is embedded in the nuclear envelope. Evidence from rapidly fixed thin nuclear membrane sections of xenopus laevis oocytes revealed fibrillar arrangements extending from the cytoplasmic and nuclear faces of the NPC (Richardson *et al.*, 1988). Proteins containing NLS are recognised by a corresponding receptor which are diffuse throughout the cytosol. Morphological studies suggest that the initial binding is at the periphery of the NPC associated with the filaments which emanate from the nuclear pore. RAN.GTP is a transport protein specifically confined to the peripheral cytoplasmic region of the NPC (Melchior and Gerace, 1995). RAN.GTP and ligand-receptor complex are bound to the NPC, where they interact and GTP is hydrolysed. This hydrolysis fuels the ligand-receptor complex's commitment to the more distal steps of transport. RAN.GDP enters the nucleus with the ligand-receptor complex, RAN is recycled along with the receptor (Gorlich *et al.*, 1996; Izaurralde and Adam 1998; Hieda *et al.*, 1999).

Proteins containing short sequences of basic amino acids are transported to the nucleus with the help of importin α. All proteins identified to date as

having a physiological function in the nucleus contain an NLS. This does not just affect cellular proteins. The Adenovirus DNA binding protein, polymerase, E1a, fibre and pVI contain nuclear localisation sequences. These are important in the control of the viral life cycle ensuring that these proteins are in the correct compartment to carry out their function.

1.9 Aims

The aim of the study was to further characterised the spatial and temporal control of the Adenovirus type 2 protease and develop our understanding of the activation mechanism of the protease. One interesting fact we to try and resolve was the shift of the protease between 24-28 hpi from the cytoplasm to the nucleus (Vaughan, 1997).

We also wanted to develop an assay to assess the relative binding affinities of peptides which were designed to further our understanding of the role or the individual amino acid's role within the activating peptide.

2. Materials and Methods

This chapter describes all the procedures and methods used throughout the duration of the study. At the beginning of each section a brief theoretical explanation of the underlying methods is presented. Chemicals and reagents were all purchased at the best available grade from Sigma unless otherwise stated.

2.1 Production of Recombinant Proteins

2.1.1 Agarose Gel electrophoresis

Agarose gel electrophoresis was used to separate DNA according to its molecular weight. The negatively charged phosphate backbone of the DNA electromigrates to the anode. To each DNA sample 5 x concentrated loading buffer was added (40%w/v sucrose, 5x TBE [0.225M Tris-Borate, 5mM EDTA], 5 units/ml RNAse (Qiagen) and 0.05% bromophenol blue). A stock solution of 5x TBE running buffer was prepared and diluted accordingly. Gels used ranged from 0.5% to 4% (w/v) agarose TBE buffer with 0.00005% (w/v) ethidium bromide to mark DNA. When 0.5 % (w/v) gels were prepared a 2 mm underlay of 2% (w/v) agarose was used (Sambrook *et al.*, 1989). Samples were run at a constant 80V until the dye front had migrated 2/3 through the gel. Separated bands were visualised using UV illumination, and the appropriate bands excised with a clean scalpel.

2.1.2 Determination of DNA Purity and Concentration

For the purpose of Prokaryotic and Eukaryotic expression studies and DNA sequencing the purity and concentration of DNA need to be determined accurately. The most common procedure for assessing the purity of DNA preparations is to assume a ratio of the absorbance 260:280 nm of 2 is indicative of pure DNA.

The concentration of DNA was determined either on 1% agarose gels comparing unknown samples to a series of standards, or assuming that an absorbance of 1 at 260 nm was equivalent to 50µg/ml (Sambrook et al., 1989) of DNA.

2.1.3 PCR

PCR was used to amplify DNA sequences and perform DNA mutagenesis. Oligonucleotides were designed to complement the flanking regions of DNA and PCR reactions simulated using Amplify software (Engels, 1992). All primers were purchased from Oswell. PCR cycle consists of a denaturing, annealing and elongation step. Denaturation at 72°C separated the DNA strands, the sample was cooled to the annealing temperature which varies from primer to primer. Annealing allows the primer to specifically bind the target DNA. Incubation at 95°C permits elongation of DNA strands by Vent DNA polymerase.

Primers were diluted to a stock concentration of 10µM. DNA samples were ethanol/glycogen precipitated (Sambrook *et al.*, 1989) and resuspended in sterile DNAse free water. A typical reaction mixture volume was 100µl, and comprised of 10µl 10x ThermoPol buffer (New England Biolabs), 2µl Vent polymerase, 5µl of each primer, 10µl dNTP solution (0.5 mM each of d ATP, dCTP, dGTP, dTTP), 2µl DNA (20-200ng) and 66µl water.

The reaction mixture was incubated at 95°C for 5 min to separate the strands and then underwent 40 cycles of amplification. Each cycle comprised of an incubation at 95°C for 45 s, X°C (optimised annealing temperature), followed by one at 72°C for 45 s and a final incubation at 72°C for 5 min which allowed exonuclease activity to increase the efficacy of the DNA copy. DNA fragments were then separated on, and extracted from, 1-3% agarose gels depending upon the size of the fragment.

2.1.4 Primers

Primers were designed using primer design software. The following primers were used throughout the study. Bold ***italic*** text indicates the unique restriction sites.

For pUHD10-3

5′ Eco R1 Protease	5′ CAACAAC***GAATTC***CGCCATGGGCTCCAG 3′
3′ Bam H1 Protease	5′ CGGCGC***GGATCC***CAAGGGTGGGG 3′
Internal Protease F	5′ CGCCTGTGGACTATTCTGCT 3′
Internal Protease R	5′ AGCATAATAGTCCACAGGCG3′
pUHD10-3 upstream	5′ CAGAGCTCGTTTAGTCAACCG 3′
pUHD10-3 downstream	5′ AAAGCAATAGCATCACAAATTTCA
5′ Eco R1 pVI	5′ GGTCCTGTAA***GAATTC***GGTAGAATGGAAGAC 3′
3′ Bam H1 pVI	5′GGCGCGACAT***GGATCC***ATACATGACACACATA
3′	
Internal pVI F	5′ GGTCTGGGCGTTAGGATACA 3′
Internal pVI R	5′ GTCCTTGCGGGATGTTCTTCT 3′

For GFP

5′ Xho 1	5′ CCG***CTCGAG***GCTCCAGTGAGCAGGACTG 3′
3′ Xba 1	5′ TGC***TCTAGA***TTACATGTTTTTCAAGTGACA 3′

2.1.5 DNA Gel Extraction Protocol

DNA from agarose gels was extracted essentially as described (Qiagen, 1999). Excised bands were weighed and 300µl QG1 added per 100mg of band and incubated at 50°C for 10 min to melt the agarose. One gel volume of isopropanol was then added, mixed and the resultant suspension applied to a QIAquick spin column and centrifuged at 13000rpm for 1 min (MSE Microcentaur). The column was washed once with 0.5ml QG1 buffer and once

with 0.75ml PE buffer with the eluate discarded between washes. Residual PE buffer was eluted at 13000 for 1 min and the column air dried for 5 min Either 50µl EB buffer or sterile ddH_2O was applied to the column and incubated for 5 min at room temperature before DNA was eluted from the column by two centrifugation steps at 13000 rpm for 1 min.

2.1.6 DNA Sequencing

All plasmid constructs which needed to be amplified were transformed into *E. coli* DH5α. DNA and primers were supplied to Alex Houston at the University of St. Andrews DNA sequencing unit. Samples were subject to 25 cycles of PCR amplification in the presence of ABI Prism BigDye Terminators. The termination reaction mix contains the 4 standard DNA bases along with 4 terminator bases which are equivalent to each base with a specific dye label which when incorporated into DNA terminate the polymerisation. Elongation proceeds with AmpliTaq DNA polymerase at 96°C elongation terminates at the site of dye terminator incorporation. Termination is random and produces a range of lengths of PCR products. These products are separated by long range bis-acrylamide electrophoresis (ABI 377 DNA sequencer).

DNA sequences were viewed using Edit View software and sequences compiled with DNA Strider. Insertion of the correct sequence was confirmed by comparison of the sequences returned to the known DNA sequence from the Blast Database.

2.1.7 Cloning

Inserting a gene into a plasmid is a two step process. After PCR amplification of the gene of interest the product contains specific restriction sites which are cleaved with restriction endonucleases, as was the target plasmid. These restriction enzymes recognise specific palindromic sequences of DNA and

produce cohesive ends complementary to those in the plasmid, which were ligated with T4 DNA ligase enzyme.

2.1.8 Source of Oligonucleotides

5′ phosphorylated oligonucleotides of HPLC purity were ordered from Oswell DNA service (www.oswel.com).

2.1.9 Restriction Endonuclease Digestion of DNA

Typical reaction mixture contained 5µl DNA (1mg/ml), 20 units of each restriction enzyme, (Promega and New England Biolabs), 5µl 10 x buffer, 0.5µl BSA (1mg/ml), and 35.5µl water. The mixture was incubated at 37°C for 4 h.

2.1.10 Hybridisation of Oligonucleotides.

In order to insert DNA sequences of less than 60bp into a plasmid complementary oligonucleotides were designed and annealed rather than being PCR amplified from original DNA. 50 µl of each complementary oligonucleotide (diluted in 1 x TE [10mM Tris-HCl, 1mM EDTA pH 8] buffer to a concentration of 100ng/µl) was mixed, boiled for 5 min, and cooled to 25 °C below their melting temperature. Hybridised oligonucleotides were stored at -20 °C. The oligonucleotides were designed with cohesive overhanging ends and directly cloned into the target plasmid.

2.1.11 Ligation

Prior to the ligation procedure, vector and insert DNA concentrations were estimated on 1% agarose gels using known standards.

Ratios 1:3. 1:7 and 1:9 of plasmid to insert were used for ligation. A typical reaction mix consisted of 2µl 10x concentration ligase buffer, 400 units T4 DNA ligase (Promega and New England Biolabs) in a final volume of 20µl in

sterile HLPC grade water (Rathburn) and incubated at room temperature for 4 h..

2.1.12 Purification of DNA in preparation for electroporation

The products of ligation were electroporated due to the high efficiency of transfection. Chemically competent cells were used for all other transformations using the heat shock technique.

Ligated DNA preparations were precipitated to remove traces of salts and other contaminants which would interfere with electroporation. The presence of salt would short circuit the electroporation chamber producing an arc of electricity resulting in bacterial death.

To 20µl of ligation mixture 180µl water, 400µl of 100% ethanol and 2µl of glycogen (20mg/ml) were incubated for 30 min at -70°C. Samples were centrifuged at 14000 rpm, 4°C for 30 min, and the DNA pellet washed twice with 70% ethanol 4 °C. DNA was air dried for 5 min and resuspended in 10µl of sterile water.

2.2 Preparations of Competent Cells

In order to transform plasmid into bacterial cells for protein over-expression or plasmid DNA amplification, prokaryotic cells were made competent (i.e. were treated to prepare them to receive DNA). Cells were prepared to be either electro or chemically competent as outlined below.

2.2.1 Preparation of Electrocompetent Prokaryotic Cells

For electrocompetence a single DH5α colony was selected and incubated for 8 h. at 37°C in 10ml Luria Broth (Luria-Bertani medium (LB: 1% bacto-tryptone, 0.5% bacto-yeast extract, 1% NaCl, pH7 (Sambrook et al., 1989)) at 250 rpm (orbital shaker). 500 µl of the inoculum was used to seed 500ml of LB and incubated at 37°C until the absorbance at 600nm reached 0.4-0.6 . The inoculum was placed on ice for 2 h., cells were harvested at 4000 rpm

(Beckman JA 4 rotor) and washed twice with cold sterile 1mM HEPES. Cells were resuspended in 2 ml 30% glycerol, snap frozen in liquid nitrogen and stored at -70 °C.

2.2.2 Preparation of Chemically Competent Cells

A single colony of BL21 DE3 from a stock agar plate was incubated at 37 °C overnight in 5 ml of LB. 1 ml of this was used to inoculate a 50 ml culture of LB at 37 °C and incubated until the absorbance at 600nm reached 0.4-0.6. Cells were centrifuged (IEC Centra-3R) at 3000 rpm for 10 min at 4 °C and the supernatant discarded.

The cell pellet was resuspended with 10 ml of ice cold filter sterilised 0.1M $MgCl_2$ and centrifuged as above. Supernatant was discarded and the pellet resuspended with 4 ml of filter sterilised ice cold 0.1 M $CaCl_2$ and left on ice for 30 min

2.3 Transformation Techniques

Two methods were used to deliver DNA across the cell wall into the bacteria.

(i) 40 µl of electrocompetent DH5α cells were incubated on ice with 10 µl of purified ligation mixture for 15 min in a 2mm electroporation cuvette (EQUIBIO). Electroporation was carried out at a voltage of 2500, and a current of 201 A for 5ms. After electropration, cells were resuspended in 500 µl of LB and incubated at 37 °C for 45 min before being spread on agar plates containing the appropriate selective antibiotic.

(ii) 200 µl of chemically competent cells and 10 ml of DNA (200ng DNA) were mixed and incubated on ice for 30 min, Cells were heat shocked at 42°C for 90 sec and placed on ice for 5 min. 30 and 100 µl aliquots of cells were spread on agar plates with the appropriate selective antibiotic. Plates were incubated at 37°C overnight.

2.3.1 Agar Plates

LB-Agar was prepared by adding 15g of bacto-agar per litre Luria-Bertani medium and autoclaving followed by the addition of the appropriate selective antibiotics where necessary when the agar was sufficiently cool so as to avoid degradation of the antibiotic. 20 ml of agar was poured into each sterile Petri dish.

2.4 DNA Purification

Two methods of purification were used depending upon the amount of DNA required.

2.4.1 Small Scale DNA Purification (Mini-Prep)

Mini-Prep procedure was used to prepare DNA for sequencing and analytical restriction digests.

Essentially the protocol as described in the 1997 edition of the Qiagen Mini-Prep manual was followed. Single colonies were selected from overnight agar plates and used to inoculate 4-5 ml of LB (plus appropriate antibiotic) and incubated at 37°C for 16 h. The remaining culture was centrifuged at 13000 rpm for 1 min, the supernatant discarded and the pellet resuspended in 200 µl P1(Qiagen). Cells were lysed with 200 µl P2 (Qiagen) for 5 min at room temperature. Lysis was terminated by the addition of 200 µl of P3 buffer (Qiagen). The resulting precipitate was centrifuged at 13000 rpm (MSE Microcentaur) for 10 min and the supernatant applied to a Mini-Prep column (Qiagen). DNA bound to the column was washed with 750 µl PE buffer (Qiagen) and then centrifuged as above to remove any residual ethanol. The column was air dried for 5 min and DNA eluted in either 50 µl of EB buffer (Qiagen) or sterile DNase free water. DNA was stored at -20°C.

2.4.2 Large Scale DNA Purification (Maxi Preps)

DNA Maxi preparations were carried out essentially as described in the Qiagen Maxi-prep guide (Feb. 1995). This procedure was used to prepare large stocks of plasmid for cloning, Eukaryotic and prokaryotic expression studies and DNA sequencing. Typically 0.75-1 mg of DNA per 500ml bacterial culture was recovered.

A single colony selected from an overnight agarose plate or a scraping from a glycerol stock was used to seed 10 ml LB (antibiotic) which was incubated for 8 h. at 37°C. 500 µl were transferred to 500ml of L-Broth (antibiotic) incubated for 16 h. at 37°C. Cells were harvested at 6000rpm (Beckman JA-17 rotor) for 20 min and resuspended in 10 ml of P1/RNase. Cells were lysed with 10 ml P2 for 5 min and the lysate neutralised with 10 ml of buffer P3. The resulting suspension was incubated on ice for 15 min then centrifuged for 30 min at 15000 rpm (Beckman JA-17 rotor). The supernatant was filtered though two layers of muslin into a Qiagen 500 column pre-equilibrated with 20 ml of buffer QBT. The column was washed twice with 30ml of buffer QC. DNA was eluted with 15ml QF and precipitated with the addition of 10.5 ml propan-2ol. DNA was pelleted at 4000 rpm for 30 min (Sigma Laborzentrifugen 3K10, 5500 rotor) and washed twice with 70 % ethanol. The DNA pellet was air dried and resuspended in DNase free water to a final concentration of 1mg/ml and stored at -20°C.

2.5 Glycerol Stocks

Bacterial cells were stored in 30% glycerol at -70°C. The glycerol prevents the formation of ice crystals which would otherwise puncture bacterial cell walls. Stocks were prepared by adding 300 µl of sterile glycerol to 700 µl of bacterial culture in polypropylene tubes. Samples were gently mixed by pipette action, snap frozen in liquid nitrogen and stored at -70°C.

2.6 Prokaryotic Expression and Purification of Recombinant Adenovirus Type 2 Protease

2.6.1 Production and Purification of Recombinant Proteases

A recombinant source of Adenovirus protease was essential to provide reproducibly an abundant source of high quality enzyme for studies into activation. It also negated the need to work with large volumes of virally infected cell cultures. Wild type protease, C104A and C122A mutants used in this work had already been cloned into pET vectors and so the relevant plasmid was transformed into the relevant cell line for expression or amplification.

2.6.2 Production of Recombinant Protease for Enzymatic Assays

The Adenovirus type 2 protease gene was amplified using PCR from viral DNA. Primers were synthesised to the flanking regions with unique restriction sites which enable the protease gene to be inserted into the target plasmid.

Protease was cloned into pET-11C vector (NOVAGEN) and the expression was under the control of an inducible T7 promoter. Normal laboratory strains of *E coli.* do not endogenously express T7 RNA polymerase and so expression requires the presence of chomosomally integrated T7 RNA polymerase gene. pET11-C also contains the β-lactamase gene which confers ampicillin resistance allowing selection of successfully transformed bacteria. Protein expression was then induced by the addition of isopropylthiogalactopyranoside (IPTG). IPTG induces the expression of T7 polymerase and thus the expression of the protease.

2.6.3 Purification of Protease

To purify the expressed recombinant protease one first has to extract a soluble protein and then recover it from an heterogeneous mixture. Two methods were employed to release the expressed protease. Samples were

either sonicated-which uses high frequency sound waves to disrupt the cell wall and shear DNA, or freeze-thawed in the presence of lysozyme to breakDown the bacterial peptidoglycan cell wall followed by DNase treatment to digest DNA. Samples from either treatment were centrifuged to remove insoluble particulate matter. Once the protease has been extracted, purification by FPLC using anionic and cationic exchange was performed.

2.7 FPLC Protein Purification

Fast Protein Liquid Chomatography (FPLC) (Pharmacia P-500) was used to purify recombinant protease. Samples were filtered (0.2μ) and applied to 10 or 50 ml superloops. The apparatus allowed the application of an ionic gradient to elute proteins whose appearance was recorded by an UV detector at 280nm. Initially purification involved two steps which were then refined to a single step.

2.7.1 Diethylaminoethyl Sepharose Column

DEAE-Sepharose is a weak anion exchange group, most proteins stick to the column and are thus separated from the protease on the basis of charge since at the operating pH the protease carries a net positive charge.

FPLC parameters were set as follows, pump A with buffer A -50 mM Tris-HCL pH 8, and pump B with buffer B - 50 mM TRIS-HCL/HCl, 1M NaCl pH 8, absorbance sensitivity at 1, chart recorder at 0.25cm/min and a flow rate of 1 ml/min

The supernatant was filtered (0.2μ) before being applied to a DE-Sepharose column equilibrated with buffer A. 2ml fractions were collected and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE).

2.7.2 Heparin-Sepharose and Carboxymethyl-Sepharose Columns

The Heparin and Carboxymethyl moieties are both cationic exchange groups thus interacting with proteins which carry a net positive charge at the operating pH (pH8). The negatively charged moieties of heparin resemble the phosphate backbone of DNA and consequently interact with positively charged proteins, often those implicated in DNA binding. The protease does not bind to Heparin-Sepharose at pH8, but does bind to the CM-Sepharose. The protease was eluted from the CM-Sepharose column with 150mM NaCl.

Fractions containing protease from the DEAE-Sepharose separation were pooled and run though consecutive heparin and carboxymethyl columns equilibrated with buffer A. The parameters were set as follows; flow rate - 1 ml/min chart recorder 0.25 cm/min and absorbance sensitivity 0.2. Protease was eluted 50mM Tris-HCl, 150mM, pH8. Sample purity was confirmed by SDS PAGE electrophoresis and Western blot.

2.7.3 Cell Culture, Induction and Harvesting

Bacteria were grown in either fermentor (Microfermentor, New Brunswick Scientific) or flask cultures using LB medium, or in flask culture using a Dulbecco's Modified Eagle's Medium (D-MEM Sigma Catalogue number D5671) based medium

Both methods using LB were initiated in the same way, 5 ml of LB (antibiotic) with either a single colony of *E. coli* from a stock plate or a scraping from a glycerol stock were incubated for 6-8 h at 37°C. 2 ml of this was used to seed 500ml of LB (antibiotic) which was incubated overnight at 37°C with constant shaking. This was used to inoculate 10 litre LB (fermentor). The culture was incubated at 37°C until the absorbance at 600 nm reached 0.4-0.6 at which point the temperature was reduced to 25°C and induction with IPTG final concentration of 1mM initiated for 3 h.

For the conical flask method one 500 ml culture was grown until the absorbance at 600nm reached 0.4-0.6 at which point the culture was induced for 3 h with IPTG to a final concentration of 1mM 25°C.

For growth in D-MEM a single colony was selected from a stock plate and resuspended in 1ml sterile water. An aliquot was removed and the absorbance at 600 nm determined. Appropriate dilutions were made from the remainder to allow each of 3 L-Agar/Antibiotic plates to be spread with 300µl of a suspension with an absorbance of 0.001. Plates were incubated at 37°C overnight. Cells were harvested from plates, washed with 20ml of LB/antibiotic and resuspended in 10 ml LB. An appropriate volume of this culture was used to inoculate 500ml D-MEM (Gibco), 50mM HEPES pH7.5 containing appropriate antibiotic at an initial absorbance at 600nm of 0.15.

The culture was incubated at 37°C until absorbance at 600nm was 0.4-0.6. The culture was induced with 1mM IPTG as above for 3 h.

2.7.4 Harvesting Cultures

Cells were harvested in one of two ways. Large scale fermentor cultures were centrifuged in 1 litre vessels at 4000rpm (Beckman J6-HC) for 30 min. The supernatant was discarded and the pellet washed twice with PBS. The pellet was resuspended in 90 ml of suspension buffer (50mM Tris-HCl, 5mM EDTA, 4% glycerol, pH8) and stored at -20°C.

Small scale cultures were harvested at 10000 rpm (Beckman J2-21) for 10 min Again the cell pellet was washed twice in PBS before being resuspended in 10 ml of suspension buffer and stored at -20°C.

2.8 SDS PAGE

2.8.1 Analysis of Protein Expression by SDS-PAGE

Confirmation of induction of the desired protein was provided by separating samples of the culture taken at hourly intervals using SDS-PAGE electrophoresis. The induced protein was identified by the appearance of a band whose intensity increased with time. Confirmation of the identity of specific protein bands was obtained by protein sequencing, mass spectrometry or Western blotting.

1 ml samples were centrifuged for 1 min at 13000 rpm, the supernatant discarded and pellet resuspended in 100 µl 10mM Tris-HCl, 1mM EDTA, pH8 and incubated at room temperature with 20µl 20mg/ml lysozyme. After 1 min of sonication the soluble and insoluble fraction were separated by centrifugation at 13000 rpm. Supernatant was collected and 50µl of boiling mixture added. The insoluble pellet was resuspended in 100µl of boiling mix and analysed on 15% SDS PAGE gels.

2.8.2 Gel Preparation

Protein expression was analysed using Bio-Rad Mini Gel Kit. The separating gel comprised 5.6 ml 40% acrylamide, 0.65 ml 20% bis -acrylamide, 5.6 ml 1M Tris-HCl pH8.7, 150 µl 10% SDS and 2.9ml dH_2O. Polymerisation was catalysed by the addition of 12.5µl TEMED and 50µl ammoniun persulphate 0.1mg/ml (APS). The separating gel was then overlaid with water saturated butanol which was removed by washing with 70% ethanol and then distilled H_2O once the gel had set.

Stacking gel consisted of 1.3 ml 40% Acrylamide, 700 µl 20% bis-acrylamide, 1.25 ml 1M Tris-HCl pH6.9, 100 µl 10% SDS, 6.75ml dH_2O. Polymerisation was catalysed with 25µl TEMED and 50µl APS. To each well 10-15µl of sample was

added and electrophoresis carried out at 40mA per gel for 40 min in chamber buffer (25mM Tris-HCl, 0.192 M glycine, 0.1% SDS).

2.8.3 Sample Preparation

A mixture of reducing (boiling) solution (1 part mercaptoethanol : 3 parts magic mix [5ml 1M Tris-HCl pH6.9, 10ml 20% (w/v) SDS, 5ml glycerol]) was boiled for 5 min with 15-25µl of sample.

2.8.4 Gel Staining

Gels were stained in 10 ml of Coomassie Blue stain (2.5g Coomassie Brilliant Blue R250, 454ml methanol, 92ml glacial acetic acid and 454ml H_2O) for 10 min and then destained (25% methanol v/v, 7.5% acetic acid v/v) until bands were visible.

2.9 Determination of Protein Concentration

Protein concentration was determined using SBTI (Soya bean trypsin inhibitor) standards on 15% SDS PAGE gels. In some cases Bradford's reagent (Bradford, 1976) was used to estimate the protein concentration comparing unknown samples to a SBTI standard curve. This method is based upon the binding of Coomassie Blue binding to protein and results in a shift in the absorption maximum of the dye from 465nm to 595 nm. The increase absorbance at 595 nm was monitored and plotted on a standard curve.

200 µl of Dye Reagent Concentrate was added to 800 µl of sample or standard and then mixed thoroughly. After 5 min the absorbance at 595 nm was measured against a reagent blank and protein concentration estimated from the standard curve.

2.10 Preparation of Protein Samples for Identification

2.10.1 Western Blot Technique

Western blotting technique transfers proteins from a polyacrylamide medium to nitrocellulose medium. Samples were separated as above using SDS PAGE.

The resulting gel was laid on a PVDF membrane and sandwiched between 2 sheets of 3mm paper soaked in transfer buffer (10mM CAPS (3-[cyclohexylamino] -1- propanesulfonic acid), 20% Methanol pH11) and two fibrous pads. Transfer proceeded at a constant current of 300mA for 2h.

The PVDF membrane was blocked for 10 min in blocking buffer (5% Boots dried milk powder, 0.1% TWEEN 20, in 250ml PBS). The PVDF membrane was incubated for 30 min in 20 ml of fresh blocking buffer containing primary antibody (diluted 1 in 4000) The membrane was then washed in 10 ml blocking buffer and washed three times for 10 min in blocking buffer. The membrane was then incubated for 30 min with anti-primary antibody antibody coupled to horse radish peroxidase (Amersham) diluted to a final concentration of 1 in 5000 in blocking buffer. The PVDF membrane was washed as above and finally rinsed in 0.1% TWEEN 20 in PBS. The membrane was developed by enhanced chemiluminesence (ECL). The membrane was incubated for 1 min with a mix of 1 ml of reagent 1 (reagent 1: -1% (w/v) luminol 0.44% (w/v) coumaric acid, 10% Tris-HCl pH 8.7) and 1 ml of reagent 2 (Reagent 2: -0.64% 30% H_2O_2 10 % Tris-HCl pH 8.5). The membrane was then exposed to X-ray sensitive film for 30 s, 1min and 5 min. The film was them processed using a Kodak M35 X-OMAT Processor.

2.11 In-Gel Trypsin Digestion of Proteins.

This technique was used to sequence protein bands from SDS PAGE gels and to prepare samples for mass spectrometry.

Protein bands were excised from SDS polyacrylamide gels and washed in 500µl of 100mM NH_4HCO_3 for 1 h,and then incubated in 150 µl of 100mM NH_4HCO_3, 10µl 45mM DTT at 60 °C. The sample was cooled to room temperature and 10µl of 100mM iodacetamide added. Incubation continued at

room temperature in the dark for 30 min. 500µl of 50:50 acetonitrile/100mM NH4CO3 was added to the gel slices and shaken for 1 h before the wash was discarded and the band sliced into 3 pieces. Gel slices were shrunk in 50 µl acetonitrile for 15 min, the solvent removed and gel slices dried for 30 min in a rotary evaporator at 30 °C.

Gel slices were swollen, 15µl of 25mM NH_4CO_3 containing 1µg modified Trypsin (Promega) for 15 min. A further 20µl of buffer was added and the incubation continued overnight at 37°C. The supernatant was removed and the gel slices washed twice with 50µl 10% (v/v) acetonitrile/0.1% (v/v) TFA. The supernatant and washings were combined and rotary evaporated to near dryness before resuspension in 0.1% TFA.

2.12 Peptide Synthesis

Throughout the study peptides were synthesised by Dr. G. Kemp at the University of St. Andrews peptide synthesis unit essentially described by (Atherton and Sheppard, 1988). The peptides included GVQSLKRRRCF, GVQSLKRRRCA, GVQSLKRRRCY, GVQSLKRARCF, GVQSLKRARCA, GVQSLLKRRRCA, GVRYGCQRYCY. LSGAGFSW were synthesised as described by Webster *et al* (1993) using Fmoc-polyamide chemistry (Atherton *et al*; 1988). Peptides were purified by reversed-phase HPLC using a C18 column pre-equilibrated with 0.1% v/v TFA and eluted by increasing concentrations of acetonitrile.

2.13 Mass Spectrometry

Mass spectrometry is a very sensitive technique capable of measuring the molecular weight of chemical compounds and biopolymers. Matrix-assisted laser desorption/ionisation (MALDI) permits the analysis of high molecular weight compounds by allowing the transfer of the sample from a condensed phase to the gas phase. MALDI analysis requires the sample to be presented

in a solid matrix which is ionisable. The matrix absorbs laser light energy causing the matrix material to vaporise carrying the sample with it. The charged molecules are then directed by an electrostatic lenses from the ionisation source to the mass analyser. Uncharged molecules will often interact with the matrix undergoing ionisation and become charged.

Once the molecules are vaporised they are separated in this instance by time of flight analysis. Time of flight analysis is based on the acceleration of a set of ions directed at the detector with the same amount of energy. Because the ions have the same energy but different mass they reach the detector at different times i.e. smaller ions first.

2.14 Peptide and Protein Sequencing

Peptide and protein sequencing was carried out at the St. Andrews University protein sequencing unit. The band corresponding to the protein of interest was excised and sequenced by Edman degradation (Han *et al.*, 1985). The process is comprised of three steps. Coupling of phenylisothiocyanate (PITC) to the N-terminal amino-acid, acidic cleavage of the derivatised amino acid and conversion of the unstable thiazolone complex to a more stable phenylthiohydantoin.

Protein samples were purified on 15% SDS PAGE gels and electroblotted to sequencing membranes in 10mM CAPS, 20% methanol pH11. Protein bands were stained with 0.1% amido black, excised and sent for sequencing.

Peptides for sequencing were reverse HPLC purified and 5 µl aliquots in 0.1% TFA containing 20-50 pmoles were submitted for sequencing.

2.15 Determination of Peptide Concentration

2.15.1 Fluorescamine Assay

Concentration of peptides was determined using fluorescamine assay as described previously (Cabrita, 1997). Fluorescamine (Sigma) reacts with primary amines and following excitation at 390nm a fluorescent emission at 475nm was detected. Peptide concentration was determined by comparison to a tetraglycine standard curve.

2.15.2 Capillary Electrophoresis

Capillary Electrophoresis (CE) provides a rapid and cost effective way to separate biopolymers, metal ions and analytes in respect of their mass:charge ratio. Separation is achieved by the application of an electric field across the capillary. Analytes are detected as they pass a window at the end of the capillary. Analytes enter the capillary tube in a low ionic strength buffer, which sharpens the peaks as the analyte migration slows when entering the electrolyte buffer. Thus the leading edge slows, the tail catches up resulting in a more defined peak.

Samples were separated on a Biofocus 3000 Capillary Electrophoresis system with a detection time of 14-20 min. The running sample composed 1:1:8 of analyte sample: 0.1M phosphate buffer pH2.5 (BioRad): HPLC grade water (Rathburn). Injection pressure was 10psi with a 10kV potential difference applied across a 24cm x 25µm capillary cartridge. All software used was Bio-Rad V3.01 for windows 3.1x. Electropherograms were recorded at Absorbance 200nm and stored as Braided Format Files (.BFF).

Peak area integration was carried out using Integrator software V3.01. .BFF files were converted to ACQ. (Acquisition files) and integrated using manual base line editing of the observed peaks. Peak area and elution times were

recorded and used for the calculation of initial rates of enzyme activity as previously reported (Cabrita 1997 b) .

2.16 Modification to Peptides

2.16.1 Oxidation of Peptides

Peptides were dissolved in 0.1 M $NaHCO_3$ to a final concentration of 1mg/ml and shaken overnight at room temperature before being purified using reverse phase HPLC.

2.16.2 Reduction of Peptides

Peptides were dissolved in 0.1 M $NaHCO_3$ and a 3 molar excess of DTT was added. Nitrogen gas was bubbled through the sample for 2 min before stirring at room temperature for 45min. The pH was adjusted to 4-5 with 0.2 M acetic acid and the peptide purified using reverse phase HPLC on a C18 column.

2.17 Proteolysis Assay

The synthetic peptide substrate assay was developed by Webster *et al.* (1989a) based on the cleavage sites of proteins pVI and pVII. Digestion of LSGAGFSW produces LSGA and GFSW peptides which can be separated and quantified using capillary electrophoresis.

The reaction mixture typically comprised of 10 µl of protease (0.02-0.06mg/ml), 10 µl activating peptide (100 molar excess over protease) and 25 µl buffer (50 mM Tris-HCl, 10 mM EDTA, 2 mM 2-mercaptoethanol, pH 8) which was equilibrated to 37 ^{o}C before adding 5 µl of LSGAGFSW (2mg/ml). 10 µl aliquots were removed at 5, 10 15, 20 and 30 min after addition of substrate and the reaction stopped with an equal volume of 1% TFA. 80µl of water was added prior to analysis by Capillary Electrophoresis.

2.18 Non-reducing, Non-Denaturing Native Gel Electrophoresis

Mini gels were prepared using a Bio-Rad Gel Kit. Separating solutions were comprised 4.5 ml 40% acrylamide, 2.4ml 2% bis-acrylamide, 3.45 ml H_2O, 1.5 ml 17.4 mM potassium acetate buffer pH4. Gels were set with the addition of 10 µl TEMED and 150µl of 1mg/ml (w/v) ammonium persulphate.

The stacking gel consisted of 1 ml 40% bis acrylamide, 0.56 ml 2% bis-acrylamide, 5.2 ml H_2O, 1. ml of 17.3 mM potassium acetate buffer pH4. Gels were set with the addition of 10 µl TEMED and 200 µl of 1mg/ml (w/v) ammonium persulphate.

2.19 Crystallisation Trail

Crystallisation's trials were initiated using the hanging drop vapour diffusion technique. Purified Adenovirus C104A protease was concentrated to 5.6 $mgml^{-1}$ (McGrath, 1996) on Centricon 10 K cut off membranes. 1µl protease was mixed with the mother liquor on a cover slips and sealed in chambers (Nuncon 24 well plates) containing 1 ml of mother liquor. Magic 50 and Magic 48 kits (Hampton Research, H2-110 and H2-112 respectively) were used for the screening. A second set of conditions using 1.2-1.8M sodium acetate as the precipitant and 100mM HEPES pH 7.5-8.5 were initiated (McGrath et al, 1996).

2.20 Protein-Protein Interactions

2.20.1 Use of GST-Protease Fusion Proteins

This technique used Adenovirus type 2 protease fused to GST as a bait for protein-proteins interaction. pGEX23K is a GST protease fusion vector (N-terminal GST) under the control of *tryp/lac* promoter. Expression is induced as with the pET system by the addition of IPTG. Fusion proteins are readily purified from bacterial cell lysate by affinity chromatography over glutathione-Sepharose resin, making use of the ability of GST to bind the

tripeptide glutathione. This binding is reversible in the presence of reduced glutathione which serves to remove proteins from the column. The system can also be used to purify proteins making use of the thrombin cleavage site located at the C-terminus of GST.

In outline, HeLa cell extracts were incubated with glutathione-Sepharose beads loaded with GST-protease fusion. The beads were then washed and adhering proteins eluted prior to identification by sequencing or mass spectrometry.

GST-fusion pellets were thawed and resuspended in 20ml PBS/0.5M NaCl, 2mm EDTA, 1mM HCl and 2Mm Benzamidin, before being sonicated 6 times for 30 s with 30 s incubations on ice in between. Triton X100 was added to a final concentration of 1%. Insoluble particulate matter was collected at 17000 rpm (Beckman JA 17) and the supernatant retained.

Glutathione beads were prewashed with two column volumes of 0.2 M NaOH, and 6 column volumes of PBS/0.5 M NaCl and incubated overnight at 4°C with either fusion protein supernatant or pure GST (kindly donated by E. Jaffray).

These beads were then incubated a 4°C with the HeLa cell protein extract. After 16 h the beads were washed sequentially with 6 column volumes PBS / 50mM Tris-HCl, 0.5 M NaCl, 10 mM NaOH pH 8. Proteins were eluted from the beads with 1ml of elution buffer PBS / 50mM Tris-HCl, 0.5 M NaCl, 10 mM NaOH, 10mM glutathione pH 8. These proteins were concentrated by precipitation in TCA (Trichloroacetic acid) as per method 2.20.3.

2.20.2 HeLa Cell Extract Preparation

100 litres of HeLa S3 spinner cells were harvested at 4000rpm (J6-HC). The resulting pellet was washed twice in PBS and resuspended in 4 times pellet

volume of GST incubation buffer (PBS containing 5mM NaF, 20mM Tris-HCl , 1mM Na orthovandate, 1mM EDTA with 1mM of each of TLCK, Pepstatin, Leupeptin, Bestatin and 100mM Pefablock). Cells were sonicated for 6 x 30s periods and incubated on ice in-between. Particulate matter was collected at 14000 rpm (Beckman JA-14). Supernatant was then ultracentrifuged for 45 min at 40000rpm (Beckman L-60 centrifuge, 42.1 rotor), filtered (0.2m) and applied to the Sepharose beads.

2.20.3 Trichloroacetic Acid Precipitation

TCA was added to eluted bead samples to a final concentration 10% v/v. Samples were centrifuged at 25000 rpm (Beckman TL-100 centrifuge, TLA 100.2 rotor) for 1 h at 4°C. The resulting pellet was washed twice with acetone and centrifuged as above. The pellet was dried and resuspended in 7µl of water and 3µl of boiling mix. Samples were analysed by 15% SDS PAGE electrophoresis.

2.21 Eukaryotic Expression Systems

2.21.1 Maintenance of Eukaryotic Monolayers.

All Eukaryotic cells were maintained at 37°C in a humid 5% CO_2 atmosphere in 75 cm^3 flat bottom flasks. Cells were passaged at 80% confluence. Media was removed and monolayers washed with 10ml PBS before incubation with 2ml trypsin for 5 min. Trypsinisation was terminated with 10 ml of D-MEM supplemented with 10% Foetal Calf Serum (FCS). Cells were centifuged for 3 min at 1000 rpm and resuspended in D-MEM 10% FCS.

2.21.2 Transfection of HeLa Cells.

Throughout the study two methods of transfection were used, electroporation in 4mm cuvettes (EQUIBIO) for transient transfection studies

and LipofectACE (Gibco) to establish stable integrated DNA transfections. Transient transfections rely on episomal expression whereas with stable transfections DNA integrated into the host chromosome.

2.21.2.1 Transient Transfections

For electroporation 80% confluent monolayers were incubated in 10% FCS enriched D-MEM medium 1 h before trypsination. Cells were resuspended in 10%FCS/D-MEM 15mM HEPES and centrifuged for 5 min at 1000 rpm. Supernatant was discarded and cells resuspended in 15mM HEPES/D-MEM at a cell density of $5x10^6$ per ml.

50µl of 210mM NaCl containing 8-40µg of plasmid DNA was mixed with $2x10^5$ cells and electroporated (GenePulser II Bio-Rad) 950 mFD, 240V for 35ms. Cells were resuspended in 5ml of 15 mM HEPES D-MEM and centrifuged at 1000 rpm for 5 min, resuspended in 4 ml D-MEM and incubated on ethanol sterilised coverslips for a minimum of 16 h. Duplicate plates were prepared in 20mm diameter plates (NUNCLON) for Western blot analysis.

2.21.2.2 Stable Transfections

LipofectACE is a mixture of lipids which fuse with the extracellular membrane facilitating the delivery of DNA in to the target cell. 100ng of DNA was added to 200µl of OPTIMEM (Gibco) and mixed using a vortex mixer. 10 µl of LipofectACE and 190 µl OPTIMEM were mixed by gentle pipetting action before both were mixed together and incubated at room temperature for 15min The transfection mixture was added to 50% confluent monolayer in 6 well plates along with 1.6ml of OPTIMEM and incubated at 37°C/5%CO_2. OPTIMEM was replaced by a modified MEM and incubated as before. Selective antibiotics were added (CLONTECH Manual, PT30001-1).

2.21.3 Immunocytochemistry

Slides were washed with PBS containing 0.5mM $CaCl_2$ 1mM $MgCl_2$ and fixed with 3% paraformaldehyde. After washing three time with PBS, slides were incubated 2 x 15 min with 0.1M Glycine PBS and permeablised with 0.1% Triton X 100 PBS. Slides were blocked with PBS containing 0.2% BSA and primary antibody in PBS/BSA at 1/200 dilution was added for 30 min. After removal of primary antibody and washing with PBS/BSA, cells were incubated with anti-mouse FITC conjugated antibody and DAPI at 1/200 dilution. Cells were viewed at 100x magnification under oil immersion with a Nikon Microphot-FXA microscope.

3. Results

3.1 Introduction

The main objective of the work detailed here was to try and develop our understanding of the activation mechanism of the protease and interpret the results in the context of developing an antiviral agent. To achieve this several characteristics of the activating peptide needed to be elucidated such as the importance of the KRRR motif, the constraints upon the length of the activating peptide and the importance of the C- terminal residue. Further insight might be gained by investigating the interaction of the activating peptide from the EDS virus (Egg drop syndrome). The sequence of the EDS activating peptide is the one most different from that of Adenovirus type 2.

The chapter starts with a general overview of the techniques used to express and purify the protease and moves on to try and assess the relevant properties of the C-terminal residue, the effect of the EDS activating peptide on Ad2 protease, the reported activation by VEGGS peptide with Wt type and the mutant C104A proteases, and finally the role of the conserved KRRR motif in activation. We then describe the development of a binding affinity assay to assess the protease/activating peptide complex formation.

We then look at the general effects of protease expression *in vivo* in a series of studies designed to determine whether or not the KRRR motif, which is conserved throughout human serotypes of the Adenovirus activating peptide, is capable of conferring nuclear localisation. Finally an attempt to develop a controllable, inducible nuclear localisation expression system is described.

3.2 Preparation of Adenovirus Protease from Bacterial Cultures

Wild type (Anderson, 1990) and the mutant protease C104A (Jones *et al.*, 1996) had previously be cloned into pET11C. 10 ng of each of these plasmids were

transformed into BL21 (DE3) using the heat shock method. Because of the variable expression between batches of protease 12 colonies from each plate were selected, grown at 37°C in LB/antibiotic for 2h and protease expression induced for 2h with 1 mM IPTG. Figure 3.01 shows Wt protease samples separated by 15% SDS PAGE and examined on PVDF membrane by the Western blot technique. Bands corresponding to protease expression were resolved with R11 polyclonal antisera an antipeptide antibody raised against the N-terminal 17 residues of the protease, and the colony giving the most intense band was used to prepare a glycerol stock for all subsequent protease preparations. In the case of the Wt protease colony 12 was selected.

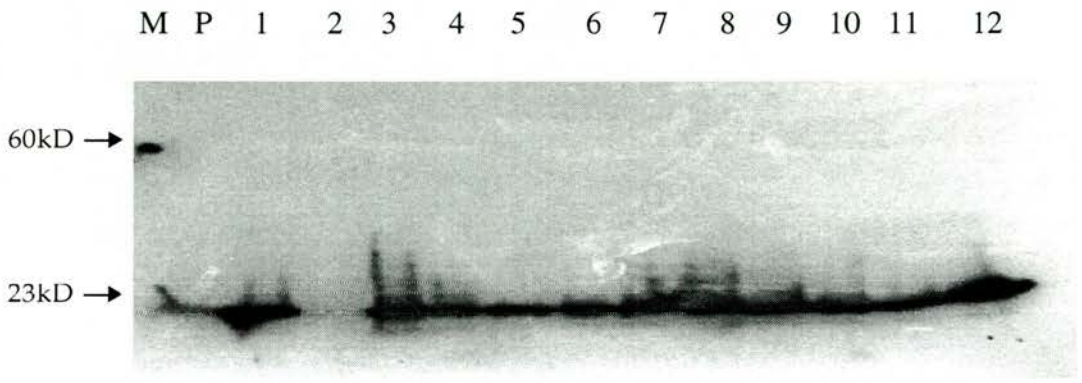


Fig 3.01 Expression test. 12 colonies were selected from a transformation plate grown for 2 h in L-Broth plus antibiotic. Protease expression was induced with 1mM IPTG for 2 h. Cells were harvested, sonicated and separated using 15% SDS PAGE. Gels were electrotransferred and detected with polyclonal antisera R11. Lane -1 pre-stained molecular weight marker with HSA, lane 2, a protease standard marker, and lanes 3-14 colonies 1-12. There was no detectable protease expression from colony 2, colony 1 and 12 gave the best expression and 12 was used for preparations of all future wild type protease expressions. Similar expression studies were carried out for C104A (results not shown).

3.3 Expression of the Protease

Recombinant Ad 2 protease was originally purified from disrupted virus particles (Trembley 1983), the authors reported that there was enough protease for 200 assays and that the protease remained active for over 1 year. This process was time consuming and required work with the live virus. In 1990 the Wt protease gene was cloned into pET 11C (Anderson, 1990) nd over expressed in *E. coli.*. The protease was recovered from the bacterial cell pellet and purified for use in activation studies. This process took much less

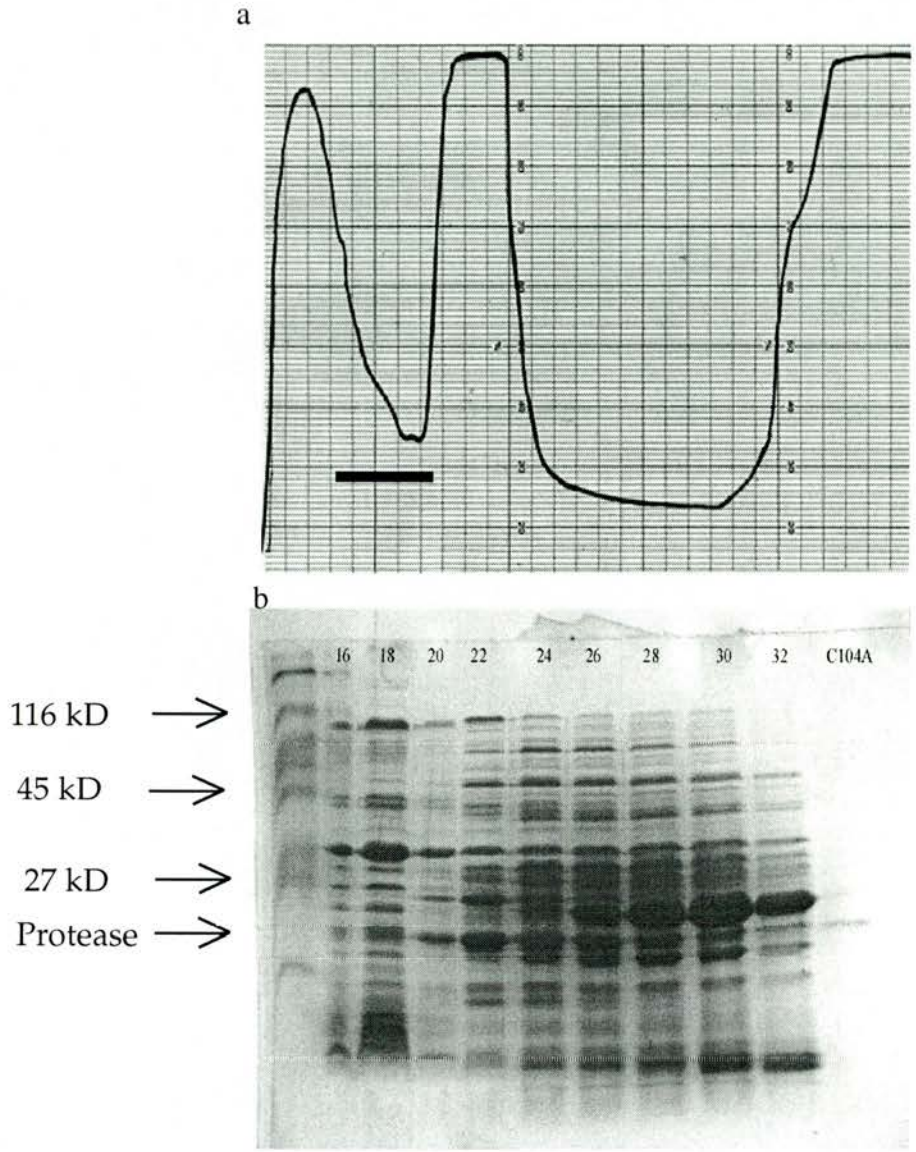


Figure 3.02 DEAE Purification. Part a shows the elution profile of proteins from a 500ml culture of BL21(DE3) protease expression in D-MEM. Protein was detected at 280 nm, chart recorder speed 0.25cm/min, absorbance sensitivity 2. Protein extract flowed through the DEAE column at 1ml/min. The black bar represents the region corresponding to protease elution. Part b, 15% SDS PAGE gel (resolved with Coomassie brilliant blue) of the even numbered fractions 16-32 corresponding to the bar in part a. The protease band was first seen in fraction 18 and was eluted by fraction 30. Fractions 18-30 were pooled and passed through consecutive heparin-Sepharose and CM-Sepharose columns.

time than the live virus method and had the advantage of giving a greater yield of protease.

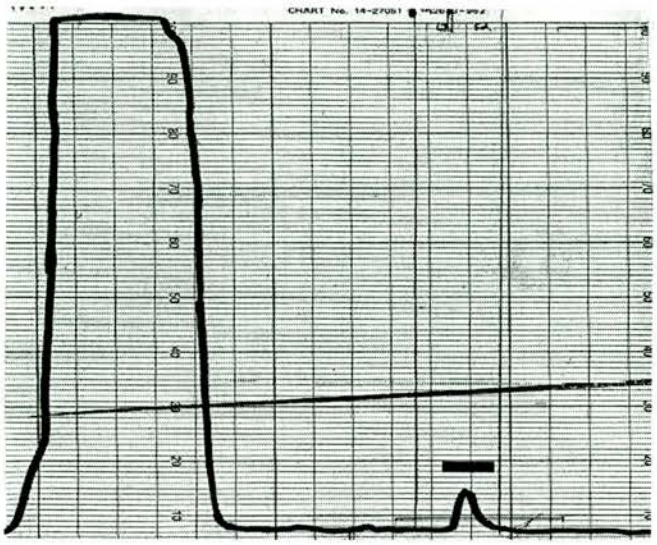
In this study the expression of the protease was tracked by taking 1 ml samples at hourly intervals postinduction for 5 h. The samples were prepared for electrophoresis and examined by Western blotting to assay the optimum expression conditions (data not shown). It was found that protease expression peaked approximately 3 h post-induction and that there was no advantage gained in leaving the induction to progress any longer. At this point the cells were harvested and the pellet stored at -20°C.

3.4 Protease Purification

The first step of purification was separation on a DEAE-Sepharose column pre-equilibrated with 50mM Tris-HCl pH 8. A typical profile of the elution from a DEAE column is shown in figure 3.02a. The protease was eluted from the DEAE column consistently from fraction 18 through to fraction 30 which is indicated in figure 3.02a by a thick black line. From lane 22 to 32 a band of approximately 30 kD appeared and was identified as β-lactamase. This enzyme was expected to be produced in such a manner in response to the selective antibiotic and was a useful control for sequencing and mass spectrometry of bacterial extracts

Fractions 18-30 were pooled and separated on consecutive heparin-Sepharose and CM-Sepharose columns pre-equilibrated with 50 mM Tris-HCl pH8. The elution profile is shown in figure 3.03a. Proteins eluted in the first peak represent those which did not interact with the CM and heparin columns at the operating pH and buffering conditions. The protease at pH 8 carries a net positive charge which interacts with the negative charge carried by the CM moiety. Initial studies illustrated that the protease did not interact with the Heparin column, however, other contaminating proteins did (results not

a



b

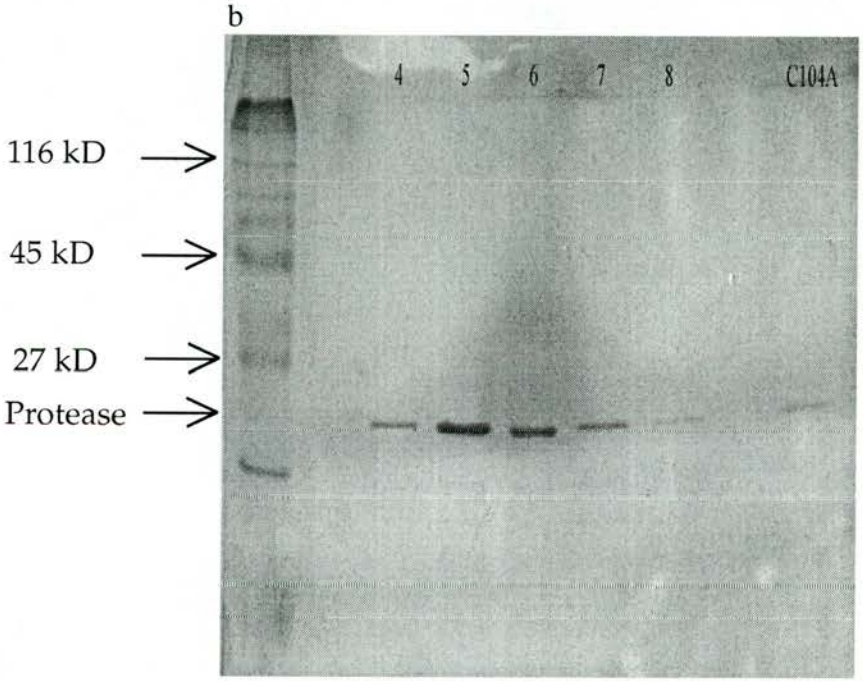


Figure 3.03 Part CM/Heparin Purification. Part a shows the elution profile at 280nm for the passage of pooled fractions 18-30 from figure 3.02 b. Eluate from the DEAE separation was passed through the column in 50mM Tris-HCl pH 8 at 1ml/min. The protease bound the CM column and was eluted with 150mM NaCl, 50mM Tris-HCl pH8. Part b is a 15% SDS PAGE gel of fractions 4-8 (from the CM column) corresponding to the black bar in part a. The protease band was first seen in fraction 4. Bands were protein sequenced to confirm protease presence. Protein concentration was then determined using Bradfords assay and SBTI standards. The protease was stored at -70°C.

shown). As a consequence, protease elution buffer (50 mM Tris-HCl 100-150mM NaCl pH8) was only applied to the CM column to minimise any possible contaminants which could have been eluted from the heparin column. Figure 3.03b shows a Coomassie stained 15% SDS PAGE, fractions 5, 6 and 7 (corresponding to the second peak in figure 3.03a) were sequenced and confirmed to be Adenovirus protease. The eluates were pooled, aliquoted and stored at -70°C. Protease concentration was determined using SBTI standards on SDS PAGE or by Bradfords reagent assay. From 1 litre of culture 1-1.5 mg of protease was recovered. Protease was diluted to 0.02 mg /ml for the subsequent protease activity assays, or concentrated onto Centricon 10 k cut-off membranes to 5.6mg/ml for crystallography trials.

3.5 Development of a Single Step Purification Method

Originally it had been necessary to check all eluates from the DEAE and CM column to identify fractions containing the protease. Identification procedures involve SDS PAGE and Western blot techniques which were time consuming and consequently affected the quality of the protease. Protease was found to consistently elute from the DEAE and CM column between fractions 18-30 and 4-8 respectively. The following method describes a new procedure to purify the protease in a single step and is summarised schematically in figure 3.04. The advantage being that single band purity on SDS PAGE gel resolved with Coomassie took approximately 2 h.

Bacterial cell extracts were prepared as previously described. The extract was applied to a 50mM Tris-HCl pH8 equilibrated DEAE-Sepharose column (30 x 1 cm) for 18 minutes at a flow rate of 1 ml/min and the eluate was discarded. At 18 minutes the flow from the DEAE-Sepharose column was diverted to consecutive heparin-Sepharose and CM-Sepharose columns maintaining a flow rate of 1 ml/min for 12 min. After 12 minutes the DEAE-Sepharose

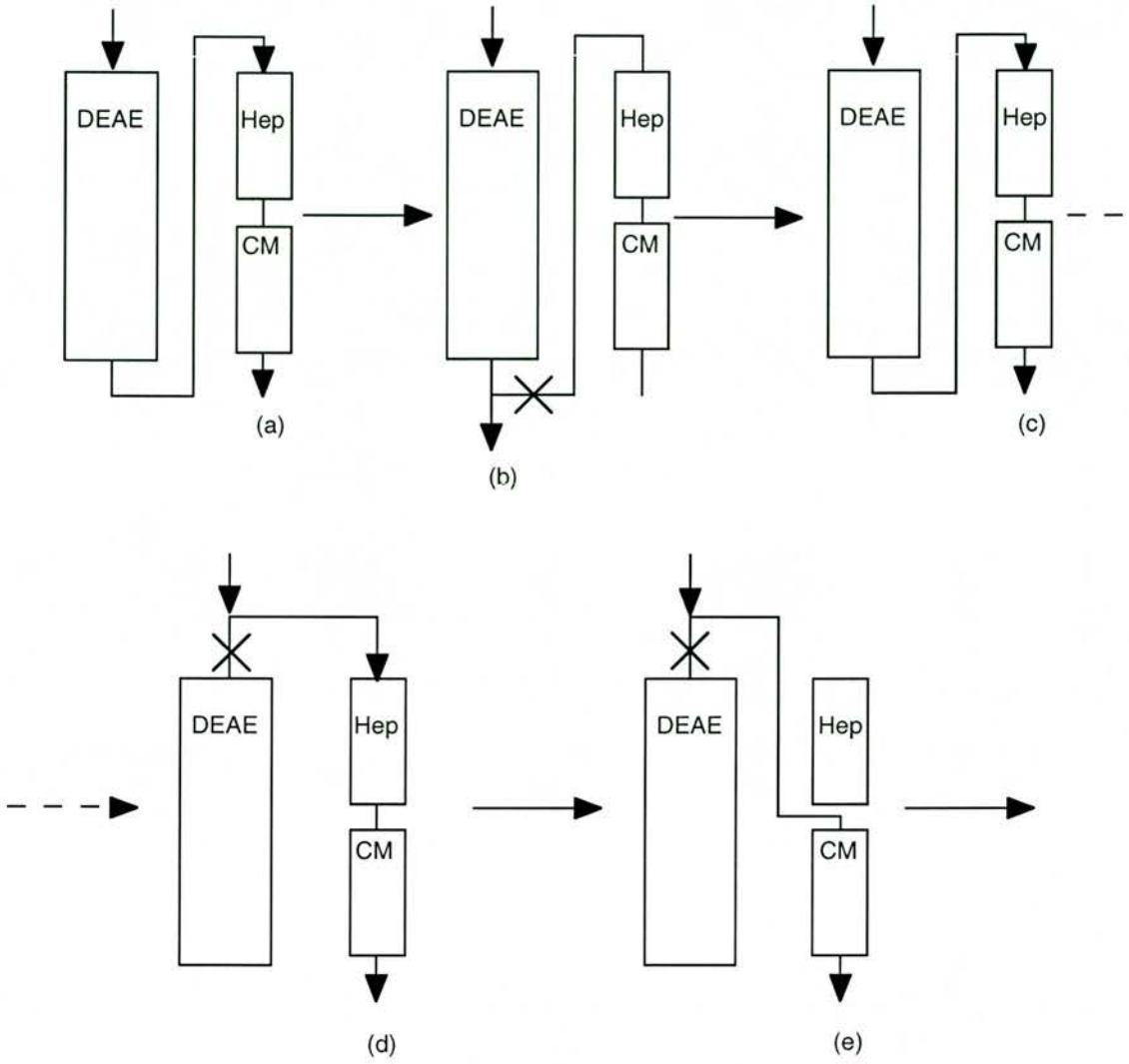


Figure 3.04 Single Step Purification. Schematic representation of the single step purification method: (a) shows DEAE, heparin and CM column equilibration with 50mM Tris-HC1 pH 8 for 30 min at a flow rate of 1ml/min; (b) - the loading of bacterial cell extract on to DEAE column for 18 minutes. (c) - eluate from DEAE is loaded on to consectutive heparin-Sepharose and CM-Sepharose columns for 12 minutes; (d) - buffer by-passes the DEAE column and enters only the heparin and CM column to remove unbound proteins; (e) - protease eluted from the CM column only, with 50mM Tris-HCl 100-150mM NaCl pH 8.

column was diverted and 50mM Tris-HCl pH8 continued to pass through the heparin-Sepharose and CM-Sepharose columns for 30 minutes (or until the absorbance returned to baseline values).

Elution buffer (50mM Tris-HCl 100-150mM NaCl pH8) was added to the CM-Sepharose column for 10 minutes at a flow rate of 1ml/min. Ten 1 ml fractions were collected and prepared for SDS PAGE and Western blot analysis to identify the fractions containing protease, and the concentration determined using Bradfords assay. 50µl aliquots were stored at -70°C. This method was employed for both Wt protease and the mutant C104A protease purification.

3.5.1 Yield from Purification

Unknown concentrations of Wt and C104A were compared to SBTI standards to assess concentration of protein using the Bradfords assay. Typical protease yield for the L-Broth method was 0.75 mg and that of D-MEM was 1.5 mg per litre of cell culture.

3.6 Activity Studies

Recruitment of an activating peptide to achieve maximal cleavage potential by the Adenovirus protease represents a novel mechanism of controlling the activation of a proteolytic enzyme. The fact that the activating peptide conforms in most part to a consensus of GxxxxKRRRC(F/Y) offers an intriguing insight into the control of activation. Here we investigated the role of the 4 conserved basic amino acids KRRR, the C-terminal aromatic residue, the effect of extending the peptide length, the effect of the newly reported activation ligand VEGGS on Wt and C104A protease, and serospecificity of the activating peptide.

In a typical assay 10µl of protease (0.02mg/ml) was incubated with 10µl of oxidised activating peptide (200mM) and 25 µl of 1mM β-MeOH (2-

Mercaptoethanol), 10mM EDTA, 50mM Tris -HCl pH 8 for 10 minutes at 37°C. 5µl of substrate (LSGAGFWS) was added and samples taken at 5 minutes intervals to 20 minutes followed by a final time point at 30 min. Digestion assays were terminated with the addition of an equal volume of 1% TFA.

Initial rates were calculated from integrated peak areas derived from substrate digestion assays separated by capillary electrophoresis for 10 min using peaks corresponding to undigested substrate, LSGA and GFSW. The peak areas were estimated using Biofocus integration package version 3 (fig. 3.05). To account for differences in the activity of the protease from batch to batch, initial rates of digestion of the test peptides were expressed as a percentage of that achieved by a Wt activating peptide in a duplicate control assay.

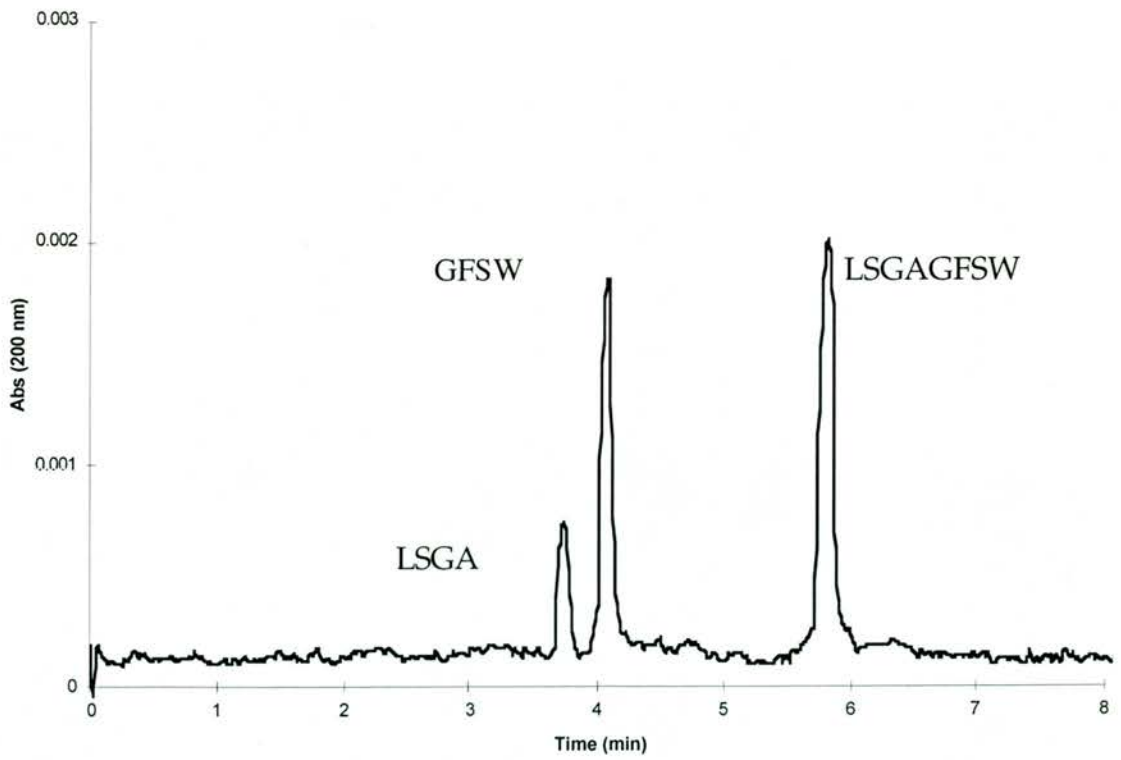


Figure 3.05 CE electropherogram. of a 30 min protease digestion. The protease was activated with Wt activating peptide for 10 minutes at 37°C before substrate was added. Samples were taken at 5, 10, 15,10 and 30 minutes and separated by capillary electrophoresis.

3.6.1 Sample Initial Rate Calculation

The digestion of synthetic substrate represented an irreversible pseudo first order reaction where the rate (v) is described as the change in substrate concentration (ds) within a certain time interval (dt), expressed mathematically as :-

$$v = -\frac{ds}{dt}$$

Since the rate (v) for an irreversible first order reaction is proportional to substrate concentration :-

$$v = ks$$

k is the known as the rate constant and is independent of the substrate concentration. The initial rate will be the product of the rate constant and the initial concentration of substrate (So), therefore to calculate k the above two equations can be combined to give:-

$$-\frac{ds}{dt} = ks$$

Integration of this will give the equation for a straight line graph of the form y = Ax, where y = ln (S/So) against time with a straight line -k, the initial rate. k was calculated using the array formula below in Microsoft Excel.

	A	B	C
	Time	S/So	ln(S/So)
1	0	1	0
2	5	0.9826	-0.017553159
3	10	0.9688	-0.031697087
4	15	0.9511	-0.050136069
5	20	0.9428	-0.058901108
6	30	0.9126	-0.09145761

The array formula {=linest(C2:C7,A2:A7,FALSE,TRUE)} was applied to calculate the initial rate of enzyme activity and the following output .

	D	E
2	Slope (-k)	-0.003077896
3	Standard error	0.0000598
4	Correlation coeff.	0.994473096
5	F statistic	899.6656329
6	Regression sum of squares	0.005235751

The example shows an initial rate of 0.3% substrate digestion per minute. Throughout the study a correlation coefficient of 0.90 or greater was acceptable, figures below this value were discarded from the data set.

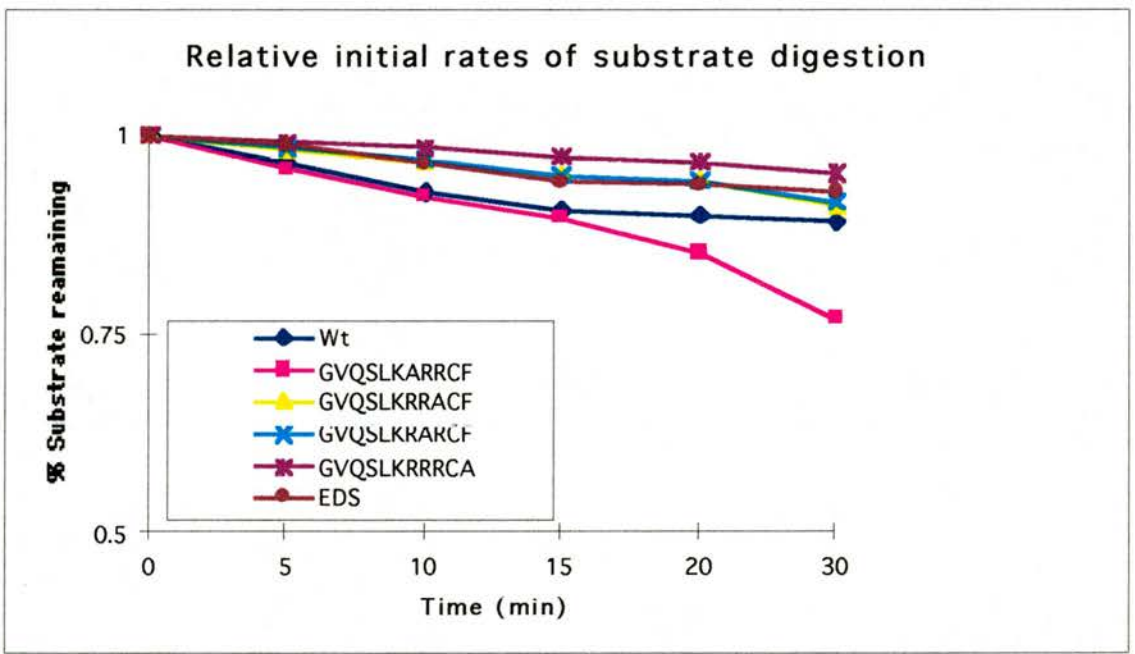


Figure 3.06. Initial rate calculations were performed for each peptide and expressed as a per cent of that achieve by the wild type peptide in duplicate assays. The figure shows a representative sample of initial rates of digestion for each of the active peptides. No activity was detected with GVQSLLKRRRCF and VEGGS. The slope of the line was calculated using the array formula {=linest(C2:C7,A2:A7,FALSE,TRUE)} and represented the initial rate of enzyme digestion.

Duplicate assays activated with Wt activating peptide were carried out at the same time. To calculate the relative initial rate, reactions with Wt activating peptide were assumed to be 100% active and the mutant activating peptide activity calculated as a function of that. Figure 3.06 illustrates a representative initial rate for each of the active peptides. Rates calculated from the array formula were expressed as a per cent Wt peptide in duplicate assays. This

method allowed the comparison of assays with different peptides at different times with different batches of protease to be comparable.

3.6.2 Effect of Extended Activating Peptide

To investigate the effect of extending the activating peptide the mutant peptide GVQSLLKRRRCF was synthesised by Fmoc chemistry and purified using reverse phase HPLC. This peptide was used to assess the constraints imposed by the binding groove to the incoming activating peptide. GVQSLLKRRRCF peptide is 1 amino acid longer than conventional activating peptide (analogous to 10 mer peptides binding MHC class 1 receptors). In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate GVQSLLKRRRCF, and a control activated with GVQSLKRRRCF. Samples were taken at 5, 10, 15, 20, and 30 min. Protease activity was not detected with this form of the peptide.

3.6.3 Effect of C-terminus Residue of the Activating Peptide

Figure 4.03 from the Discussion shows the wild type activating peptide sequences published to date. One important observation was the conservation of an aromatic amino acid at the C-terminal of the activating peptide. Morgan and McAdon (1980) implicated pi orbitals in the stabilisation of protein tertiary structure involving cysteines and methionine residues. To assess the importance of the aromatic side chains in activation, the peptide GVQSLKRRRCA was synthesised by Fmoc chemistry and purified using reverse phase HPLC. In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate GVQSLKRRRCA, and a control activated with GVQSLKRRRCF. Samples were taken at 5, 10, 15, 20 and 30 min and digested substrate peptide fragments separated by capillary electrophoresis. The resultant peak areas

were integrated and the RCA peptide was found to activate the protease to 0.41 ±0.08 the activity Wt activating peptide.

3.6.4 Effect of VEGGS

It had been reported that the peptide VEGGS from a phage library display experiment was sufficient to activate the protease to comparable level of that achieved by the Wt activating peptide (Diouri *et al.*, 1996). The authors concluded that the activation was at a site other than that of the conventional activating peptide. To confirm this observation VEGGS was synthesised using Fmoc chemistry and HPLC purified. In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate peptide acetyl-VEGGS-amine peptide, and a control activated with GVQSLKRRRCF. Samples were taken 5, 10, 15, 20 and 30 min. No activity was recorded during this time course. The assays were repeated increasing the incubation length to 16 h and again no activity was found.

If the sequence acetyl-VEGGS-amide did activate the protease at a site distinct from that of the conventional activating peptide one would expect that a mutant lacking C at 104 would still be activatable. The above assay was repeated using C104A mutant protease, samples were taken at 5, 10, 15, 20, 30 min and 16 h. No activity was observed.

3.6.5 Effect of the EDS Activating Peptide

To see if activating peptides from another serotypes were able to cross activate Adenovirus type 2 protease GVRYGSQRYCY responsible for the activation of the Avian 127 protease was synthesised by Fmoc chemistry and purified by reverse phase HPLC. EDS activating peptide represented the most evolutionary distant relation to the Ad2 activating peptide. In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate peptides GVRYGSQRYCY and

GVQSLKRRRCF. Samples were taken at 5, 10, 15, 20 and 30 min and plotted against substrate digestion. The initial rate of activity with GVRYGSQRYCY peptide was compared to the initial rate attained with Wt activating peptide was 0.42 ± 0.04.

3.6.6 Effect of the Conserved Basic Amino Acids

It was apparent from sequences of activating peptide so far identified that the mammalian serotypes conformed to a consensus sequence of GxxxxKRRRC(Y/F). The fact that 4 basic amino acids are conserved amongst most of the *Mastadenovirus* was an intriguing concept. Were they important in the activation mechanism, involved in the stabilisation of the disulphide bond formation, or were they important in the regulation of protease activity in a compartmentalised fashion? That is to say does the activating peptide acting as a signal to confer nuclear localisation of the protease or pVI? Mutant peptides were synthesised by Fmoc chemistry and purified using reverse phase HPLC with each basic residue being substituted by alanine.

3.6.6.1 Peptide GVQSLKARRCF

In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate peptides GVQSLKARRCF, and GVQSLKRRRCF as a control. Samples were taken at 5, 10, 15, 20 and 30 min and separated as before by capillary electrophoresis. Substrate digestion was plotted against time and the initial rate calculated. The initial rate as compared to that of the protease activated by GVQSLKRRRCF was 1.70±0.36.

3.6.6.2 Peptide GVQSLKRARCF

In a typical assay three substrate digestions were followed with time, where the protease (0.02mg/ml) was activated by duplicate peptides GVQSLKRARCF, and GVQSLKRRRCF as a control. Samples were taken at 5,

10, 15, 20 and 30 min and separated as before by capillary electrophoresis. Substrate digestion was plotted against time and the initial rate calculated. The initial rate as compared to that of the protease activated by GVQSLKRRRCF was 0.69± 0.07.

3.6.6.3 Peptide GVQSLKRRACF

A typical assay comprised of three substrate digestions being followed with time, where the protease (0.02mg/ml) was activated in duplicate assays with the peptides GVQSLKRRACF, and GVQSLKRRRCF as a control. Samples were taken at 5, 10, 15, 20 and 30 min and separated as before by capillary electrophoresis. Substrate digestion was plotted against time and the initial rate calculated. The initial rate as compared to that of the protease activated by GVQSLKRRACF was 0.71±0.28.

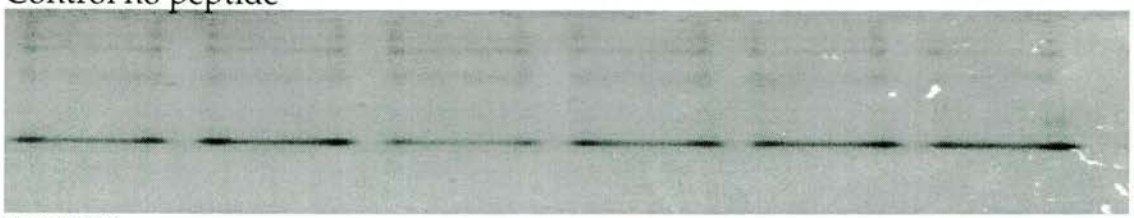
In summary the most dramatic effect was the 70% increase in activity observed by the interaction of the protease with GVQSLKARRCF peptide. Substitution of either of R8 or R9 to A resulted in a maximum of 31% decrease in substrate digestion.

3.7 Investigation into the Relative Binding Affinities of Activating Peptides and the Protease

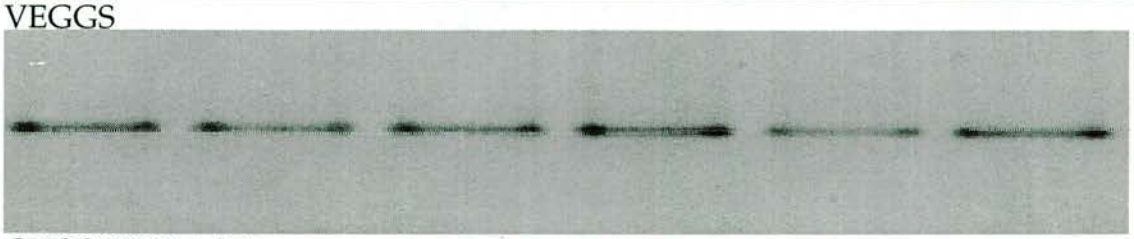
The reason for GVQSLLKRRRCF's failure to activate the protease was unknown. Did the peptide fail to interact with the protease or had the peptide bound to the protease and strapped it in to an inactive conformation? To understand this we developed a rapid qualitative assay which with further refinements could become quantitative.

The assay separates protease/activating peptide complexes based upon their mass to charge ratio. Serial dilutions of each of the above peptides from 250μM-19μM were incubated with protease for 10 min in the absence of

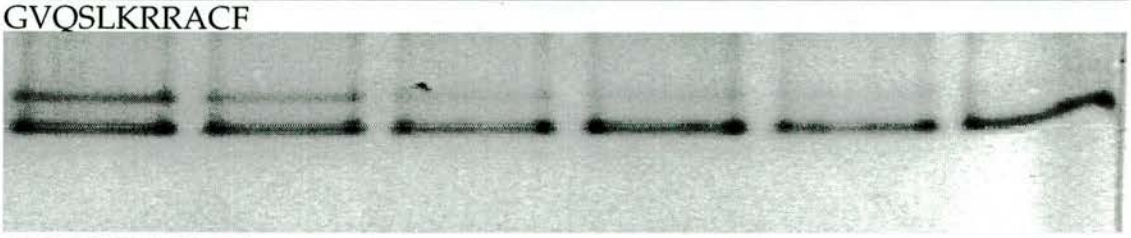
Control no peptide



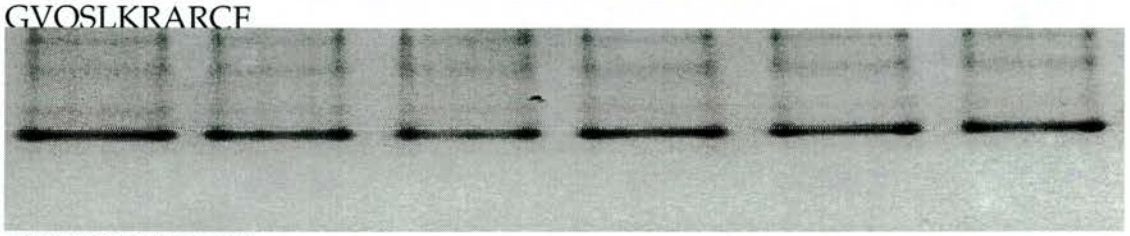
VEGGS



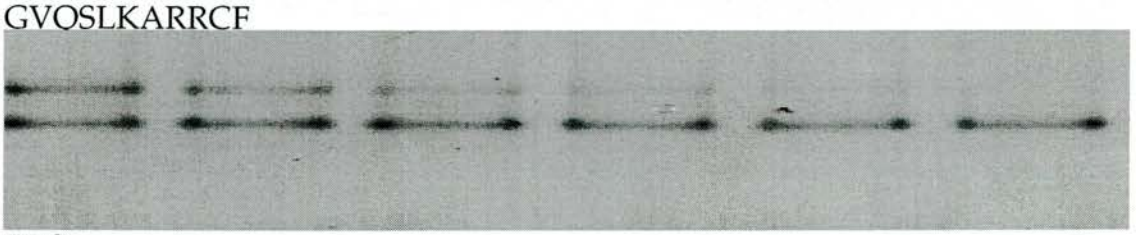
GVQSLKRRACF



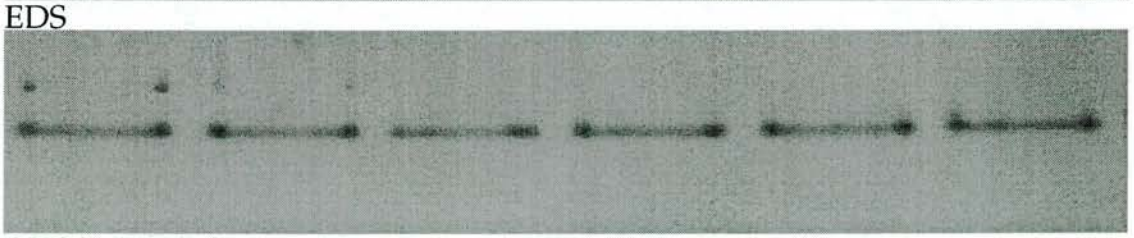
GVQSLKRARCF



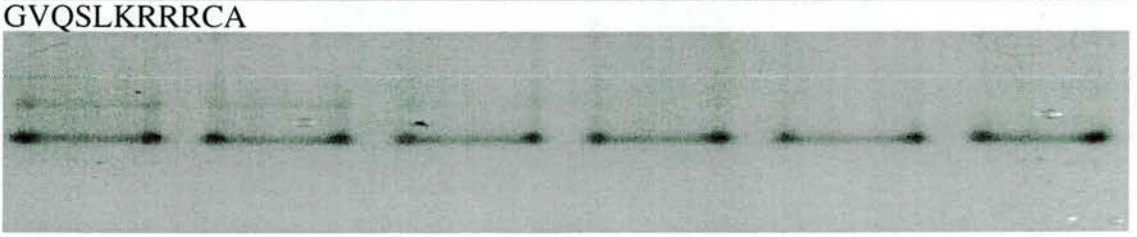
GVQSLKARRCF



EDS



GVQSLKRRRCA



Wt Activating peptide

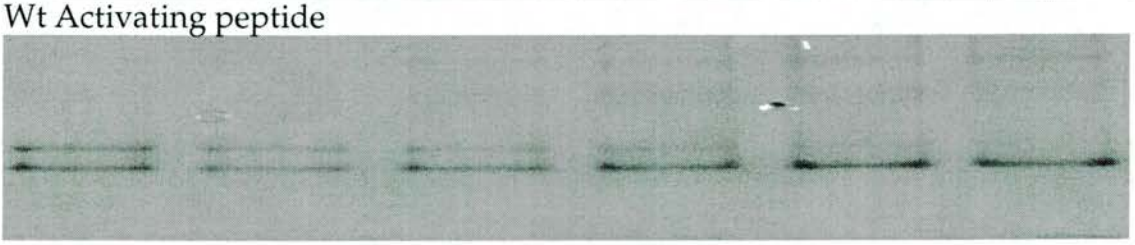


Fig 3.07 Relative Binding Affinity Assays. The band in the control experiment represents lysozyme (0.02mgμl), the internal control. 5 μl of protease (0.02mg/ml) was incubated with 5μl oxidised form of each of the peptides shown for 10 min at 37°C before being loaded on to 15% native gels and separated for 45 min at 180V. The upper band in GVQSLKRRRCF, GVQSLKARRCF, GVQSLKRARCF, GVQSLKRRACF, GVQSLKRRRCA and Wt activating peptide represents the formation of a protease /activating peptide complex. VEGGS did not heterodimerise with the protease, nor did GVQSLLKRRRCF (data not shown). As the protease / peptide complex band intensity decreased, the smearing effect of the protease increased. Activating peptide concentration from left to right represents a doubling dilution series from 250μM of peptide to 19μM.

substrate at 37°C. Samples were separated on native gels and bands resolved with Coomassie brilliant blue.

In the absence of Wt activating peptide the protease appeared as a smear on 15% SDS PAGE gels. The smearing effect was not observed in the presence of wild type activating peptide with concentrations ranging from 250μM-19mM. Here we investigate two possibilities, firstly was the inability of the mutant activating peptides to activate the protease due to the fact that they did not interact with the protease, and secondly if there was an interaction was this proportional to amount of activity observed in the activity assays? Selected gels from this system are presented in fig. 3.05.

3.7.1 Extended Activating Peptide

A double serial dilution of GVQSLLKRRRCF was incubated with Wt protease for 10 min in TE buffer. Samples were prepared for analysis on 15% non-denaturing, non-reducing polyacrylamide gel and separated at constant 180V for 45 minutes. To investigate the potential of an extended peptide's ability to bind the activating peptide groove GVQSLLKRRRCF was incubated with Wt protease for 10 min at 37°C. The protease appeared as a smear upon resolution with Coomassie brilliant blue staining suggesting that the inactivity reported earlier was due to the failure of the peptide to bind with the protease (data not shown).

3.7.2 VEGGS

Protease bands were resolved with Coomassie brilliant blue. The protease appeared as a smear in all lanes of the gel (fig 3.07). This suggested that VEGGS failed to interact with the protease, and in part would explain why there was no digestion of substrate in the activity assays.

3.7.3 The EDS Activating Peptide

A doubling serial dilution of GVRYGSQRYCY peptide was prepared in reduced TE buffer and incubated with Wt protease for 10 min at 37°C. Samples were loaded onto 15% non-denaturing/reducing polyacrylamide gels and separated at constant 180V for 45 min. Two faint dumb bell shaped bands were observed on one (fig.3.07). On all other occasions heterodimerisation of peptide and protease were evident with 250μM and 125 μM activating peptide. This would be consistent with reduced activity observed with the EDS peptide due to a less stable heterodimer complex formation.

3.7.4 GVQSLKRRRCA

A doubling serial dilution of GVQSLKRRRCA peptide was prepared in reduced TE buffer and incubated with Wt protease for 10 min at 37°C. Samples were loaded onto 15% non-denaturing/reducing polyacrylamide gels and separated at constant 180V for 45 min. Heterodimerisation of peptide and protease was evident with 250μM and 125 μM activating peptide (fig 3.07). This would be consistent with reduced activity observed with the RCA peptide due to a less stable heterodimer complex formation.

3.7.5 GVQSLKRRACF

A doubling serial dilution of GVQSLKRRACF was prepared in reduced TE buffer and incubated at 37 °C for 10 min with Wt protease. Samples were then separated on 15 % non-denaturing non-reducing gels for 45 min at a constant 180V. As expected from the activity studies this peptide did bind to the protease albeit to a lesser extent than observed with the Wt activating peptide (fig 3.07). Bands were not seen below 63 μM of the peptide.

3.7.6 GVQSLKRARCF

A doubling serial dilution of GVQSLKRARCF was prepared as above and separated on 15% non-denaturing non-reducing gels. As expected the protease and peptide interacted. Again the interaction was not to the same extent as the Wt peptide but clearly observable until below 63 μM peptide (fig 3.07).

3.7.7 GVQSLKARRCF

A doubling serial dilution series of the peptide GVQSLKARRCF was incubated at 37 °C for 10 min with Wt protease . The peptide was sufficient to induce the formation of a distinct protease band when the non-denaturing non-reducing gels were stained with Coomassie brilliant blue stain. The interaction illustrated in figure 3.07 is equivalent to that of the Wt activating peptide. The complexed species band disappears below 19μM peptide.

3.7.8 GVQSLARRRCF

A doubling serial dilution series of the peptide GVQSLARRRCF was incubated at 37 °C for 10 min with Wt protease. The peptide did induce the formation of distinct protease bands when the non-denaturing non-reducing gels were stained with Coomassie brilliant blue stain. The interaction did not occur below 63 μM peptide (fig 3.07).

3.8 Scatchard Analysis

Band intensities of the protease/peptide heterodimer from non denaturing/non reducing gels stained with Coomassie brilliant blue were estimated using NIH Image version 1.62. The ratio of bound/free protease peptide complex was plotted against bound. The results are shown in figure 3.08 and table 3.01. The results show that a method to estimate binding affinity could be developed from the method described above, however, the

results were not considered to be reliable enough to form the basis of a discussion.

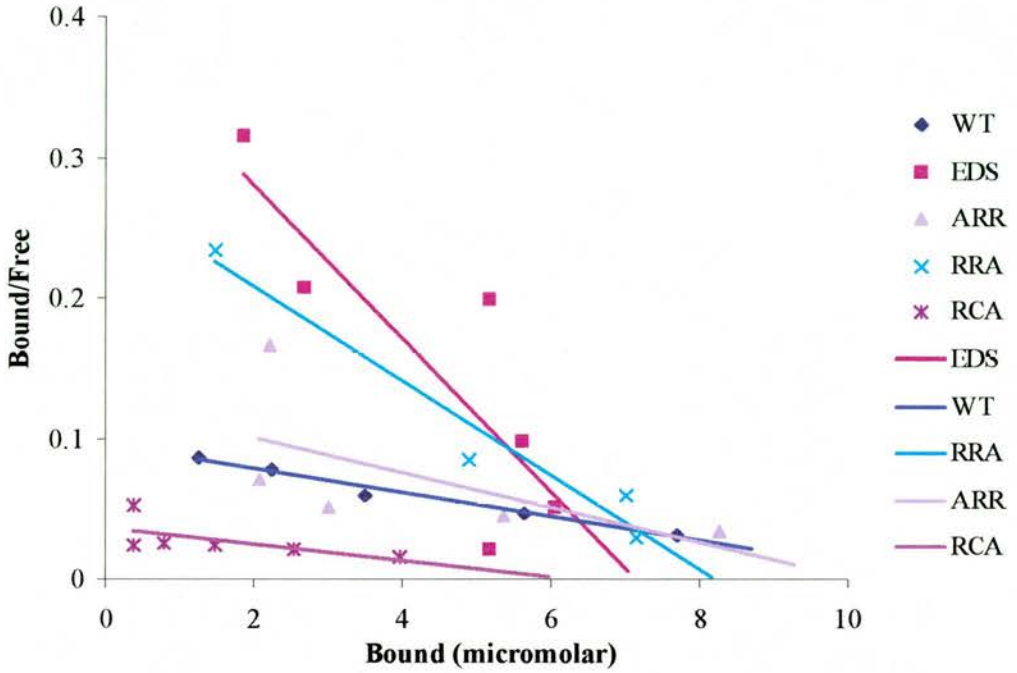


Fig 3.08. Scatchard analysis of protease/activating peptide binding. The figure shows preliminary data for Wt (GVQSLKRRRCF); EDS, activating peptide from EDS virus; RRA (GVQSLKRRACF); ARR (GVQSLKARRCF) and RCA (GVQSLKRRRCA). Concentration of protease in the binding assay was 7.7 μM.

Activating Peptide	K_d (μM)	No. of sites
Wild type (GVQSLKRRRCF)	133.1	1.45
EDS Activating Peptide GVRYGSQRYCY	13	0.93
GVQSLKARRCF	127	1.31
GVQSLKRRACF	234	1.06
GVQSLKRRRCA	88	0.82

Table 3.01. Binding parameters derived from Scatchard analysis.

3.9 In vivo Expression Studies

To develop our understand of the spatial and temporal control of the Adenovirus protease during infection we decided to try and develop an *in vivo* expression system. Previous work within our group (Vaughan, 1997) had shown that during the lytic cycle of infection, at 28 hpi, protease was almost exclusively retained in the nucleus associated within amorphous nuclear inclusions. Why and how does this translocation occur? This compartmentalisation of the protease was not thought to be a result of diffusion since the protease was not evenly distributed between the nuclear and cytoplasmic compartments.

3.9.1 Tet-On and Tet-Off Expression System.

To further our understanding of the spatial and temporal control of the Adenovirus type 2 protease we attempted to stably integrate the protease and pVI genes in 293 and HeLa cell lines using the Tetracycline On and Off system developed by Clontech. Expression of the protease would be under the control of the tetracycline responsive plasmid pUHD10-3 (Gossen and Bujhard, 1992). In this system the level of protease expression could be controlled and a panel of activating peptides and potential anti-viral agents could be introduced into an *in vivo* system.

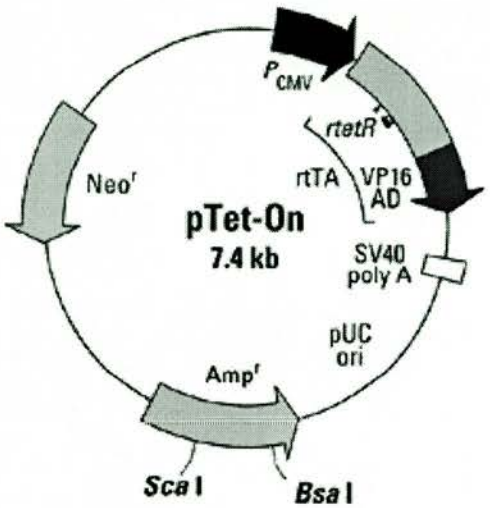


Fig 3.09 Tet-on /Tet-off plasmid (reproduced from the Clontech website) pTet-on is stably transfected into the cell line of interest and constitutively expresses reverse tet-responsive transcriptional activator (rtTA). Interaction of rtTA with Tet/Dox allows recognition and interaction with the tet-responsive element in pTRE (fig 3.10) (Gossen *et al,*. 1995).

In the Tet system Eukaryotic cell lines are transfected with two plasmids. The first of which expresses a transcription regulator (Figure 3.09) which in the case of the Tet-on system (Gossen 1995) produces reverse Tetracycline responsive transcriptional activator (rtTA) and in the Tet-off produces tTA. The tet-responsive transcriptional activator (rtTA) is active in the presence of Tet/Dox where as the tTA is active in the absence of Tet/Dox. The transcription activators regulate the expression of a second plasmid pUHD10-3 or pTRE2 (Figure 3.10) which contains the Tet Responsive Element (TRE) which responds to tTA and rtTA.

pTet-On and pTet-Off plasmids are stably integrated into the cell lines with continued selection assured by neomycin presence in the selective media.

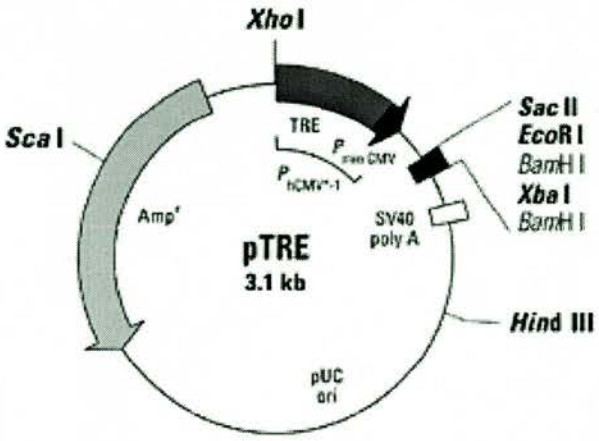


Fig 3.10 Tet responsive plasmid. pTRE2 (reproduced from the Clontech website) (pUHD10-3 Gossen and Bujhard 1992) contains TRE (tet-responsive element) which responds to tTA and rtTA. The gene of interest is cloned down stream of a promoter whose expression is controlled by the Tet-off and Tet-on plasmids tTA and rtTA.

3.9.1.1 Cloning of pVI and Adenovirus Type 2 Protease into pUHD10-3

The protease and pVI (nucleotides 21778-22389 and 18100-18750 respectively of the Ad2 genome) were cloned into pUHD10-3 using the primers detailed in method 2.1.4. Figure 3.11 shows pVI and protease cleaved from pUHD10-3 and separated on 1% agarose gel to confirm the insertion of the pVI and the protease genes. Incorporation of the protease and pVI genes was confirmed by ABI PRISM BigDye Terminator Cycle Sequencing.

3.9.1.2 Development of Stably Integrated Cell Lines

The mean lethal dose of antibiotic required to kill 50% of cells was calculated to determine the levels of antibiotic necessary to select for the stable integration of the protease and pVI genes into 293 and HeLa cell lines. A series of trials with individual and then cocktail of selective antibiotics were performed using doxycycline (dox), hygromycin (hyg) and geneticin (gen). The following conditions were used throughout the study, 1µg/ml dox, 50 µg/ml hyg, 200µg/ml gen.

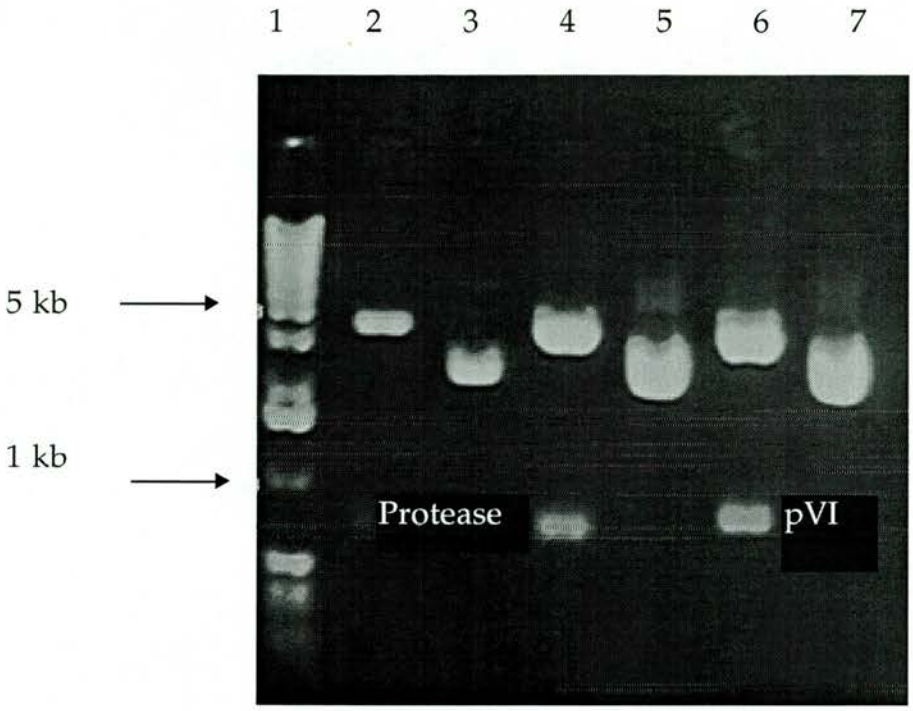


Figure. 3.11. Protease and pVI restriction digestion. pUHD10-3 minipreps with protease and pVI inserted were restriction digested for 2 hours and separated on 1% agarose gel. Lane 1, 1kB DNA ladder(Gibco); lane 2, uncut pUHD10-3 23k; lane 3, and 5, pUHD10-3; lane 4, pUHD10-3 23k cut with EcoR1 and Bam H1; and lane 6, pUHD10-3 pVI cut with EcoR1 and Bam H1. Protease migrated further than pVI as expected because of the difference in molecular weight.

3.9.1.3 Morphology of the transfected cell

Cells transfected with control DNA and untransfected cells illustrated a common morphological appearance, which throughout this study will be referred to as the normal phenotype. Cells when viewed at 100 x magnification under oil were flat and irregularly shaped. The nucleus was large and occupied approximately one third of the two dimensional image. Dead and dying cells were rounded in appearance if still attached to the substrate, or else had detached and were floating in the media.

3.9.1.4 Expression of the Protease and pVI

pUHD10-3 23K and pUHD10-3 pVI were transfected into HeLa and 293 cell lines using the lipofectACE protocol. The lipofectACE reagent is a liposome formulation which forms a lipid-DNA complex capable of interacting with cell membranes and delivering DNA into the host cell.

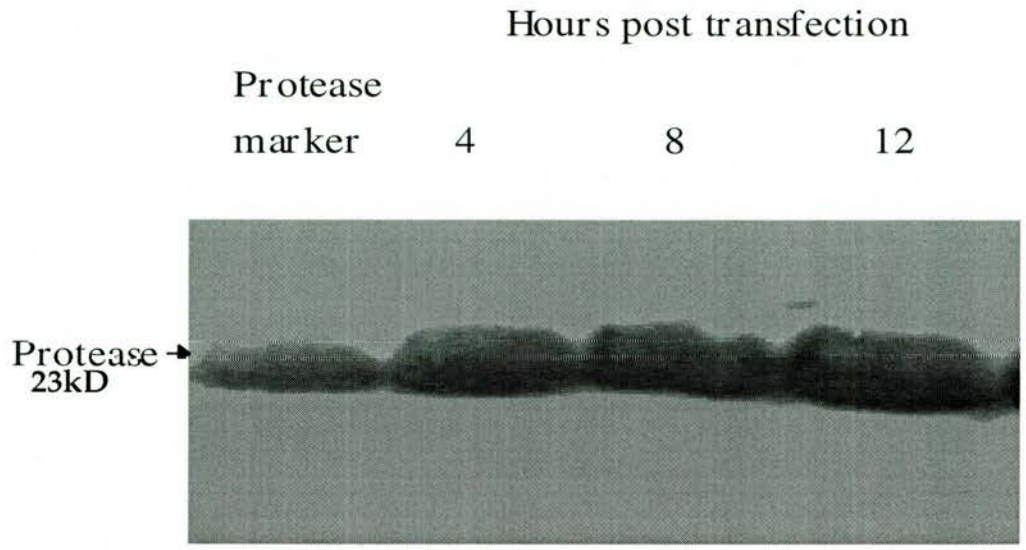


Fig 3.12 Expression of the Protease in Tet-off 293 system. 10 µg per million cells of DNA was was expressed from pUHD10-3 23k in the Tet-Off system. Expression of the protease lead to cell detachment from the monolayer as a result of cell death. Samples of cells were taken at 4, 8 and 12hpt to confirm protease expression. Samples were run on 15% SDS PAGE gel and examined by Western blotting. Protease was detected as before and pVI with anti pVI antibody kindly donated by W. Russell University of St. Andrews. The gel was run top to bottom.

Transfections were allowed to proceed for 4-16 h in OPTIMEM with DNA ranging from 100ng-10µg. The transfection mixture was replaced with serum enriched D-MEM and protein expression selected against with the absence or presence of Tet/Dox for the Tet-on and Tet-off systems respectively.

When OPTIMEM was replaced with complete media cells were viewed under 40 x magnification. Cells transfected with the protease and pVI resembled those of the control cell and of the cells pre-transfection i.e. the normal phenotype. Cells were healthy in appearance and adhered to the petri dish surface. By 16 hours post-transfection (hpt) cells were seen to round up and detach from the monolayer. This was observed for both pVI and protease transfections. When the protease and pVI were co-transfected the rounding up and detachment of the cells occurred later. Cell death was thought to be a result of the cocktail of antibiotics needed to select for the plasmid and control expression, as well as the toxic nature of protease and pVI which appeared to

be capable of a basal level of expression. Figure 3.12 is an example of a Western blot from transfected cells at 4, 8 and 12 after removal of OPTIMEM. The amount of protease expressed increased as the amount of cell death increased. Cells could not be prepared for immunocytochemistry because cells detached from the coverslips during the primary wash.

3.10 Protein-Protein Interactions

To see if the protease was interacting with a protein of cellular origin to facilitate nuclear translocation GST-protease fusion construct was purified using glutathione-Sepaharose beads. Figure 3.13 illustrates the experimental procedure used to identify a specific protein-protease interaction. A 100 litre HeLa S3 culture was pelleted and a total protein extract prepared. This extract was continually passed through consecutive columns of Sepharose, Sepharose-GST and Sepharose-protease beads for 16 h at 4°C and at a flow rate of 0.5ml/min. Figure 3.14 shows the resulting gel from each of the columns.

Two bands were excised form the SDS PAGE gel and sent for sequencing. These bands were identified as GST-protease fragments thought to have been cleaved by the protease as previously described (Houde and Weber, 1990; Webster and Kemp, 1993)

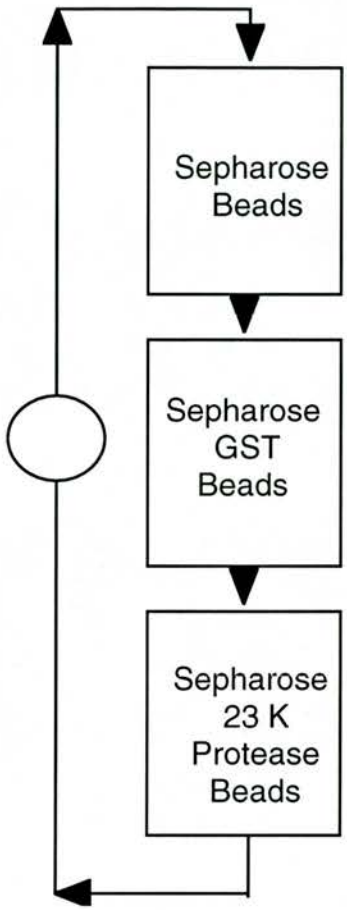


Figure 3.13 GST Pulldown Apparatus. A total HeLa cell protein extract was prepared fom 100 litres of culture and continually cycled through the columns overnight at 4°C. Proteins were eluted from the column with 10mM glutathione and precipatated with 10% final concentration TCA. Eluates were run on 15% SDS PAGE gels (fig 3.14.) specific bands were excised, trypsin digested and solublised in 0.1% TFA. Samples were submitted for protein sequencing or mass spectrometry to identify the protein bands.

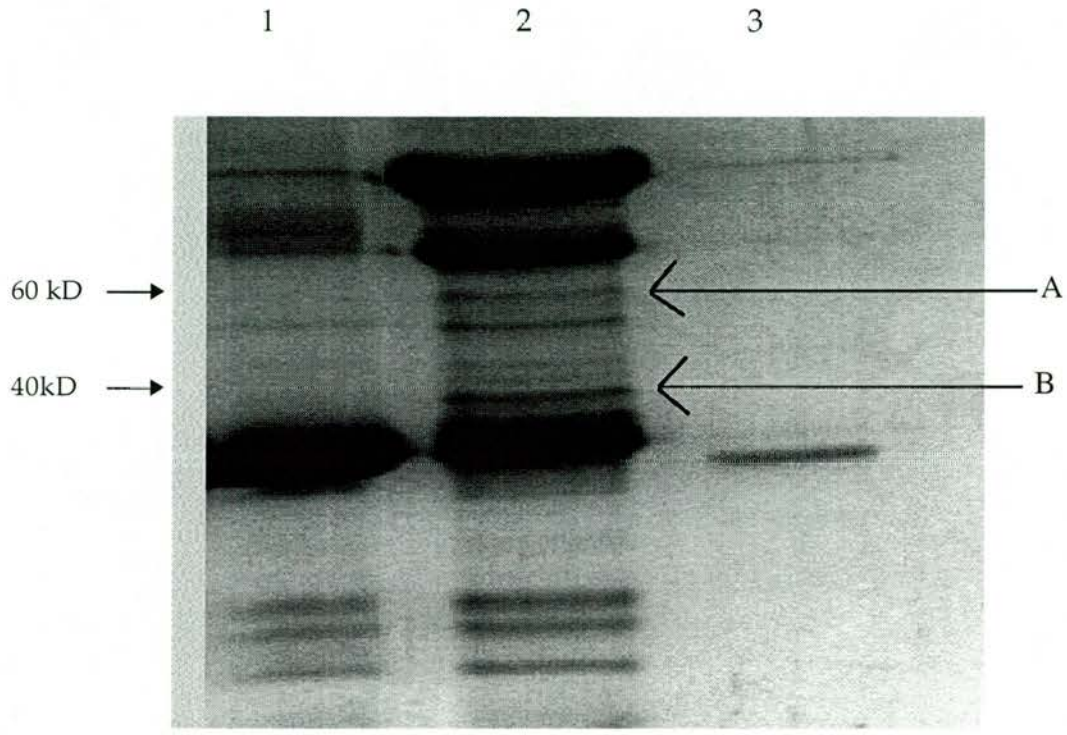


Figure 3.14 GST Protease as a bait for Potential Protein-Protease Interactions. 15% SDS PAGE gel separation of proteins eluted from; lane 1, Sepharose-GST Bead; lane 2, Sepharose-GST Protease fusion protein; and lane 3, Sepharose-Glutathione beads. A total HeLa cell protein extract was prepared and incubated with the beads overnight a $4^{O}C$ in continuous flow at 0.5ml/min. Proteins were eluted from the column in 10mM reduced glutathione and concentrated by precipitation with TCA. Bands were stained with Coomassie Brilliant Blue. Arrow A and B denote bands excised and digested with trypsin, resultant bands were extracted and analysed by protein sequencing and mass spectrometry.

3.11 Cloning of Wild Type Protease and Mutants C104A, and C122A into pCDNA3

No cellular protein could be identified to interact with wild type protease from GST protein-protein interaction assay, we therefore decided to perform a series of transient expressions of protease and its mutants in the $HeLa^{CD}$ cell line using pCDNA3.1 + and - vectors. The advantage of this system was that no antibiotic selection was required to control expression. Using this system, the Adenovirus protease and its mutants could be expressed with and without the activating peptide and their localisations determined.

The wild type protease was PCR amplified and cloned into pcDNA3.1+ at EcoR1/BamH1, whilst C104A and C122A were sub-cloned from pET11-C into pCDNA3.1- at Xba 1 Bam H1 sites. The only difference between + and - vectors was the orientation of the MCS. Products from cloning were electroporated into DH5a and 6 colonies from the resulting agar plates were selected to prepare DNA for sequencing. Maxi preps of DNA were prepared and stored in ddH_20 at a final concentration of 1mg/ml.

Figure 3.15 shows pCDNA 3 contructs with WT, C104A and C122A protease cloned. Insert DNA was excised from the constructed plasmids (EcoR1 and Bam H1 for WT, and Xba 1 and BamH1 for C104A and C122A). As expected the insert band appeared at approximately 650 bp as compared to the 1kb DNA ladder. To confirm that the inserts had been successfully inserted into the plasmid samples were sequenced.

HeLa cells were transfected with 4µg per million cells of each of pCDNA3.1+ and - as controls, pCDNA3.123k, pCDNA3.1C104A and pCDNA3.1C122A. Expression the protease and the two mutants was allowed to continue for 8, 12, and 16 h and lead to the characteristic cell rounding and detachment from the monolayers.

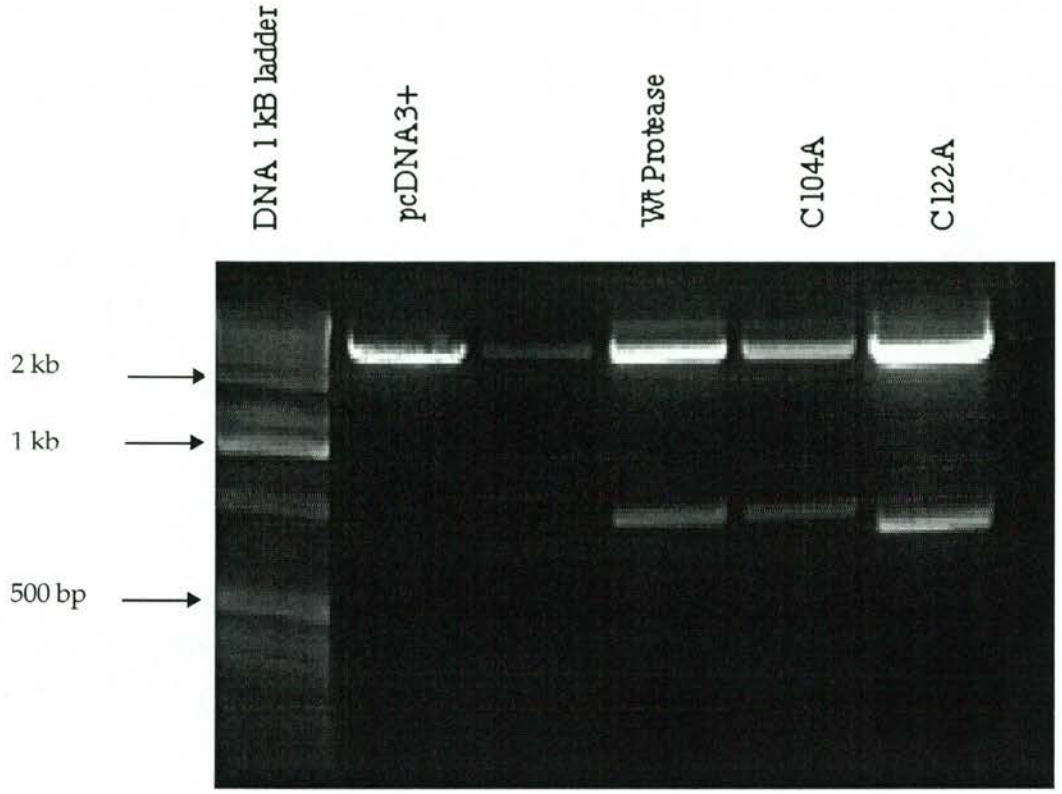


Figure 3.15 1% Agarose Gel Showing Protease, C104A and C122A excised from pCDNA3. 1% agarose gel showing, lane 1 DNA 1 kB ladder, lane 2 pcDNA vector, lane 4, digestion products of pcDNA3 23k cut with EcoR1/BamH1, lanes 5 and 6 respectively pcDNA3C104A and pcDNAC122A cut with Xba1 and Bam H1. Digestions were allowed to progress for 4 hours. The lower bands represent the protease gene and the upper band vector. Gel was run at constant 8 V for 1 h. All fragments were approximately correct size and were submitted for DNA sequencing.

Antibodies to the Adenovirus protease were kindly donated by T Vaughan and used in localisation studies. However, the monoclonal antibody to the protease produced a high level of background fluorescence and was thus unsuitable for immunocytochemistry. That protease expression has occurred was confirmed by Western blot analysis using polyclonal antisera R11 shown in figure 3.16.

Wt protease expression led to cell death, cells were seen to round up and detach from the monolayer. Control transfections with pCDNA+ and - did not undergo this detachment and coverslips grew to confluence in 36 h.

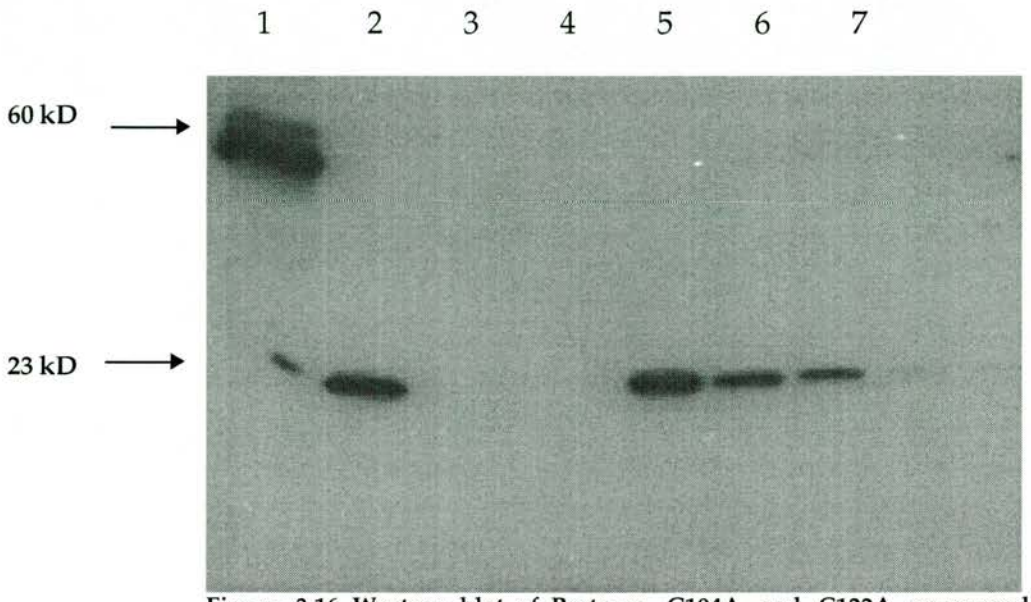


Figure 3.16 Western blot of Protease, C104A and C122A expressed from pCDNA 3.. Cells were transfected with 4 μg of DNA using electroporation. Lane 1, pre-stained standards; lane 2, Wt protease marker sample; lane 3, blank; lane 4, negative control protease; lane 5, Wt-protease; lane 6 C104A; lane 7 C122A. Total cell extracts were separated on 15 % SDS PAGE gels and Western blotted.

The detachment of the cells was concluded to be as a result of protease expression. DAPI staining revealed DNA to be condensed and fragmented within Wt protease and C104A transfected cells. Condensed nuclei were not observed with C122A transfections or controls. The fact that the toxic effect of the protease was observed in the absence of antibiotics is, in some part, confirmation that the cell death observed in the development of stable transfection in the Tet-On/Off system was probably attributable to protease expression control problems.

3.11.1 Cytokeratin Cleavage

It has been reported that the Adenovirus protease was responsible for the cleavage of cytokeratin at the end of infection (Chen *et al.*, 1993). To assess whether this cleavage could occur in the absence of activating peptide HeLa cells were transfected with 4μg per million cells of each of pCDNA3.1+ and - as controls, pCDNA3.123k, pCDNA3.1C104A and pCDNA3.1C122A as above. Cells were prepared for immunoflourescence with anti-cytokeratin

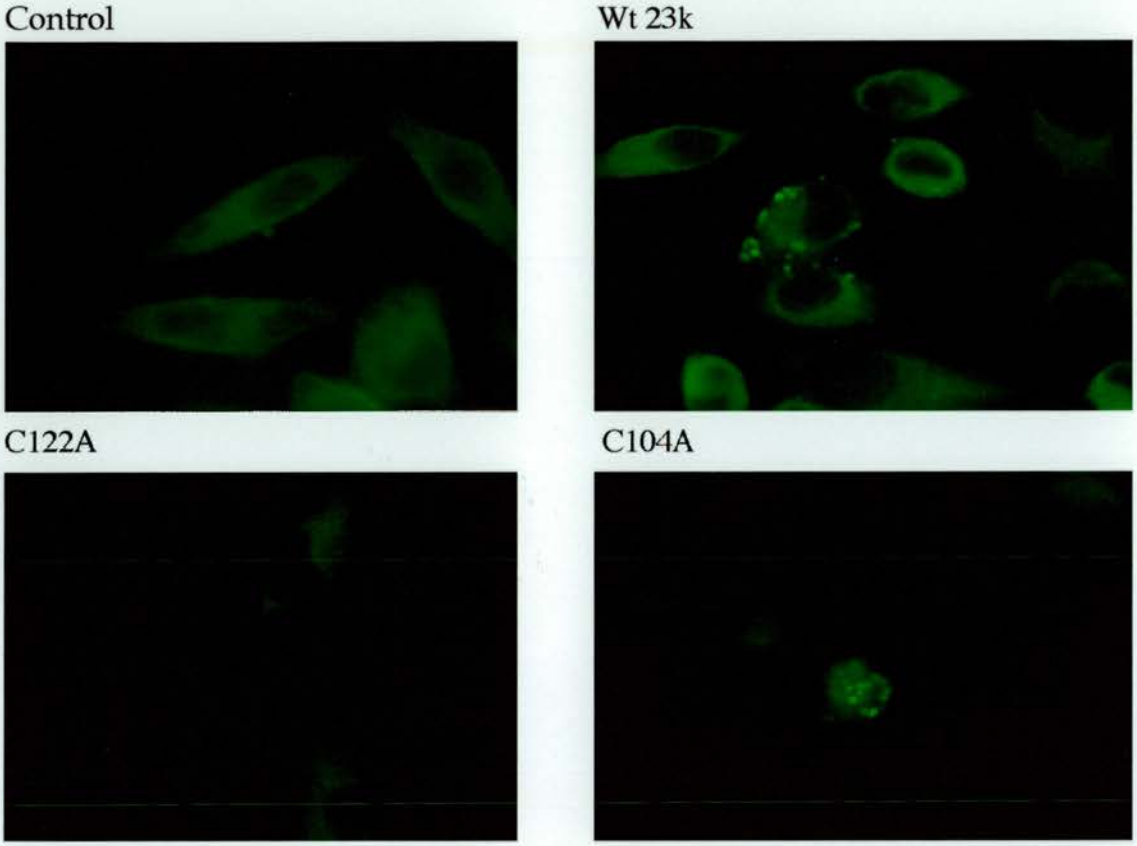


Fig 3.17. Expression of Wt-Protease, C104A and C122A Mutants. 4μg of each plasmid was transfected per million HeLa^{CD} cells. Expression was allowed to continue for 16 h after which cells were permeablised, fixed and probed with anticytokeratin monoclonal antibody (Anticytokeratin 8.13 clone Sigma). Expression of Wt protease and the mutant C104A led to cell rounding and more dense perinuclear cytokeratin staining - consistent with cytokeratin cleavage.

monoclonal antibody (K8.13 Sigma). Protease expression was confirmed from total protein cell extract prepared in duplicate to the above experiment and results confirmed that protease expression had occurred.

From figure 3.17 we can see that expression of the Wt protease lead to cell rounding, a loss of polarity and dense perinuclear cytokeratin staining - consistent with that of cytokeratin cleavage. A similar pattern was observed with C104A, cells were no longer spindle shaped with diffuse filamentous patterns of cytokeratin, rather cytokeratin was localised in the perinuclear region forming heavily stained globular structures comparable to those reported elsewhere (Chen *et al.*,1993; Zhang and Schneider, 1994 ; Strafenbiel *et al.*, 1986). C122A did not induce these phenotypic changes although the expression of the protease did result in some cell death as comapred to the control.

3.11.2 Activating Peptide and Nuclear Translocation

Due to observations which illustrated the toxic nature of the protease, we decided to adopt an indirect approach to further our understanding of the spatial control of the protease. The activating peptide contains a run of 4 positive amino acids KRRR which could potentially act as a nuclear localisation signal. To test this hypothesis oligonucleotides corresponding to the activating peptide were cloned C-terminally of pyruvate kinase. The activating peptide amino acid sequence and mutants thereof were reverse translated into the most nondegenerate corresponding DNA codon sequences with the addition of a stop codon. The sense oligonucleotides are depicted in table 3.02

Mutant	Sense oligonucleotide
SV40 NLS	GATCCGCTCCAAAGAAGAAGCGCAAGGTGGAATA GT
KRRR	GATCCTCTCTTAAACGTCGTCGTTGTTTTTAGG
KARA	GATCCTCTCTTAAAGCTCGTGCTTGTTTTTAGT
AARA	GATCCTCTCTTGCTGCTCGTGCTTGTTTTTAGT
ARRR	GATCCTCTCTTGCTCGTCGTCGTTGTTTTTAGT
KARR	GATCCTCTCTTAAAGCTCGTCGTTGTTTTTAGG
KRAR	GATCCTCTCTTAAACGTGCTCGTTGTTTTTAGG
KRRA	GATCCTCTCTTAAACGTCGTGCTTGTTTTTAGG

Table 3.02 shows the sequence of oligonucleotides synthesised containing the point mutations with in the NLS region of the activating peptide derived from C-terminal of pVI. Complementary oligonucleotides were hybridised and cloned into the Xba1/BamH1 C-terminally of pyruvate kinase.

After hybridisation of the oligonucleotides, samples were separated on 3% agarose gels (fig 3.18) to confirm that the complementary strands had annealed. Hybridised oligonucleotides were then cloned into pCDNAPK (kindly donated by Manuel Rodriguez, University of St. Andrews) at Xba/BamHI. The pCDNAPK construct contains part of the gene coding for pyruvate kinase with a C-terminal MCS. After the hybridised oligonucleotides had been cloned into the vector they were sent to the DNA sequencing unit at The University of St. Andrews.

Products from ligation were transformed into DH5α and plasmid preparations for transfections carried out with 4μg of DNA per million $HeLa^{CD}$ cells (Interleukin 2 independent cell line donated by Catherine Dargemont, Institut Currie-CNRS, Paris) and expression continued to a maximum of 16 h.

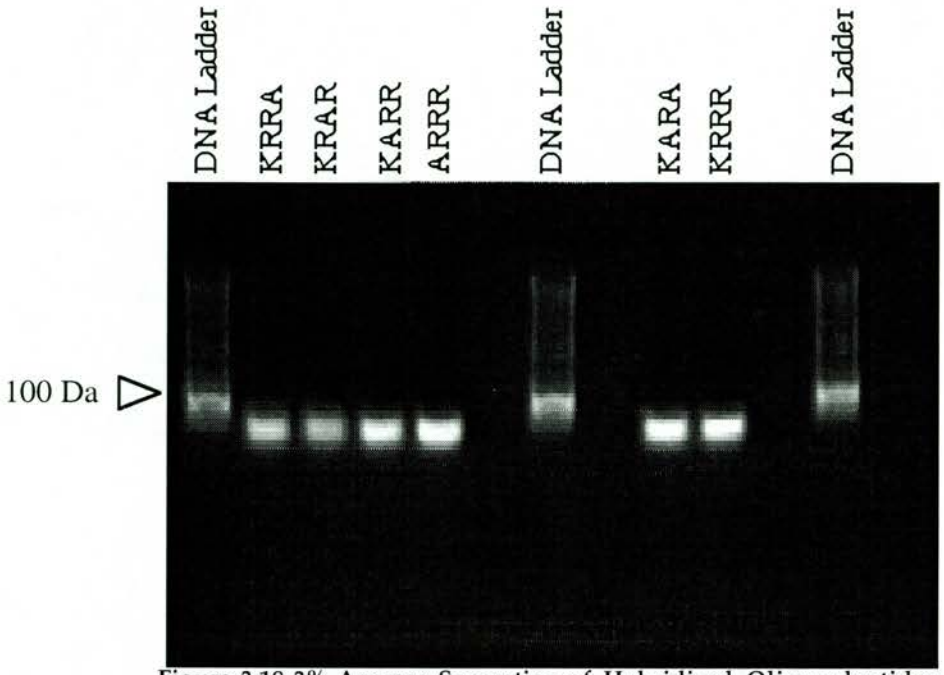


Figure 3.18 3% Agarose Separation of Hybridised Oligonuleotides. Oligonucleotides were boiled for 5 min and cooled to 25°C below their melting temperature to allow annealing. Hybridised Oligonucleotides were then cloned C-terminally of PK. Plasmid preparations were sequenced and transfected into DH5α.

The initial experiments were conducted to see if the activating peptide could target the construct to the nucleus. Clearly if KRRR was present pyruvate kinase was detected in the nucleus (Fig 3.19), however if the motif was mutated to AARA or KARA pyruvate kinase was exclusively cytoplasmic. The NLS from SV 40 large T Antigen as expected was localised to the nucleus (data not shown). Pyruvate kinase alone was unable to enter the nucleus as was the case with the other mutants. Figure 3.20 illustrates a Western blot of transfections carried out with the SV 40 NLS, the activating peptide NLS, pyruvate kinase and pyruvate kinase with no NLS. The blot was developed with anti-pyruvate kinase antibody (9E10CD kindly donated by Catherine Dargemont), and confirmed that the immunofluorescence pattern detected when the activating peptide/pyruvate kinase chimer was expressed in HeLa was attributable to the KRRR motif.

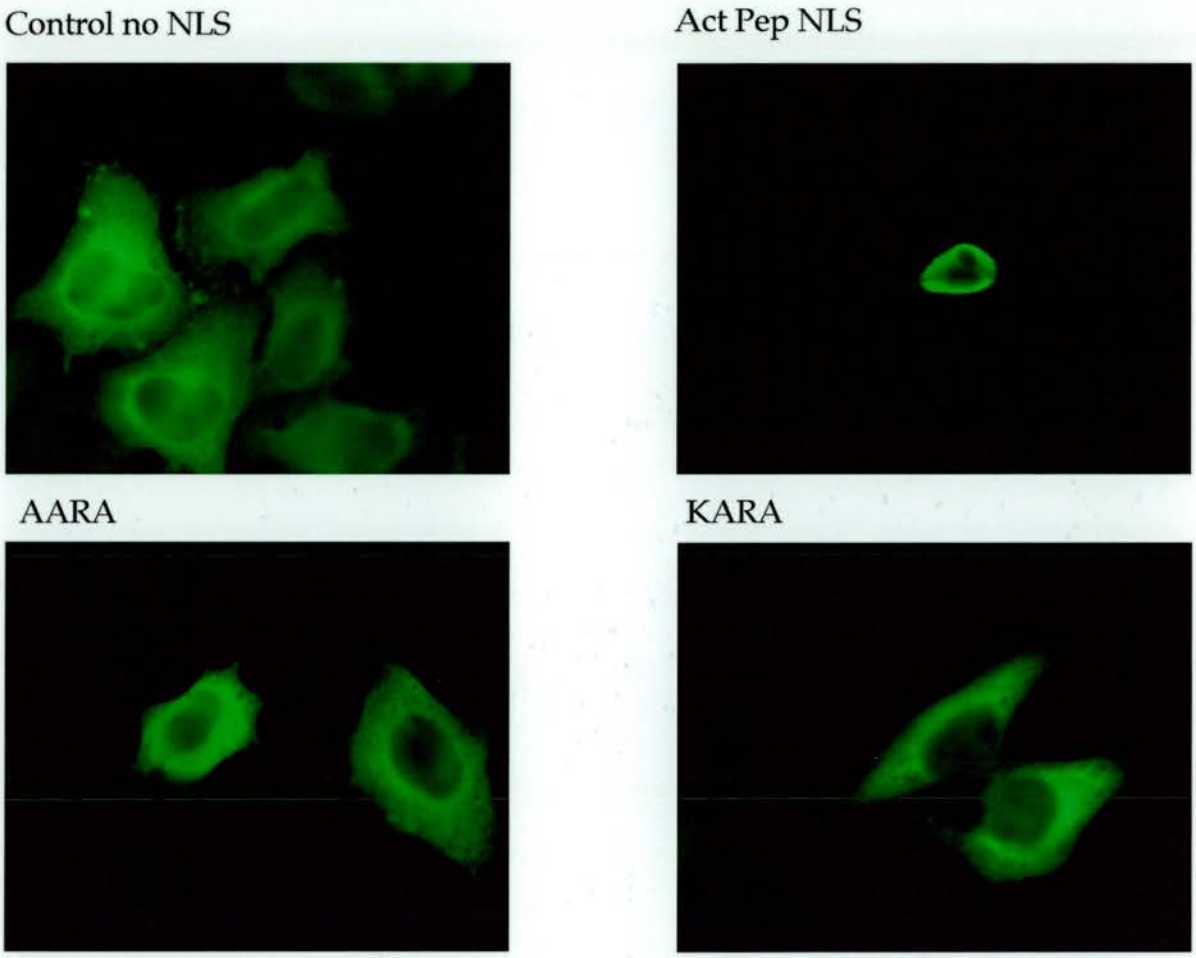


Fig 3.19. Initial NLS Study. Immunofluorescence of pCDNAPK constructs containing the potential NLS of the activating peptide. 4μg of plasmid per million cells was transfected into HeLa^{CD} cells and expressed for 16 h. Immunofluorescence was developed using 9E10CD antibody (Catherine Dargement, Institut Currie-CNRS, Paris, France) to pyruvate kinase. The activating peptide sequence was suffiecient to confer localisation of PK to the nucleus. Both AARA and KARA were unable to target the chimeric protein to the nucleus. As expected PK was maitained within the cytoplasmic compartment in the absence of a NLS.

To delimit the minimum sequence for nuclear localisation, constructs were designed to contain the sequence KRRA, KRAR, KARR and ARRR. Nuclear localisation was observed with KRRA and KRAR. However, staining in the KRAR expression study was predominantly cytoplasmic.

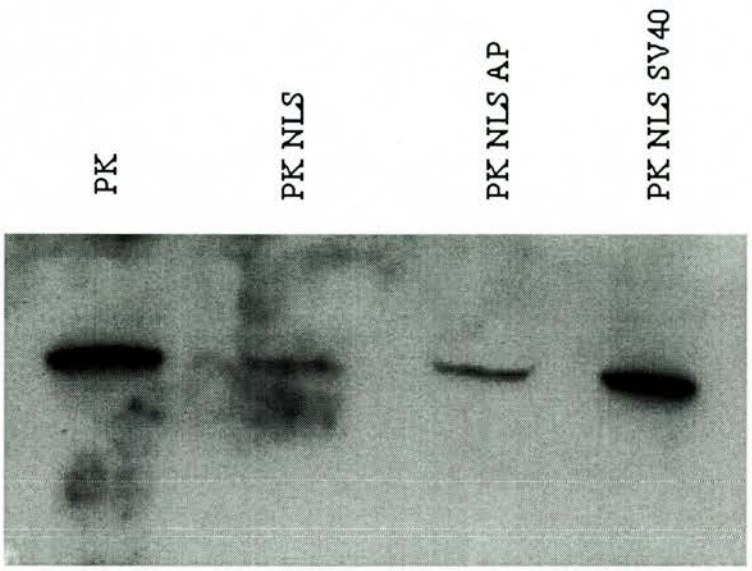


Figure 3.20 Western Blot of PK Expression. Western blot of cell extracts. Lane 1, pyruvate kinase; lane 2, Pyruvate kinase with N-term NLS; lane 3, pyruvate kinase with KRRR from activating peptide; lane 4, pyruvate kinase with SV 40 NLS. Transfections were allowed to proceed for 16 h and were initiated with 4μg DNA per million cells. Cells extracts were separated on 15% SDS PAGE gels .and examined using monoclonal antibody 9E10CD. The gel was run top to bottom.

3.11.3 Can Activating Peptide be used to Target Chimeric Proteins to the Nucleus?

The formation of the heterodimer of activating peptide (Ding *et al*, 1996) occurs along a non-contiguous region of the protease resulting in the formation of a β-sheet complex. If the minimum region of interaction could be determined and fused to pyruvate kinase could one target the construct to the nucleus in a temporally controlled manner simply by adding activating peptide to the culture media?

Using the crystal structure of the Adenovirus protease three potential regions of interaction were identified: amino acids 103-112, 98-114 and 98-148. 103-112 and 98-114 represented the areas where the activating peptide and protease

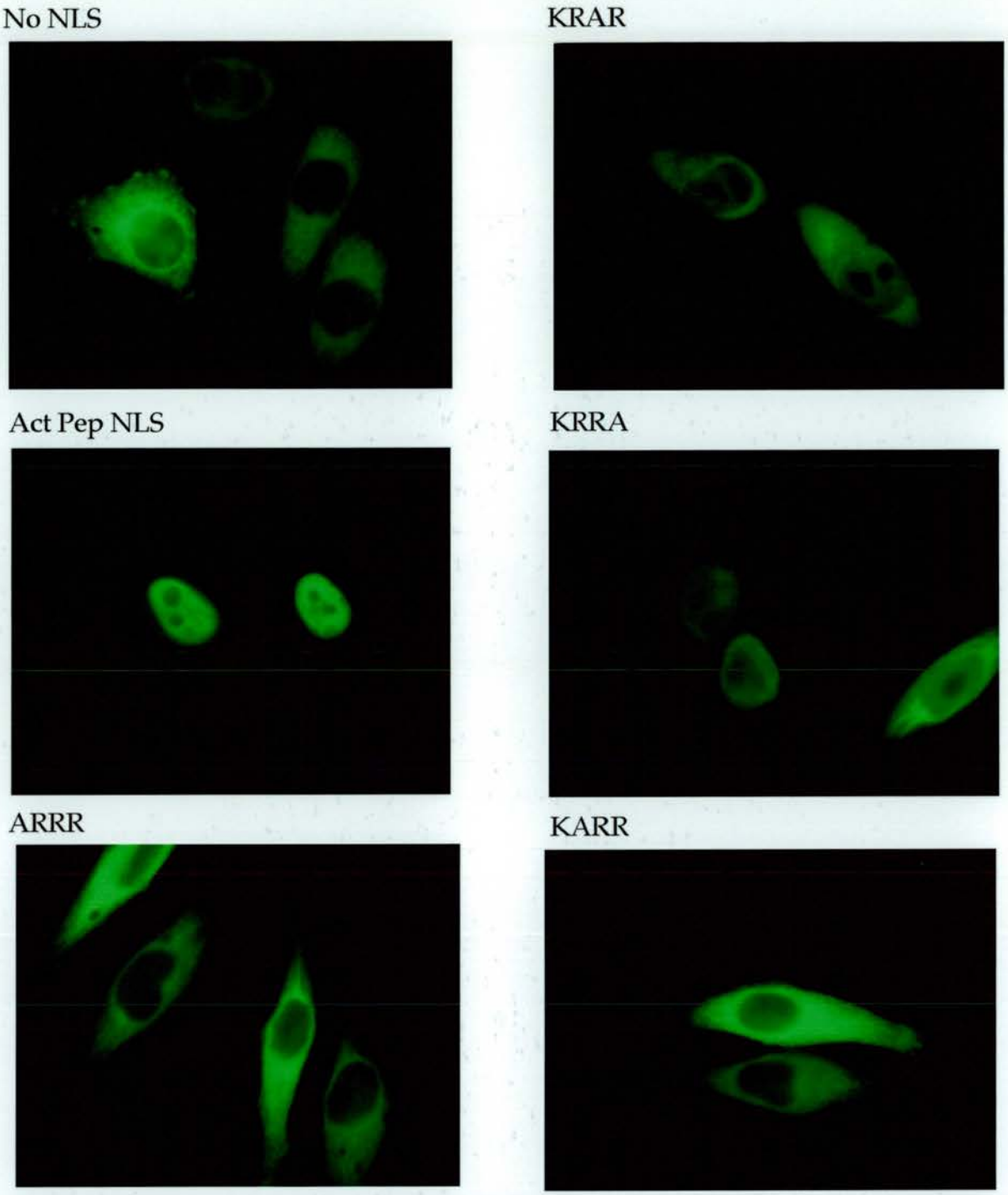


Fig 3.21. Single Amino Acid Substitutions. Single amino acid substitutions to delimit the minimum NLS at 100x magnification under oil. HeLa^{CD} cells were transfected with 4μg of each plasmid per million cells and left to express for 16 h. Immunofluoresence was carried out using 9E10CD antibody (donated by CatherineDargemont, Insitut Currie-CNRS) at 1:200 dilution. As expected there was no nuclear localisation with the negative control(top left), KARR or ARRR. KRRA and KRAR were detected in the nucleus, albeit to a lesser extent than the Wt activating peptide sequence.

most closely interact. 98-148 was selected to allow for any potential secondary structural requirements. Oligonucleotides were designed corresponding to regions 103-112 and 98-114, and PCR primers to amplify the region 98-148. Products from these reactions were cloned into Xba 1 /BamHI sites of pCDNAPK. The primer and oligonucleotides constructed are shown in Table 3.03 a&b.

Table 3.03 a. Hybridisation oligonucleotides.

Protease region	Sense Oligonucleotides
103-112	GATCCCGCTGTATAACGCTGGAAAAGTCCACCA ATGAT
98-114	GATCCGCTTCTTCCCCCGACCGCTGTATAACGCT GGAAAAGTCCACCCAAAGCGTGTGAT

The resulting plasmid constructs were sequenced and the expected results returned. Transfections of these plasmids were initiated with 4µg DNA and cells left to adhere to coverslips for 4 h before activating peptide was added to the media for 16 h. Initial results illustrated nuclear translocation of the constructs following the addition to a final concentration of 10mM GVQSLKRRRFC to the media had occurred. However, this result was not reproducible, even when varying the concentration of activating peptide across 2 log scales, and the length of assay. The preliminary investigations results are shown in figure 3.22.

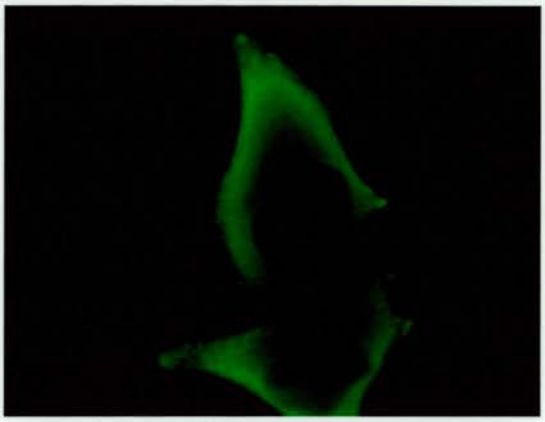
Table 3.03 b. PCR oligonucletides.

Protease region	primer	Oligonucleotide sequence
Forward 98-148		CGCGGATCCGCTTCTTCCCCCGACCGCTGT
Reverse 98-148		CGCTCTAGATCAGTTCATGGTGGGGTTGTGATC

Table 3.03 a and b The Olignucleotides Sequences used for the construction of pCDNAPK103-114, pCDNAPK98-112 and pCDNAPK98-148. Table a shows the sense oligonucletides which were hybridised and cloned into Xba/Bamhl sites at the C-terminus of pyruvate kinase. Table b shows the oligonucleotide sequences used to prime the PCR reaction to amplify 98-148 of the Adenovirus protease.

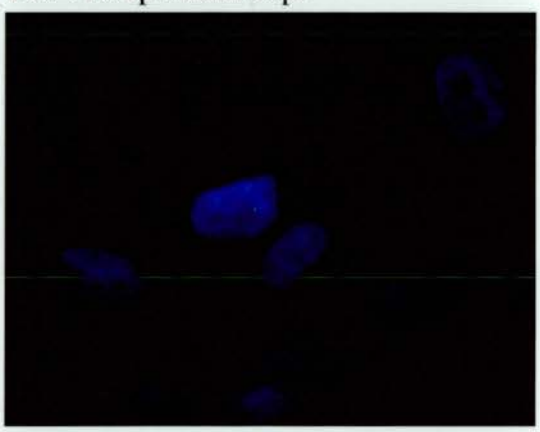
PK-pCDNA3

103-112 PK Control

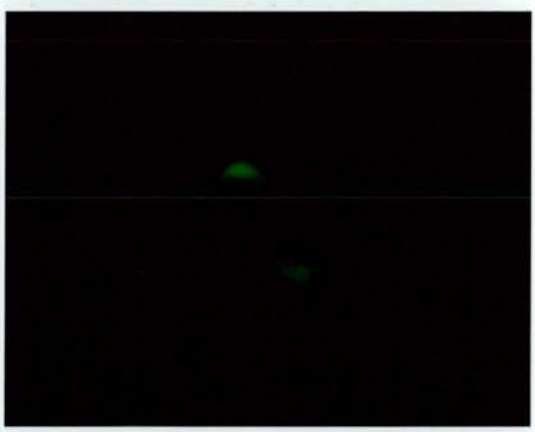


103-112 ap media

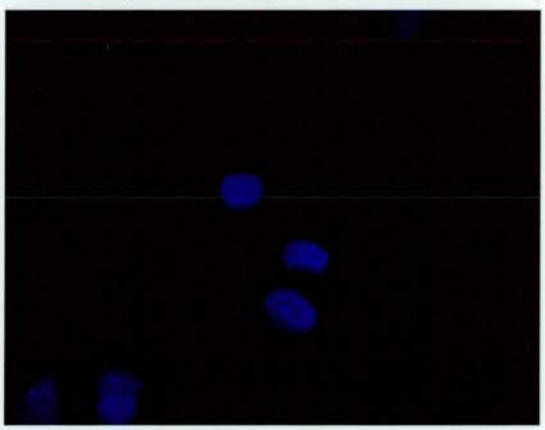
103-112 ap media dapi



98-114 ap media



98-114 ap media dapi



98-148 ap media

98-148 ap media dapi

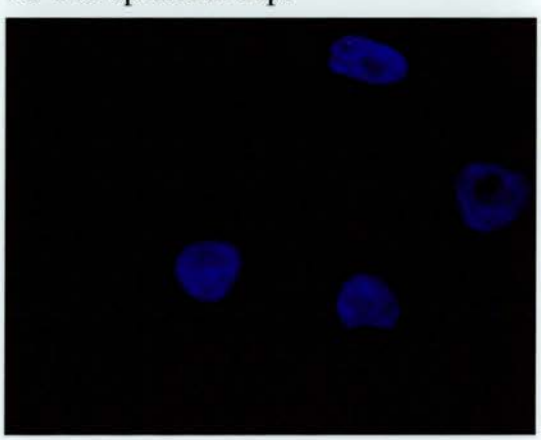


Fig 3.22 Inducible Translocation System. 4 μg per million cells of each plasmid was transfected into HeLa ^{CD} cell lines and left for 16 h. Activating peptide was added to the culture media to a final concentration of 10mM. Cells were fixed in 4% paraformaldehyde, permablised and probed with 9E10CD monoclonal antibody to pyruvate kinase. The chimeric protein was detected in the nucleus of cells transfected with pCDNAPK98-114, pCDNAPY103-112 and pCDNA98-148. Staining was cytoplasmic in the pCDNA3 control experiment and controls of the test chimers (in absence of activating peptide (not shown here)). Dapi staining was used to delimit the nucleus.

4. Discussion

4.1 Nuclear Localisation Studies.

Cellular compartmentalisation is essential for the efficient physiological control of metabolism. The nuclear envelope forms one of these compartment boundaries ensuring that the replication of genetic material in the nucleus is separated from translation in the cytoplasm. Macromolecules can either diffuse into the nucleus, (at a rate inversely proportional to their molecular weight) or be actively transported. Generally proteins with a molecular weight less than 50 kD form an equilibrium across the membrane, however complexes with larger molecular weights can be targeted to the nucleus in an energy dependent manner. Translocation of these species is energy dependant and, if cells are chilled transport is abolished (Breeuwer and Goldfarb 1990). Proteins smaller than 50 kD have been identified as carrying nuclear localisation signals (NLS), e.g. HIV tat and Histone H1 so it is reasonable to assume that the protease may have an NLS.

During Adenovirus infections it was noticed that the protease (23kD) did not equilibrate between the compartments even though its size would suggest that it should (Vaughan, 1997). It was shown that the protease was either in the cytoplasm, 24 hpi or in the nucleus by 28 hpi. What was the signal for this translocation as simple equilibrium kinetics did not explain this shift? The protease is required in fulfil functions in both cellular compartments, i.e. in the nucleus to cleave precursor polypeptides in the maturing virion (Anderson 1973 and Webster *et al.*, 1989), and in the cytoplasm to cleave the cytoskeleton to assist the virus in escaping the host cell at the end of infection (Chen *et al.*, 1993; Zhang and Schneider 1994).

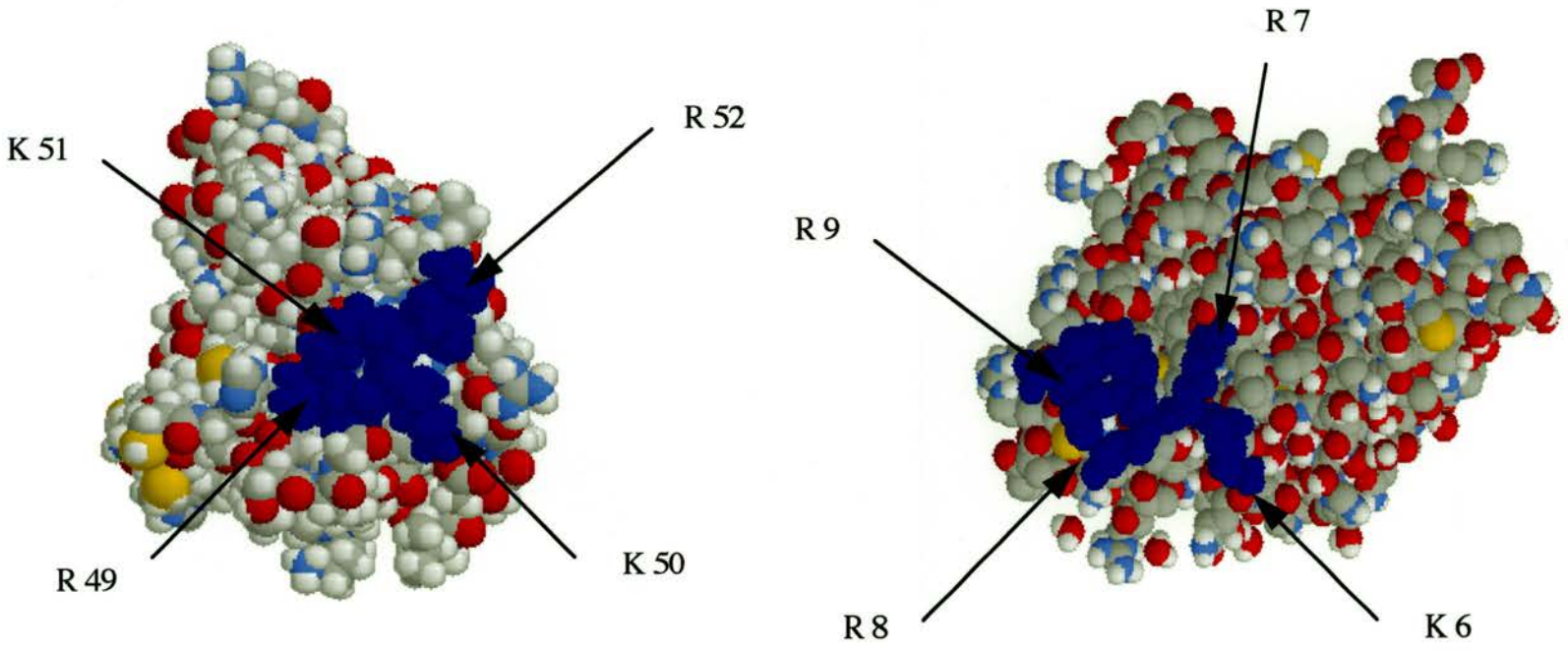


Fig 4.01 Comparison of the NLS of HIV Tat protein and that of the protease pVI heterodimer. Despite the difference in size between the two proteins the overall conformation homology of the NLS's are stikingly similar. Tat is displayed on the left and the protease heterodimer on the right. Amino acid labeling of the protease refers to the order they appear in the activating peptide.

Evidence presented here may implicate the novel activation mechanism also doubling as a nuclear localisation signal targeting the protease to the nucleus. The activating peptide amino acid sequence corresponding to the last 8 residues was translated into the genetic code and cloned onto the C-terminus of pyruvate kinase construct in pCDNA (pCDNAPK). Figure 3.16 clearly illustrates that the peptide sequence, KRRR, is sufficient to confer nuclear localisation to pyruvate kinase. If K6, R7 and R8, or R7 and R8 are simultaneously mutated to A the nuclear localisation observed with the KRRR construct is abolished. As expected the NLS of SV 40 large T antigen targeted the protease to the nucleus. Thus from these preliminary experiments the sequence KRRR was concluded to be a nuclear localisation signal.

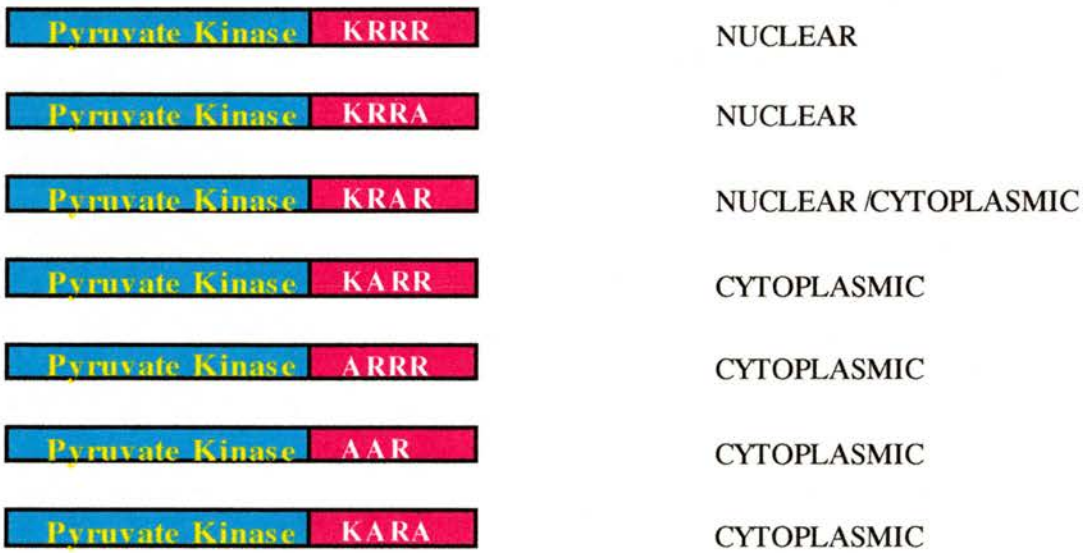


Fig 4.02 Chimeric Pyruvate Kinase NLS Constructs. Summary of the important residues in nuclear localisation. K6 and R7 were the most important residues in conferring nuclear localisation to pyruvate kinase, substitution of either these residues abolished the sequences ability to target the construct to the nucleus. KRRA and to a lesser extent KRAR were able to send the construct to the nucleus.

In an attempt to delimit the minimal NLS a series of single mutations were engineered into to the sequence. Figure 3.17 illustrates the localisation of

these constructs, figure 4.02 above summarises the observations from figures 3.16 and 3.17.

KRRA is sufficient to confer nuclear localisation to pyruvate kinase, although the effect was not as great as that observed with KRRR. The KRAR construct also showed some nucleoplasmic staining, however, staining intensity was predominantly in the cytoplasm. If K6 or R7 were mutated to A nuclear localisation was abolished. We conclude that K6 and R7 were the absolute requirements for nuclear localisation of the pyruvate kinase.

Human	type	2	GVQSLKRRRCF
Human	type	5	GVQSLKRRRCF
Human	type	12	GVKSLKRRRCY
Human	type	31	GVKSLKRRRCY
Human	type	40	GVKSLKRRRCY
Human	type	41	GVKSLKRRRCY
Canine	type	1	GLSNVKRRRCF
Canine	type	2	GLSNVKRRRCF
Equine	type	1	GLHGVKRRRCFY
Murine	type	1	GLQPIKRRRCF
Ovine	type	287	GVNFNTRRYCY
Avian	type	127	GVRYGSQRYCY
Avian	type	1	GVATATRRMCY
Avian	type	10	GVNVSSRRLCY

Fig 4.03 C-terminal pVI Alignment. Alignment of the C-terminal sequences of pVI from 14 serotypes of the Adenovirus. Equine type 1 and avian type 10 were obtained from translation of DNA upstream from hexon gene where pVI is usually located. Sequences were taken from N.C.B.I. ENTREZ Browser.

The exact role of the KRRR sequence during Adenoviral infections is not clear. There are several theories that one could propose at this early stage of the investigation to assign a function to this motif.

Conceivably since the activating peptide is derived from the last 11 amino acids of pVI one may postulate that pVI may be targeted to the nucleus by this sequence. Figure 4.03 is a compilation of the known and putative activating peptide sequences to date in 14 serotypes of Adenoviruses and illustrates the conservation of the KRRR motif throughout mammalian serotypes. However, pVI contains a second putative NLS identified through a

homologous sequence found in Ad 2 E1a proteins (Hong and Engler, 1991). Is it uncommon to have two nuclear localisation signals with in one protein? There are other instances in which two or more NLS are found within a protein either as repeats of the same sequence or as an unrelated sequence. It is generally accepted that the effects by presence of multiple NLS's are additive (Lanford *et al.*, 1986; Dworetzky *et al.*, 1988; Lanford *et al.*, 1990). In yeast, MATα2 has two NLS, one of which is amino terminally located. Removal of the amino terminus does not affect the nuclear localisation of the protein, however, removal of the internal homeodomain significantly reduces the efficiency of transport (Hall *et al.*, 1990) suggesting that they fulfil functionally distinct roles, possibly during the nuclear translocation pathway. The primary structural considerations here are very similar to those observed within pVI.

Clearly the solution of pVI crystal structure would prove advantageous in defining where the C-terminal region lies in the tertiary structure. Because of the basic nature of the residues involved, and the fact that the protease can gain access to cleave the pVI to iVI one would expect this region to be exposed. Crystallisation of the whole precursor polypeptide could prove difficult due to the toxic nature of the first 48 amino acid which were responsible for cell lysis during expression (Mathews and Russell, 1994).

Suggesting that the protease interacts with pVI in the cytoplasm goes against the generally accepted idea that the protease only interacts and becomes active within the assembling virus particle (Cotten and Weber 1995). The protease and pVI are both coded in the L3 transcript and could conceivably interact in the cytoplasm during translation. If they did interact in the cytoplasm the protease would be able to achieve the active conformation prior to entering the nucleus and potentially prematurely process the

precursor polypeptide before they are all in place. So it is unlikely that the interaction of the protease with activating peptide is how the Adenovirus targets the protease to the nucleus.

If it is the case that pVI and the protease remain apart until they are within assembling virions then what is the implication of an NLS activating peptide? Presumably if the protease enters the nucleus via simple diffusion alone what is to stop the protease leaving the nucleus, which at some point it must since it is responsible for the cleavage of cytokeratin. The second possibility is that the interaction between pVI and protease directs the protease to the nucleus at a specific time during infection. If we compare (fig 4.01) the HIV tat protein NLS with the protease/activating peptide complex we can see the striking similarity they share in both shape and size. The NLS's are depicted in blue, tat is comprised of $RKKR^{-52}$ the activating peptide $KRRR^{-9}$. The idea that a protein recruits an NLS is novel, although some proteins have been reported to "piggy back" other proteins to gain entry to the nucleus.

The activating peptide could fulfil two roles here. The interaction of the activating peptide with the protease is thought to occur transiently, i.e. when the protease has cleaved the substrate, the activating peptide and protease dissociate. In the absence of activating peptide the protease is free to diffuse out of the nucleus, however if in the bound and active form the protease would preferentially be housed within the nucleus since any complex which diffused out of the nucleus would be a target back to the nucleus by the nuclear transport machinery. To date no protein involved with nuclear transport has been shown to contain a consensus protease cleavage site. Here the activating peptide would be acting in such a way as to force the equilibrium in the favour of the active complex remaining in the nucleus.

The second role would be an indirect role. It has been shown that the protease can cleave proteins in the absence of activating peptide (Webster *et al*, 1993, Mathews and Russell 1994, 1995) (discussed later). Cleavage in the absence of the activating peptide may occur at a preferential site and may explain the cleavage kinetics of the cytokeratin network at the latter stages of infection. Clearly if there is no control over the diffusion of the active protease premature cleavage of the cytokeratin could occur which would be suicide for the virus.

Alternatively this motif may be a coincidence and the importance of the conserved KRRR residues throughout human Adenovirus types may be necessary for achieving the active conformation of the protease. Often in over expression studies, the fact that the protein is over may actually lead to the protein or cell it is expressed in behaving atypically.

An interesting extension of this work would be to assess if the activating peptide sequence that targeted the chimer to the nucleus would be capable acting as penetratins (Peck and Isakce, 1998). Could one detect the chimeric protein within HeLa cells if it was added to the culture media.

4.2 Cleavage of the cytoskeleton

In an attempt to see if protease could be directed to the nucleus by the addition of activating peptide to the media of cultured cells the protease, and two mutants C104A, C122A were constructed. Cells were transfected with 4μg of DNA and left for 16 h before being fixed and prepared for immunoflourescence. The OA10B monoclonal antibody (protease) was kindly donated by T Vaughan, however, the antibody produced a very high amount of background fluorescence and results were inconclusive.

One observation which was surprising was that the protease appeared to be toxic to the cells. Expression of the control plasmid did not induce cell death,

however expression of Wt protease and C104A did lead to cell death. The expression of C122A did lead to limited cell death although not to the extent of that observed with the Wt and C104A. Under oil immersion microscopy cells appeared rounded, having lost their polarity and the filamentous appearance associated with the normal phenotype.

Duplicate assays were performed and the cells probed with an anticytokeratin monoclonal antibody. Figure 3.14 illustrates the results of the expression study. The cytokeratin network appeared to be being degraded in the presence of Wt protease and the mutant C104A in the absence and presence of activating peptide. Duplicate transfections were carried out and cells were fractionated and extracts prepared for SDS PAGE and Western blot analysis with polyclonal antisera to the protease. Protease expression was confirmed and its was presumed that the cytokeratin reorganisation may have been a consequence of protease expression. To confirm this observation our lab is currently trying to establish a cytokeratin cleavage assay system. However, one would expect this to be a challenging prospect given the insoluble nature of the cytokeratins. Cytokeratin K18 and K7 both contain consensus protease cleavage sites. In the case of K18 cleavage is reported at the LXGX↓G site and not at the IXGG↓X site (Chen *et al.*, 1993), K7 contains two LXGX↓G site (Devereux *et al.*, 1984). The fact that the protease is thought to be active in this study in the absence of activating peptide is not novel. Webster *et al* (1993) showed that the protease was capable of cleaving pTP to iTP in SF9 cells in the absence of activating peptide. Also the processing of pVI to iVI must occur in the absence of activating peptide. An interesting phenomenon of all these cleavages is that they take place at the consensus XGXG site (Webster *et al.*, 1997; Mathews and Russell 1994 and 1995) in the absence of activating peptide.

Both Webster *et al* (1989) and Diouri *et al* (1996) reported that the protease illustrated a preference for cleavage at XGXG sites. Considering the preferential cleavage site evidence with the cleavages in the absence of activating peptide one could develop a theory about the control of the viral life cycle.

The cleavage of cytokeratin during lytic infections could occur in the absence of activating peptide. The cleavage sites in cytokeratins K7 and K18 are both the XGXG and their is no evidence that the XGGX site in K18 is cleaved. The activating peptide which contains an NLS could be exclusively retained in the nucleus of the infected cell and thus the only way the virus could break out of the cell in a system like this would be if the protease cleaved the cytokeratin in the absence of activating peptide at a preferential site. If this was the case protease not complexed with activating peptide would be capable of diffusing away from the nucleus to perform this function.

4.3 Development of an inducible in vivo *nuclear translocation assay*

Could one exploit the unique nature of the protease/activating peptide interaction in order to develop a temporally inducible *in vivo* protein expression system. If a protein chimer between pyruvate kinase and the region of the protease which interacts with the activating peptide could be directed to the nucleus of a cell simply with the addition of the activating peptide to the culture media we would have a very powerful experimental system for understanding the role of nuclear proteins. Ishibashi *et al* (1994) developed a recombinant plasmid carrying a chimeric cDNA which encoded for a fusion protein of the oestrogen receptor and human O6-methylguanine-DNA methyltransferase (ER:MGMT). The construct was introduced into MGMT deficient cell lines and they developed a controlled nuclear translocation with the addition of oestrogen. Since then a variety of groups

have used this technique to study the function of nuclear proteins (Rehberger *et al.*, 1997; Zhu et al., 1998). One draw back of this system is the size of the receptor fused to the target protein. These experiments detailed here were designed to produce an alternative method with a smaller foreign fusion region.

From the crystal structure 3 regions were selected as the likely minimum region of interaction between the protease and activating peptide. The DNA sequence corresponding to amino acids 103-112, 98-114 and 98-148 of the protease were cloned C-terminally of pyruvate kinase and expressed in $HeLa^{CD}$ cells.

Activating peptide was added to the media of cultured cells to see if the construct would be directed to the nucleus. In the initial experiments all three constructs were detected to some extent in the nucleus (fig. 3.20). However, in subsequent experiments we were unable to repeat this. The second and third trials were found to have been conducted with reduced activating peptide. This was thought to be one reason for not being able to repeat the result. Cabrita *et al* (1997) had shown that the oxidised form of the peptide was the favoured species compared to reduced peptide for binding the protease. The assay was repeated with oxidised peptide, but nuclear localisation was not observed.

One possibility considered was that the peptide was not penetrating the cell membrane in the subsequent experiments. In order to confirm that the activating peptide could enter the cell we attempted to label the activating peptide with Flourescein. However, we were unable to find a species that corresponded to this modified peptide using mass spectroscopy.

What next? Preliminary investigations should look into the ability and efficiency of the transport of activating peptide across the cell membrane.

There is no conceivable reason not to assume that the activating peptide crossed the membrane if we consider the use of penetratin (Peck and Isacke, 1998) and the results reported by Rancourt *et al* (1995) in the rescue of ts 1 viral infections. Penetratin is a 16 amino acid peptide corresponding to the third helix of Antennapedia homeodomain which not only efficiently translocates across mammalian plasma membranes but can also act as a carrier for internalisation of cargo peptides (Derossi *et al.*, 1994, 1996; Theodore *et al.*, 1995; Hall *et al.*, 1996; Peck and Isacke 1998). The penetratin sequence (RQIKIWFQNRRMKWKK) is basic and not too dissimilar to that of the activating peptide. Rancourt showed that ts1 Adenoviral infections could be rescued with the addition of activating peptide to the culture media. Both pieces of evidence support the theory that the activating peptide should be able to traverse the cell membrane.

Micro-injection of the peptide directly into cells expressing the constructs could also prove conclusive in either confirming or dismissing that the translocation observed in the initial studies could be induced by the addition of activating peptide.

4.4 Protease activity and binding assay studies

In an attempt to further our understanding of the protease/activating peptide interaction and the importance of the individual residues in the protease activity, binding assay studies were performed.

It was important to develop a system which allowed us to assess if the inability of a peptide to activate the protease was a result of an inability of bind the protease, or the ability to bind and strap the protease into an inactive conformation.

The system employed with further development would provide a widely applicable qualitative test to asses the peptides' affinity for the protease. Here

we analysed the data as a double log plot of peptide concentration against the ratio of protease to lysozyme in the assays. A preliminary Scatchard analysis was also perfomed, clearly further repetition of the experiments are required, as well as developing the paramenters presented here therefore the data from the Scatchard analysis was not considered here.

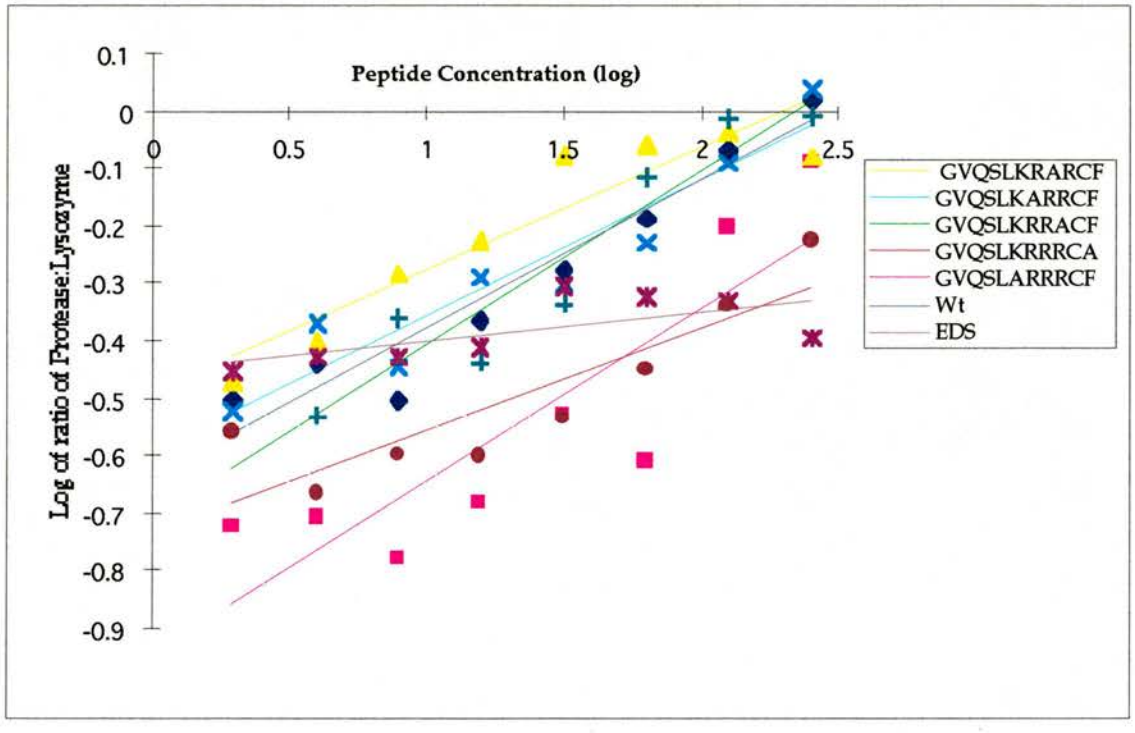


Fig 4.04 Relative Binding Affinity of Various Peptides. The graph illustrates the relative binding affinities of the peptides used throughout the study. The graph shows that most of the peptides intercept the x axis around the same point. The two outliers were EDS and RCA peptides and were also shown to activate the protease to a lesser extent.

Three peptides did not cluster around the log value of -0.3, they were EDS, RCA and ARRRCF (figure 4.04). However, the gradient of the plot for the peptide ARRRFC is approximately the same as that observed in the majority of peptide assays. EDS and RCA activated the protease to the weakest extent and were expected to interacted more weakly. The results of the affinity assays were then compared to those from the activation studies (figure 4.05). We shall consider the peptides in turn discussing results from the affinity binding assay and activation studies. We should bear in mind that the binding

affinity assay at this earlier stage in development is only an indicative way to assess the extent to which peptides were able to interact with the protease.

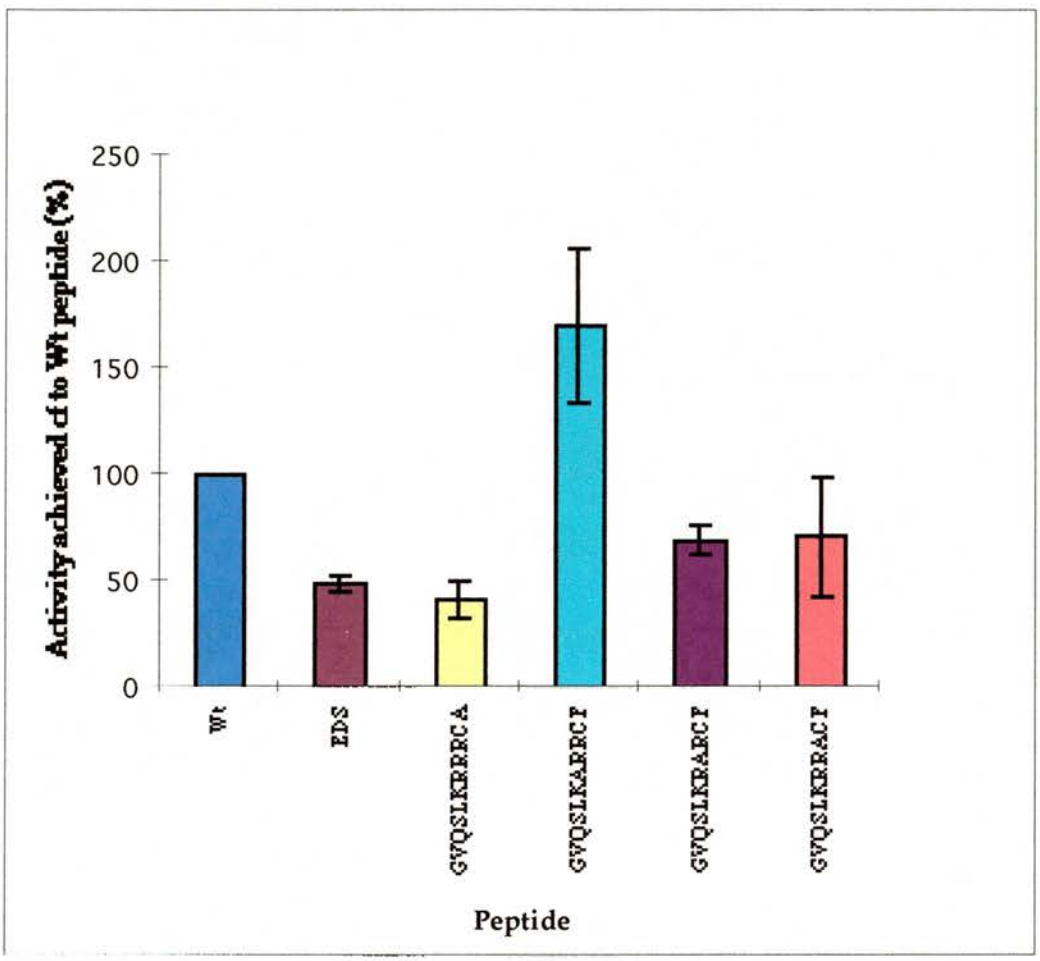


Fig 4.05 Relative Protease Activity for Selected Peptides. From the graph we can see that GVQSLKARRCF activated the protease to a greater extent than the Wt activating peptide, all other peptides lead to a reduction in the protease's activity. GVQSLKRRRCA and EDS activated the protease to the least extent. GVQSLKRARCF and GVQSLKRRACF activated the protease to similar extents.

Scatchard analysis was performed, however, the data presented here is at a very preliminary stage and would benefit from further repetitions to be confident of the results. We have demonstrated a model system to study the binding affinity of various peptides to the protease.

4.4.1 GVQSLLKRRRCF

GVQSLLKRRRCF peptide was designed to investigate the flexibility of the protease/activating peptide groove. If one could, in-effect, wedge a larger than usual peptide in the groove and prevent the interaction with Wt

activating peptide one potentially would have the effect of a specific antiviral agent. The rationale behind this system was derived from the similarity with MHC 1 antigenic peptide presentation. The binding groove usually presents peptides of 8-10 amino acids in length. Activation studies showed that protease was not activated by the peptide which was probably as a result of its inability to interact with the protease illustrated in the binding assays. In this assay the protease was not resolved to one band with the addition of a 150 molar excess of activating peptide. Protease appeared as a smear characteristically observed when peptides were not added to protease samples in the native gel system employed. The fact that the peptide did not interact with the protease was not all that surprising since the MHC class one molecule is designed to take a variety of size of antigenic peptides, where as the Adenovirus protease, with the exception Equine type 1, has only ever show preference for a peptide length of 11 amino acids.

4.4.2 VEGGS

Initially we investigated the effect of activating the protease with the peptide sequence VEGGS (Diouri *et al.*, 1996) which was identified from phage display pull down library. In our system, with detection of substrate digestion via capillary electrophoresis we could not repeat the results observed by the authors who used a more sensitive fluorogenic system with R110 labelled substrates to measure substrate cleavage. Again to confirm if the inactivation of the protease was due to the inability of the peptide to bind the protease we prepared the affinity binding assay with a doubling serial dilution of VEGGS. One other important difference between the assays was the source of the protease. The authors purify their protease from virus infected Hep 2 monolayers which may allow a post-translational modification which does not occur in our recombinant expression system in *E. coli*.

4.4.3 EDS

Could the activating peptides from the most distant related serotypes activate the protease of Adenovirus type 2? Avian 127 or EDS peptide (GVRYGSQRYCY) was used to test this hypothesis. The EDS peptide was shown to activate the protease to approximately 42% of that observed in the Wt activating control assay. This result was surprising because of the difference observed between their primary structures. Only 4 of the highly conserved residues are present in this peptide. C10 of this peptide is surrounded by 2 Y residues which may increase the stability of the activating peptide/protease complex in EDS (Morgan *et al.*, 1978 and 1980). EDS protease does not have P at 137, therefore one would expect the tertiary structure to be different and consequently explain the difference in the primary structures of the peptide. It would be interesting to take this work a stage further and investigate the effect of the Adenovirus type 2 activating peptide on the EDS protease. Such assays may also shed light on the importance of P137. The fact that ts 1 infections were rescued by the addition of an excess of activating peptide (Rancourt *et al.*, 1995) to monolayer media presents a variety of possibly explanations. The author explained the result as a function of the solubility of the protease. However with the results presented above possibly implicating the activating peptide as a nuclear localisation signal for the protease one could draw an alternative conclusion. The excess activating peptide may be required to force the equilibrium to the formation of the activating peptide/protease in the absence of P137. Although the authors did report that in controls without excess activating peptide protease was detected in the nucleus of cells, and not packaged within virus particles. This result would be consistent with the idea that the activating

peptide was responsible for the interaction with viral DNA during viral packaging.

4.4.4 GVQSLKRRRCA

The effect of mutating F11 to a lad to partial activity of approximately 42 % of that produced with Wt activating peptide. The peptide binding assay showed only limited binding ability with RCA when binding was detected it was, only present in the first and second lane (results not shown). Limited activation was expected with this as residue 11 of the activating peptide is always an aromatic residue (fig 4.03). Figure 4.06 shows the position of F11 and R8 of the activating peptide sandwiching the disulphide bond. This stacking of pi orbitals of electrons may be important in protease/activating peptide complex formation and control of activation. Morgan *et al* (1978 and 1980) in a survey of 22 globular proteins found 211 Sulphur-pi orbital interactions. The authors found an abundance of R and Y residues in close proximity to cysteine. The importance of these sulphur-aromatic interactions in globular proteins is thought to be concerned with the stabilisation of the disulphide interchanges. The presence of positive side chains such as R (211 occasions) was thought to enhance the stability of the interaction in an indirect manner.

We see a similar phenomenon in the crystal structure of the protease with activating peptide. F11, R8 and R103 appear to sandwich the S-S bond between C104 of the protease and C10 of the activating peptide. If we refer to the pile up of the protease, figure 1.07, and activating peptides, figure 4.03, we see that these residues are highly conserved throughout the differing serotypes. The importance of this observation is limited at the moment. Further investigation with point mutations at these residues would be required to confirm this hypothesis.

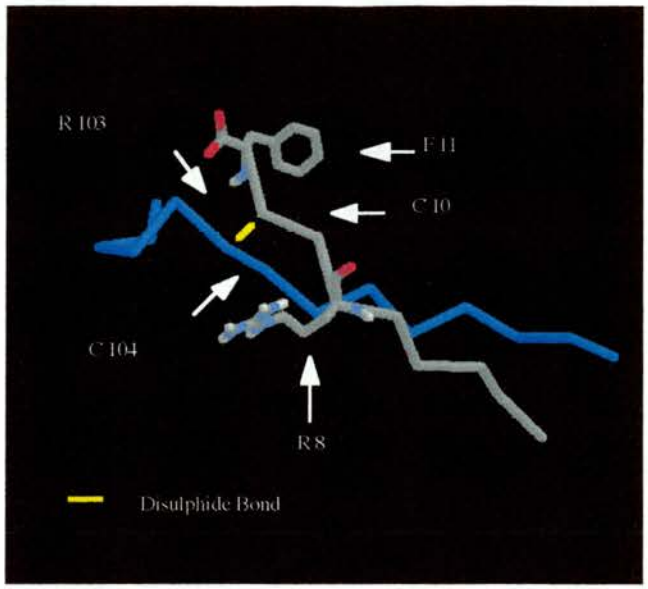


Fig 4.06 Pi Stacking around Disulphide Bonds. The figure displays the aromatic ring structure of F11 and R8 side chain. We can clearly see that the likely position of the pi orbitals would sandwich the s-s bridge formed between C104 of the protease and C10 of the peptide. Mutation of F11 to A resulted in activity of 42% of that of the Wt protease. Mutation of R8 lead to approximately a 40 % reduction in the activity of the protease.

The stability of the internal disulphide bond of the oxidised activating peptide form must be overcome by the environment offered by the protease to the activating peptide. Perhaps the environment around C104 is important in facilitating the rearrangement of the disulphide bonds and reducing the oxidised peptide favouring the protease/activating peptide complex formation. R8, F11 and R103 may offer such an environment, and in part explain why they are highly conserved.

4.4.5 The Importance of the four Basic Amino Acid Residues in Activation and Binding of the Protease

The activation of the protease achieved with basic amino acid mutant peptides was studied for two reasons. Firstly to investigate the importance of the positively charged amino acids in the activation of the protease and secondly to see if modification to this sequence would increase the affinity of the activating peptide/protease complex formation. If the activating peptide was the signal for the nuclear localisation /retention then a mutant peptide which

bound more competitively to the protease would be an attractive starting point for the development of a protease inhibitor.

The initial experiments were centred around K6 and R7 since the substitution of these residues to A abolished the nuclear localisation of pyruvate kinase. GVQSLARRRCF peptide has yet to be fully characterised in respect to it activation properties and relative binding affinity. Although initial results suggest that it was the least active of all the basic amino acid substitution peptides.

GVQSLKARRCF peptide produced a surprising result in that the activity achieved relative to that of the Wt peptide GVQSLKRRRCF (100%) was 170%. The binding affinity observed was also greater than that observed with the Wt peptide. How could this result be explained and what are the implications of the result? Looking at the crystal structure R7 is opposite L 107 of the protease (illustrated in figure 4.07). As the activation mechanism is not fully characterised it is feasible that the juxtaposition of these two residues acts as a destabilising switch between the protease and activating peptide. When the positively charge group R 7 is substituted to A, a hydrophobic group, the interaction between the protease and activating peptide would appear to become more stable. It may explain the observation that the binding of activating peptide is reversible and thus the need for a large molar excess was required to push the equilibrium in the heterodimer formation direction (Webster *et al.*, 1993; Mangel *et al.*, 1993; Jones *et al.*, 1996; Cabrita *et al.*, 1997). In this experiment the introduction of A opposite L could stabilise the interaction so that after the active protease complex interacts with substrate the activating peptide remains bound rather than being released.

Although the investigation needs to be extended there is a potential application of this result. It has been demonstrated that the addition of

activating peptide to Adenovirus infected Hep 2 cells was detrimental to viral reproduction, perhaps the addition of this peptide could result in more potent effects (Rancourt *et al.*, 1995; Ruzindana-Umunyana *et al.*, 2000). One possible explanation for the observed inhibition of viral growth is that the premature activation of the protease lead to processing of the precursor polypeptides before they were able to form a stable virion conformation. Another consequence of a prematurely active protease could be the degradation of the cytoskeleton inducing cell death before virions had had time to assemble.

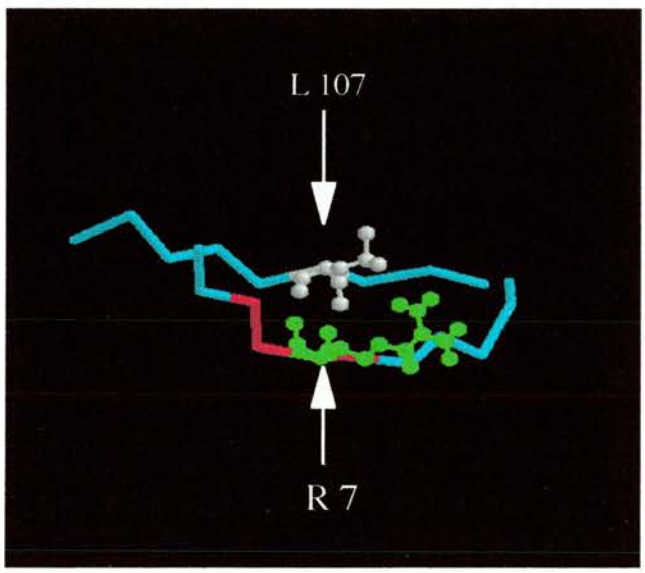


Fig 4.07 Illustrates the interaction of the side chains of L107 of the protease and R7 of the activating peptide. This interaction is believed to be responsible for the reversible binding of the activating peptide during the activation of the protease. Mutation of R7 to A lead to a more active protease and a greater binding affinity.

Closer inspection of the crystal structure of the protease/pVI dimer shows the side chain of R8 bending back on itself below a F11. We investigated here if R8 could be implicated in pi stacking about the disulphide bond. GVQSLKRARCF peptide was shown to activate the protease to 69% of that observed in duplicate GVQSLKRRRCF activated assays. The loss of activity

could be explained by the peptides failure to interact with the protease to the same extent as that of the Wt activating peptide. Considering the affinity assay results the extent to which GVQSLKRARCF was reduced as compared to GVQSLKRRRCF could be important. The decreased ability of the protease/peptide interaction could be explained by an alteration to the structure around the interaction site not favourable to reduce the oxidised form of the peptide. This is the only completely conserved R residue across all serotypes of pVI both identified and predicted from DNA sequencing results. This result implies due to the loss of only approximately 30 % of activity that the side chain bending shown in the crystal structure was an artefact of crystal packing.

Finally substitution of R9 resulted in protease activation to 72% of that of the conventional peptide GVQSLKRRRCF. R9 is the least conserved of the R within the C- terminal pVI region and is of debatable importance in the nuclear localisation of the pyruvate kinase construct.

4.5 Crystallisation Trials

In order to understand more about the inactive conformation of the protease we attempted to crystallise the enzyme several times. McGrath et al (1996) had previously co-crystallised the protease with the 11 amino acid activating peptide. Trials were initiated with 2µl of protease (5.6 mg ml) under a variety of conditions using the magic 50 and 48 kits (Hampton Research) as well as the conditions used by McGrath. No crystal were evident even after 6 months. We believe that the protease has an inherent dynamic wobble which would explain why no crystals formed. Further evidence of this dynamic nature was presented by the native gel system. The fact that the protease appeared as several bands in the absence of the activating peptide would suggest that the molecule existed in a variety of conformations which is not

conducive to ordered crystal packaging necessary for successful crystallisation.

4.6 Developing a Single Step Purification Protocol for Adenovirus type 2 Protease

Prior to the development of the single step purification protocol the whole procedure took about 2 days from fractionating the cell pellet to recovering pure protease. After several trial purifications using the DEAE-Sepharose column it was apparent that the protease consistently eluted between fractions 18-30 (fig 3.02). Each time the experiment was conducted the eluted fraction would have to be separated by SDS PAGE and protease identified using Western blot. The second purification would then be conducted using the protease containing fractions. Theses fractions would be pooled and applied to consecutive CM-Sepharose and heparin-Sepharose columns before the protease containing fractions were identified as above.

As a result of the consistent nature of the elution profile we were able to developed a method whereby the protease was eluted from the CM column within 90 minutes. This system had the advantage of being rapid and therefore the protease had less time to become oxidised which was shown to reduce its inactivity (Cabrita *et al.*, 1997). This gradual oxidation was believed to be in some part responsible for the observed differences in activity of the protease between batches (results not shown). The development of a reproducible single step procedure limited the amount of time the protease was exposed to oxidising conditions and thus gave a more consistent quality of protease.

Summary

We have demonstrated that the activating peptide, which originates from the C-terminus of pVI, contains an NLS which is sufficient to direct a chimeric protein to the nucleus. The minimum sequence to confer nuclear localisation

is KRRR. Whether this is of physiological significance during infection is unclear at present.

Interestingly these same four amino acids are of debatable importance in the activation mechanism of the protease with the exception of mutation of R7, which resulted in a 70 % increase in activity. These basic amino acids are highly conserved throughout the mammalian serotypes and begs the question why? Similarly what would happen if a larger hydrophobic group was introduced into the R7 position? If this residue is important in the activation mechanism we may have the target for the development of an antiviral agent.

Clearly the regulation of the protease is important during infection. Premature cleavage of the pre-cursor polypeptides or the cytokeratin network is lethal to the virus. Evidence presented here that the protease can cleave the cytokeratin in the absence of activating peptide suggested a finer control of activation. Certainly it would suggest that cleavage at the two consensus cleavage sites may favour LXGX↓G in the absence of activating peptide, and is supported by the cleavages reported in pVI and pTP in the absence of activating peptide.

5. References

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

Adam, S.A. and Gerace, L. (1991) Cytosolic proteins that specifically bind nuclear location signals are receptors for nuclear import. *Cell*, **66**, 837-847.

Akusjarvi, G., Alestrom, P., Pettersson, M., Lager, M., Jornvall, H. and Pettersson, U. (1984) The gene for the adenovirus 2 hexon polypeptide. *J Biol Chem*, **259**, 13976-13979.

Anderson, C.W. (1990) The proteinase polypeptide of adenovirus serotype 2 virions. *Virology*, **177**, 259-272.

Anderson, C.W. (1993) Expression and purification of the adenovirus proteinase polypeptide and of a synthetic proteinase substrate. *Protein Expr Purif*, **4**, 8-15.

Anderson, C.W., Baum, P.R. and Gesteland, R.F. (1973) Processing of adenovirus 2-induced proteins. *J Virol*, **12**, 241-252.

Anderson, C.W., Schmitt, R.C., Smart, J.E. and Lewis, J.B. (1984) Early region 1B of adenovirus 2 encodes two coterminal proteins of 495 and 155 amino acid residues. *J Virol*, **50**, 387-396.

Atherton, E. and Sheppard, R. C. (1989). Solid Phase peptide synthesis: A practical approach.In "Practical Approach Series" pp 203. *IRL Press*, Oxford.

Babich, A., Feldman, L.T., Nevins, J.R., Darnell, J.E. and Weinberger, C. (1983) Effect of adenovirus on metabolism of specific host mRNAs: transport control and specific translational discrimination. *Mol Cell Biol*, **3**, 1212-1221.

Beltz, G.A. and Flint, S.J. (1979) Inhibition of HeLa cell protein synthesis during adenovirus infection. Restriction of cellular messenger RNA sequences to the nucleus. *J Mol Biol*, **131**, 353-373.

Bergelson, J.M., Cunningham, J.A., Droguett, G., Kurt-Jones, E.A., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L. and Finberg, R.W. (1997) Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science*, **275**, 1320-1323.

Bergelson, J.M., Krithivas, A., Celi, L., Droguett, G., Horwitz, M.S., Wickham, T., Crowell, R.L. and Finberg, R.W. (1998) The murine CAR homolog is a receptor for coxsackie B viruses and adenoviruses. *J Virol*, **72**, 415-419.

Bhatti, A.R. and Weber, J. (1978) Protease of adenovirus type 2. In vitro processing of core protein. *Biochem Biophys Res Commun*, **81**, 973-979.

Bhatti, A.R. and Weber, J. (1979) Protease of adenovirus type 2: partial characterization. *Virology*, **96**, 478-485.

Blumenthal, R., Seth, P., Willingham, M.C. and Pastan, I. (1986) pH-dependent lysis of liposomes by adenovirus. *Biochemistry*, **25**, 2231-2237.

Bodnar, J.W., Hanson, P.I., Polvino-Bodnar, M., Zempsky, W. and Ward, D.C. (1989) The terminal regions of adenovirus and minute virus of mice DNAs are preferentially associated with the nuclear matrix in infected cells. *J Virol*, **63**, 4344-4353.

Bond, J.S. and Butler, P.E. (1987) Intracellular proteases. *Annu Rev Biochem*, **56**, 333-364.

Bonner, W.M. (1975) Protein migration into nuclei. I. Frog oocyte nuclei in vivo accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *J Cell Biol*, **64**, 421-430.

Borrelli, E., Hen, R. and Chambon, P. (1984) Adenovirus-2 E1a products repress enhancer-induced stimulation of transcription. *Nature*, **312**, 608-612.

Boudin, M.L., Moncany, M., D'Halluin, J.C. and Boulanger, P.A. (1979) Isolation and characterization of adenovirus type 2 vertex capsomer (penton base). *Virology*, **92**, 125-138.

Boulanger, P.A. and Blair, G.E. (1991) Expression and interactions of human adenovirus oncoproteins. *Biochem J*, **275**, 281-299.

Breeuwer, M. and Goldfarb, D.S. (1990) Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. *Cell*, **60**, 999-1008.

Bridge, E., Medghalchi, S., Ubol, S., Leesong, M. and Ketner, G. (1993) Adenovirus early region 4 and viral DNA synthesis. *Virology*, **193**, 794-801.

Brown, D.T., Westphal, M., Burlingham, B.T., Winterhoff, U. and Doerfler, W. (1975) Structure and composition of the adenovirus type 2 core. *J Virol*, **16**, 366-387.

Cabrita, G., Iqbal, M., Reddy, H. and Kemp, G. (1997) Activation of the adenovirus protease requires sequence elements from both ends of the activating peptide. *J Biol Chem*, **272**, 5635-5639.

Cabrita, G., (1997) The mechanism of activation of the adenovirus type 2 protease. **PhD Thesis**, University of St. Andrews.

Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell*, **43**, 439-448.

Challberg, M.D. and Kelly, T.J. (1989) Animal virus DNA replication. *Annu Rev Biochem*, **58**, 671-717.

Challberg, M.D. and Rawlins, D.R. (1984) Template requirements for the initiation of adenovirus DNA replication. *Proc Natl Acad Sci U S A*, **81**, 100-104.

Chatterjee, P.K., Bruner, M., Flint, S.J. and Harter, M.L. (1988) DNA-binding properties of an adenovirus 289R E1a protein. *EMBO J*, **7**, 835-841.

Chatterjee, P.K. and Flint, S.J. (1987) Adenovirus type 2 endopeptidase: an unusual phosphoprotein enzyme matured by autocatalysis. *Proc Natl Acad Sci U S A*, **84**, 714-718.

Chatterjee, P.K., Vayda, M.E. and Flint, S.J. (1986a) Adenoviral protein VII packages intracellular viral DNA throughout the early phase of infection. *EMBO J*, **5**, 1633-1644.

Chatterjee, P.K., Vayda, M.E. and Flint, S.J. (1986b) Identification of proteins and protein domains that contact DNA within adenovirus nucleoprotein cores by ultraviolet light crosslinking of oligonucleotides 32P-labelled in vivo. *J Mol Biol*, **188**, 23-37.

Chen, P.H., Ornelles, D.A. and Shenk, T. (1993) The adenovirus L3 23-kilodalton proteinase cleaves the amino-terminal head domain from cytokeratin 18 and disrupts the cytokeratin network of HeLa cells. *J Virol*, **67**, 3507-3514.

Chiu, C.Y., Mathias, P., Nemerow, G.R. and Stewart, P.L. (1999) Structure of adenovirus complexed with its internalization receptor, alphavbeta5 integrin. *J Virol*, **73**, 6759-6768.

Cotten, M. and Weber, J.M. (1995) The adenovirus protease is required for virus entry into host cells. *Virology*, **213**, 494-502.

Cuillel, M., Cortolezzis, B., Chroboczek, J., Langowski, J., Ruigrok, R.W. and Jacrot, B. (1990) Purification and characterization of wild-type and ts 112 mutant protein IIIa of human adenovirus 2 expressed in Escherichia coli. *Virology*, **175**, 222-231.

De Robertis, E.M., Longthorne, R.F. and Gurdon, J.B. (1978) Intracellular migration of nuclear proteins in Xenopus oocytes. *Nature*, **272**, 254-256.

Defer, C., Belin, M.T., Caillet-Boudin, M.L. and Boulanger, P. (1990) Human adenovirus-host cell interactions: comparative study with members of subgroups B and C. *J Virol*, **64**, 3661-3673.

Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G. and Prochiantz, A. (1996) Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. *J Biol Chem*, **271**, 18188-18193.

Derossi, D., Joliot, A.H., Chassaing, G. and Prochiantz, A. (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem*, **269**, 10444-10450.

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

Ding, J., McGrath, W.J., Sweet, R.M. and Mangel, W.F. (1996) Crystal structure of the human adenovirus proteinase with its 11 amino acid cofactor. *EMBO J*, **15**, 1778-1783.

Dingwall, C. and Laskey, R.A. (1990) Nucleoplasmin: the archetypal molecular chaperone. *Semin Cell Biol*, **1**, 11-17.

Diouri, M., Girouard, G.S., Allen, C.M., Sircar, S., Van Lier, J.E. and Weber, J.M. (1996a) New stimulation ligand of the adenovirus 2 protease. *Virology*, **224**, 510-516.

Diouri, M., Keyvani-Amineh, H., Geoghegan, K.F. and Weber, J.M. (1996b) Cleavage efficiency by adenovirus protease is site-dependent. *J Biol Chem*, **271**, 32511-32514.

Dworetzky, S.I., Lanford, R.E. and Feldherr, C.M. (1988) The effects of variations in the number and sequence of targeting signals on nuclear uptake. *J Cell Biol*, **107**, 1279-1287.

Dyson, N., Guida, P., McCall, C. and Harlow, E. (1992) Adenovirus E1a makes two distinct contacts with the retinoblastoma protein. *J Virol*, **66**, 4606-4611.

Dyson, N. and Harlow, E. (1992) Adenovirus E1a targets key regulators of cell proliferation. *Cancer Surv*, **12**, 161-195.

Engels , B (1992). Amplify, Madison.

Evans, R.M., Fraser, N., Ziff, E., Weber, J., Wilson, M. and Darnell, J.E. (1977) The initiation sites for RNA transcription in Ad2 DNA. *Cell*, **12**, 733-739.

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

Ferguson, B., Krippl, B., Andrisani, O., Jones, N., Westphal, H. and Rosenberg, M. (1985) E1a 13S and 12S mRNA products made in Escherichia coli both function as nucleus-localized transcription activators but do not directly bind DNA. *Mol Cell Biol*, **5**, 2653-2661.

FitzGerald, D.J., Padmanabhan, R., Pastan, I. and Willingham, M.C. (1983) Adenovirus-induced release of epidermal growth factor and pseudomonas toxin into the cytosol of KB cells during receptor-mediated endocytosis. *Cell*, **32**, 607-617.

Fredman, J.N. and Engler, J.A. (1993) Adenovirus precursor to terminal protein interacts with the nuclear matrix in vivo and in vitro. *J Virol*, **67**, 3384-3395.

Furcinitti, P.S., van Oostrum, J. and Burnett, R.M. (1989) Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography. *EMBO J*, **8**, 3563-3570.

Gorlich, D. and Mattaj, I.W. (1996) Nucleocytoplasmic transport. *Science*, **271**, 1513-1518.

Gorlich, D., Pante, N., Kutay, U., Aebi, U. and Bischoff, F.R. (1996) Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO J*, **15**, 5584-5594.

Gossen M, Bujard H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci*, **89** :5547-5551.

Gossen M, Freundlieb S, Bender G, Muller G, Hillen W, Bujard H (1995) Transcriptional activation by tetracyclines in mammalian cells. *Science*, **268**,1766-1769.

Grable, M. and Hearing, P. (1992) cis and trans requirements for the selective packaging of adenovirus type 5 DNA. *J Virol*, **66**, 723-731.

Green, M and Daesch, E. (1961) Biochemical studies on adenovirus multiplication IV. Isolation, puricication and chemical analysis of adenovirus. *Virology* **20**, 199-207.

Greber, U.F. (1998) Virus assembly and disassembly: the adenovirus cysteine protease as a trigger factor. *Rev Med Virol*, **8**, 213-222.

Greber, U.F., Webster, P., Weber, J. and Helenius, A. (1996) The role of the adenovirus protease on virus entry into cells. *EMBO J*, **15**, 1766-1777.

Greber, U.F., Willetts, M., Webster, P. and Helenius, A. (1993) Stepwise dismantling of adenovirus 2 during entry into cells. *Cell*, **75**, 477-486.

Grierson, A.W., Nicholson, R., Talbot, P., Webster, A. and Kemp, G. (1994) The protease of adenovirus serotype 2 requires cysteine residues for both activation and catalysis. *J Gen Virol*, **75**, 2761-2764.

Gruber, W.C., Russell, D.J. and Tibbetts, C. (1993) Fiber gene and genomic origin of human adenovirus type 4. *Virology*, **196**, 603-611.

Halbert, D.N., Cutt, J.R. and Shenk, T. (1985) Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *J Virol*, **56**, 250-257.

Hall, M.N., Craik, C. and Hiraoka, Y. (1990) Homeodomain of yeast repressor alpha 2 contains a nuclear localization signal. *Proc Natl Acad Sci U S A*, **87**, 6954-6958.

Hammarskjold, M.L. and Winberg, G. (1980) Encapsidation of adenovirus 16 DNA is directed by a small DNA sequence at the left end of the genome. *Cell*, **20**, 787-795.

Han, K., Belaiche, D., Moreau, O., and Briand, G. (1985). *Int J Biochem*, **17**, 429-445

Handa, H., Kingston, R.E. and 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. and Weber, J. (1983) Biological and structural studies with an adenovirus type 2 temperature-sensitive mutant defective for uncoating. *Intervirology*, **19**, 213-223.

Harlow, E., Franza, B.R. and Schley, C. (1985) Monoclonal antibodies specific for adenovirus early region 1A proteins: extensive heterogeneity in early region 1A products. *J Virol*, **55**, 533-546.

Harrach, B., Meehan, B.M., Benko, M., Adair, B.M. and Todd, D. (1997) Close phylogenetic relationship between egg drop syndrome virus, bovine adenovirus serotype 7, and ovine adenovirus strain 287. *Virology*, **229**, 302-308.

Hassell, J.A. and Weber, J. (1978) Genetic analysis of adenovirus type 2. VIII. Physical locations of temperature-sensitive mutations. *J Virol*, **28**, 671-678.

Hasson, T.B., Ornelles, D.A. and Shenk, T. (1992) Adenovirus L1 52- and 55-kilodalton proteins are present within assembling virions and colocalize with nuclear structures distinct from replication centers. *J Virol*, **66**, 6133-6142.

Hay, R.T., Freeman, A., Leith, I., Monaghan, A. and Webster, A. (1995) Molecular interactions during adenovirus DNA replication. *Curr Top Microbiol Immunol*, **199**, 31-48.

Hay, R.T. and Russell, W.C. (1989) Recognition mechanisms in the synthesis of animal virus DNA. *Biochem J*, **258**, 3-16.

Hearing, P., Samulski, R.J., Wishart, W.L. and Shenk, T. (1987) Identification of a repeated sequence element required for efficient encapsidation of the adenovirus type 5 chromosome. *J Virol*, **61**, 2555-2558.

Hearing, P. and Shenk, T. (1983a) The adenovirus type 5 E1a transcriptional control region contains a duplicated enhancer element. *Cell*, **33**, 695-703.

Hearing, P. and Shenk, T. (1983b) Functional analysis of the nucleotide sequence surrounding the cap site for adenovirus type 5 region E1a messenger RNAs. *J Mol Biol*, **167**, 809-822.

Hearing, P. and Shenk, T. (1986) The adenovirus type 5 E1a enhancer contains two functionally distinct domains: one is specific for E1a and the other modulates all early units in cis. *Cell*, **45**, 229-236.

Hennache, B. and Boulanger, P. (1977) Biochemical study of KB-cell receptor for adenovirus. *Biochem J*, **166**, 237-247.

Hieda M, Tachibana T, Yokoya F, Kose S, Imamoto N, Yoneda Y. (1999)A monoclonal antibody to the COOH-terminal acidic portion of Ran inhibits both the recycling of Ran and nuclear protein import in living cells. *J Cell Biol* **144,** 645-655.

Hoeffler, W.K., Kovelman, R. and Roeder, R.G. (1988) Activation of transcription factor IIIC by the adenovirus E1a protein. *Cell*, **53**, 907-920.

Hoeffler, W.K. and Roeder, R.G. (1985) Enhancement of RNA polymerase III transcription by the E1a gene product of adenovirus. *Cell*, **41**, 955-963.

Hong, S.S. and Boulanger, P. (1995) Protein ligands of the human adenovirus type 2 outer capsid identified by biopanning of a phage-displayed peptide library on separate domains of wild-type and mutant penton capsomers. *EMBO J*, **14**, 4714-4727.

Hong JS, Engler JA The amino terminus of the adenovirus fiber protein encodes the nuclear localization signal. *Virology* ;**185**, 758-767

Horwitz, M.S., Scharff, M.D. and Maizel, J.V. (1969) Synthesis and assembly of adenovirus 2. I. Polypeptide synthesis, assembly of capsomeres, and morphogenesis of the virion. *Virology*, **39**, 682-694.

Horwitz M. S. (1971) Intermediates in the synthesis of type 2 adenovirus deoxyribonucleic acid. *J Virol*, **8**, 675-683.

Horwitz M. S. (1990a) Adenoviridae and their replication. In: *Fields of Virology Volume 2.* Second edition pp 1679-1722 (B. N. Fields,D. M. Knipe and P. M. Howley, eds) *Raven Press.*

Horwitz M. S. (1990b) Adenoviruses. In: *Fields of Virology Volume 2.* Second edition pp 1723-1742 (B. N. Fields,D. M. Knipe and P. M. Howley, eds) *Raven Press.*

Houde, A. and Weber, J.M. (1987) Sequence of the protease of human subgroup E adenovirus type 4. *Gene*, **54**, 51-56.

Houde, A. and Weber, J.M. (1988a) The primary structure of human adenovirus type 12 protease. *Nucleic Acids Res*, **16**, 7195.

Houde, A. and Weber, J.M. (1988b) Sequence of the human adenovirus type 3 protease. *Nucleic Acids Res*, **16**, 11374.

Houde, A. and Weber, J.M. (1990) Adenovirus proteinases: comparison of amino acid sequences and expression of the cloned cDNA in Escherichia coli. *Gene*, **88**, 269-273.

Huang, J.T. and Schneider, R.J. (1991) Adenovirus inhibition of cellular protein synthesis involves inactivation of cap-binding protein. *Cell*, **65**, 271-280.

Ishibashi, T., Nakabeppu, Y. and Sekiguchi, M. (1994) Artificial control of nuclear translocation of DNA repair methyltransferase. *J Biol Chem*, **269**, 7645-7650.

Izaurralde E, Adam S. (1998)Transport of macromolecules between the nucleus and the cytoplasm. *RNA.*, **4**:351-364

Jans DA, Xiao CY, Lam MH. (2000) .Nuclear targeting signal recognition: a key control point in nuclear transport? *Bioessays*. **22**: 532-544.

Jones, S.J., Iqbal, M., Grierson, A.W. and Kemp, G. (1996) Activation of the protease from human adenovirus type 2 is accompanied by a conformational change that is dependent on cysteine-104. *J Gen Virol*, **77**, 1821-1824.

Keyvani-Amineh, H., Diouri, M., Guillemette, J.G. and Weber, J.M. (1995a) Electrophoretic and spectral characterization of wild type and mutant adenovirus protease. *J Biol Chem*, **270**, 23250-23253.

Keyvani-Amineh, H., Diouri, M., Tihanyi, K. and Weber, J.M. (1996) Adenovirus type 2 endoprotease: isoforms and redox effects. *J Gen Virol*, **77**, 2201-2207.

Keyvani-Amineh, H., Labrecque, P., Cai, F., Carstens, E.B. and Weber, J.M. (1995b) Adenovirus protease expressed in insect cells cleaves adenovirus proteins, ovalbumin and baculovirus protease in the absence of activating peptide. *Virus Res*, **37**, 87-97.

Lally, C., Dorper, T., Groger, W., Antoine, G. and Winnacker, E.L. (1984) A size analysis of the adenovirus replicon. *EMBO J*, **3**, 333-337.

Lanford, R.E., Feldherr, C.M., White, R.G., Dunham, R.G. and Kanda, P. (1990) Comparison of diverse transport signals in synthetic peptide-induced nuclear transport. *Exp Cell Res*, **186**, 32-38.

Lanford, R.E., Kanda, P. and Kennedy, R.C. (1986) Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell*, **46**, 575-582.

Leibowitz, J. and Horwitz, M.S. (1975) Synthesis and assembly of adenovirus polypeptides. III. Reversible inhibition of hexon assembly in adenovirus type 5 temperature-sensitive mutants. *Virology*, **66**, 10-24.

Leith, I.R., Hay, R.T. and Russell, W.C. (1989) Adenovirus subviral particles and cores can support limited DNA replication. *J Gen Virol*, **70**, 3235-3248.

Leong, K., Lee, W. and Berk, A.J. (1990) High-level transcription from the adenovirus major late promoter requires downstream binding sites for late-phase-specific factors. *J Virol*, **64**, 51-60.

Liu, G.Q., Babiss, L.E., Volkert, F.C., Young, C.S. and Ginsberg, H.S. (1985) A thermolabile mutant of adenovirus 5 resulting from a substitution mutation in the protein VIII gene. *J Virol*, **53**, 920-925.

Lonberg-Holm, K. and Philipson, L. (1969) Early events of virus-cell interaction in an adenovirus system. *J Virol*, **4**, 323-338.

Louis, N., Fender, P., Barge, A., Kitts, P. and Chroboczek, J. (1994) Cell-binding domain of adenovirus serotype 2 fiber. *J Virol*, **68**, 4104-4106.

Mangel, W.F., McGrath, W.J., Toledo, D.L. and Anderson, C.W. (1993) Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. *Nature*, **361**, 274-275.

Mangel, W.F., Toledo, D.L., Brown, M.T., Martin, J.H. and McGrath, W.J. (1996) Characterization of three components of human adenovirus proteinase activity in vitro. *J Biol Chem*, **271**, 536-543.

Mansour, S.L., Grodzicker, T. and Tjian, R. (1986) Downstream sequences affect transcription initiation from the adenovirus major late promoter. *Mol Cell Biol*, **6**, 2684-2694.

Maran, A. and Mathews, M.B. (1988) Characterization of the double-stranded RNA implicated in the inhibition of protein synthesis in cells infected with a mutant adenovirus defective for VA RNA. *Virology*, **164**, 106-113.

Mathews, M.B. (1980) Binding of adenovirus VA RNA to mRNA: a possible role in splicing? *Nature*, **285**, 575-577.

Matthews, D.A. and Russell, W.C. (1994) Adenovirus protein-protein interactions: hexon and protein VI. *J Gen Virol*, **75**, 3365-3374.

Matthews, D.A. and Russell, W.C. (1995) Adenovirus protein-protein interactions: molecular parameters governing the binding of protein VI to hexon and the activation of the adenovirus 23K protease. *J Gen Virol*, **76**, 1959-1969.

McGrath, W.J., Ding, J., Sweet, R.M. and Mangel, W.F. (1996) Preparation and crystallization of a complex between human adenovirus serotype 2 proteinase and its 11-amino-acid cofactor pVIc. *J Struct Biol*, **117**, 77-79.

Medghalchi, S., Padmanabhan, R. and Ketner, G. (1997) Early region 4 modulates adenovirus DNA replication by two genetically separable mechanisms. *Virology*, **236**, 8-17.

Melchior, F. and Gerace, L. (1995) Mechanisms of nuclear protein import. *Curr Opin Cell Biol*, **7**, 310-318.

Moran, E. and Mathews, M.B. (1987) Multiple functional domains in the adenovirus E1a gene. *Cell*, **48**, 177-178.

Morgan, R.S. and McAdon, J.M. (1980) Predictor for sulfur-aromatic interactions in globular proteins. *Int J Pept Protein Res*, **15**, 177-180.

Morgan, R.S., Tatsch, C.E., Gushard, R.H., McAdon, J. and Warme, P.K. (1978) Chains of alternating sulfur and pi-bonded atoms in eight small proteins. *Int J Pept Protein Res*, **11**, 209-217.

Mul, Y.M. and Van der Vliet, P.C. (1992) Nuclear factor I enhances adenovirus DNA replication by increasing the stability of a preinitiation complex. *EMBO J*, **11**, 751-760.

Mul, Y.M., Verrijzer, C.P. and van der Vliet, P.C. (1990) Transcription factors NFI and NFIII/oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. *J Virol*, **64**, 5510-5518.

Nakielny, S. and Dreyfuss, G. (1999) Transport of proteins and RNAs in and out of the nucleus. *Cell*, **99**, 677-690.

Nevins, J.R. (1981) Mechanism of activation of early viral transcription by the adenovirus E1a gene product. *Cell*, **26**, 213-220.

Nevins, J.R. and Darnell, J.E. (1978) Groups of adenovirus type 2 mRNA's derived from a large primary transcript: probable nuclear origin and possible common 3' ends. *J Virol*, **25**, 811-823.

Nevins, J.R. and Winkler, J.J. (1980) Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc Natl Acad Sci U S A*, **77**, 1893-1897.

Norrby, E. (1969) The structural and functional diversity of Adenovirus capsid components. *J Gen Virol*, **5**, 221-236.

O'Malley, R.P., Mariano, T.M., Siekierka, J. and Mathews, M.B. (1986) A mechanism for the control of protein synthesis by adenovirus VA RNAI. *Cell*, **44**, 391-400.

Ornelles, D.A. and Shenk, T. (1991) Localization of the adenovirus early region 1B 55-kilodalton protein during lytic infection: association with nuclear

viral inclusions requires the early region 4 34-kilodalton protein. *J Virol*, **65**, 424-429.

Oualikene, W., Lamoureux, L., Weber, J.M. and Massie, B. (2000) Protease-deleted adenovirus vectors and complementing cell lines: potential applications of single-round replication mutants for vaccination and gene therapy. *Hum Gene Ther*, **11**, 1341-1353.

Peck, D. and Isacke, C.M. (1998) Hyaluronan-dependent cell migration can be blocked by a CD44 cytoplasmic domain peptide containing a phosphoserine at position 325. *J Cell Sci*, **111**, 1595-1601.

Peters R. (1986) Fluorescence microphotolysis to measure nucleocytoplasmic transport and intracellular mobility.*Biochim Biophys Acta*. **864**(3-4):305-359.

Philipson, L. (1984a) The interplay between host and viral genes in adenovirus gene expression. *Klin Wochenschr*, **62**, 433-440.

Philipson, L. (1984b) Structure and assembly of adenoviruses. *Curr Top Microbiol Immunol*, **109**, 1-52.

Prage, L., Pettersson, U., Hoglund, S., Lonberg-Holm, K. and Philipson, L. (1970) Structural proteins of adenoviruses. IV. Sequential degradation of the adenovirus type 2 virion. *Virology*, **42**, 341-358.

Rancourt, C., Keyvani-Amineh, H., Diouri, M. and Weber, J.M. (1996) Mutagenesis of conserved residues of the adenovirus protease. *Virology*, **224**, 561-563.

Rancourt, C., Keyvani-Amineh, H., Sircar, S., Labrecque, P. and Weber, J.M. (1995) Proline 137 is critical for adenovirus protease encapsidation and activation but not enzyme activity. *Virology*, **209**, 167-173.

Rancourt, C., Tihanyi, K., Bourbonniere, M. and Weber, J.M. (1994) Identification of active-site residues of the adenovirus endopeptidase. *Proc Natl Acad Sci U S A*, **91**, 844-847.

Raychaudhuri, P., Bagchi, S., Neill, S.D. and Nevins, J.R. (1990) Activation of the E2F transcription factor in adenovirus-infected cells involves E1a - dependent stimulation of DNA-binding activity and induction of cooperative binding mediated by an E4 gene product. *J Virol*, **64**, 2702-2710.

Rehberger, P.A., Richter, K.H., Schwartz, D., Goldfinger, N., Oskato, R., Almog, N., Marks, F. and Rotter, V. (1997) Differential expression of the regularly spliced wild-type p53 and its COOH-terminal alternatively spliced form during epidermal differentiation. *Cell Growth Differ*, **8**, 851-860.

Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1988) Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell*, **52**, 655-664.

Ridgway, P.J., Hall, A.R., Myers, C.J. and Braithwaite, A.W. (1997) p53/E1b 58kDa complex regulates adenovirus replication. *Virology*, **237**, 404-413.

Roberts, B.L., Richardson, W.D. and Smith, A.E. (1987) The effect of protein context on nuclear location signal function. *Cell*, **50**, 465-475.

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

Roberts, R.J., O'Neill, K.E. and Yen, C.T. (1984) DNA sequences from the adenovirus 2 genome. *J Biol Chem*, **259**, 13968-13975.

Rosenfeld, P.J., O'Neill, E.A., Wides, R.J. and Kelly, T.J. (1987) Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Mol Cell Biol*, **7**, 875-886.

Russell, W.C. and Kemp, G.D. (1995) Role of adenovirus structural components in the regulation of adenovirus infection. *Curr Top Microbiol Immunol*, **199**, 81-98.

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

Russell, W.C., Webster, A., Leith, I.R. and Kemp, G.D. (1989) Phosphorylation of adenovirus DNA-binding protein. *J Gen Virol*, **70**, 3249-3259.

Ruzindana-Umunyana, A., Sircar, S. and Weber, J.M. (2000b) The effect of mutant peptide cofactors on adenovirus protease activity and virus infection. *Virology*, **270**, 173-179.

Sambrook, J., Frisch. and Maniatis, T. (1989) *Molceular Cloning: a laboratory manual*. Cold Spring Harbour Press, Cold Spring Harbour, NY.

Sarnow, P., Hearing, P., Anderson, C.W., Halbert, D.N., Shenk, T. and Levine, A.J. (1984) Adenovirus early region 1B 58,000-dalton tumor antigen is physically associated with an early region 4 25,000-dalton protein in productively infected cells. *J Virol*, **49**, 692-700.

Sarnow, P., Ho, Y.S., Williams, J. and Levine, A.J. (1982a) Adenovirus E1b - 58kD tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kD cellular protein in transformed cells. *Cell*, **28**, 387-394.

Sarnow, P., Sullivan, C.A. and Levine, A.J. (1982b) A monoclonal antibody detecting the adenovirus type 5-E1b -58Kd tumor antigen: characterization of the E1b -58Kd tumor antigen in adenovirus-infected and -transformed cells. *Virology*, **120**, 510-517.

Schaack, J., Ho, W.Y., Freimuth, P. and Shenk, T. (1990) Adenovirus terminal protein mediates both nuclear matrix association and efficient transcription of adenovirus DNA. *Genes Dev*, **4**, 1197-1208.

Seth, P., Fitzgerald, D., Ginsberg, H., Willingham, M. and Pastan, I. (1984a) Evidence that the penton base of adenovirus is involved in potentiation of

toxicity of Pseudomonas exotoxin conjugated to epidermal growth factor. *Mol Cell Biol*, **4**, 1528-1533.

Seth, P., Fitzgerald, D.J., Willingham, M.C. and Pastan, I. (1984b) Role of a low-pH environment in adenovirus enhancement of the toxicity of a Pseudomonas exotoxin-epidermal growth factor conjugate. *J Virol*, **51**, 650-655.

Shen, Y and Shenk, T (1994) Relief of p53 mediated transcriptional repression by the AdenovirusE1B-19kD protein or the celluar Bcl-2 protein. *Natl Acad Sci USA***91**, 8940-8944.

Shenk, T (1996). Adenoviridae: The Viruses and Their Replication. In "Fundamental Virology" (B. N. Fields,D. M. Knipe and P. M. Howley, eds) pp 979-1016. *Lippincroft and Raven.*, Philadelphia

Shenk, T. and Flint, J. (1991) Transcriptional and transforming activities of the adenovirus E1a proteins. *Adv Cancer Res*, **57**, 47-85.

Sircar, S., Keyvani-Amineh, H. and Weber, J.M. (1996) Inhibition of adenovirus infection with protease inhibitors. *Antiviral Res*, **30**, 147-153.

Sircar, S., Ruzindana-Umunyana, A., Neugebauer, W. and Weber, J.M. (1998) Adenovirus endopeptidase and papain are inhibited by the same agents. *Antiviral Res*, **40**, 45-51.

Smart, J.E. and Stillman, B.W. (1982) Adenovirus terminal protein precursor. Partial amino acid sequence and the site of covalent linkage to virus DNA. *J Biol Chem*, **257**, 13499-13506.

Stewart, P.L. and Burnett, R.M. (1995) Adenovirus structure by X-ray crystallography and electron microscopy. *Curr Top Microbiol Immunol*, **199**, 25-38.

Stewart, P.L., Burnett, R.M., Cyrklaff, M. and Fuller, S.D. (1991) Image reconstruction reveals the complex molecular organization of adenovirus. *Cell*, **67**, 145-154.

Stewart, P.L., Fuller, S.D. and Burnett, R.M. (1993) Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy. *EMBO J*, **12**, 2589-2599.

Stillman, B. (1989) Initiation of eukaryotic DNA replication in vitro. *Annu Rev Cell Biol*, **5**, 197-245.

Sundquist, B., Everitt, E., Philipson, L. and Hoglund, S. (1973) Assembly of adenoviruses. *J Virol*, **11**, 449-459.

Sung, M.T., Cao, T.M., Lischwe, M.A. and Coleman, R.T. (1983) Molecular processing of adenovirus proteins. *J Biol Chem*, **258**, 8266-8272.

Svensson, U. (1985) Role of vesicles during adenovirus 2 internalization into HeLa cells. *J Virol*, **55**, 442-449.

Tamanoi, F. and Stillman, B.W. (1983) Initiation of adenovirus DNA replication in vitro requires a specific DNA sequence. *Proc Natl Acad Sci U S A*, **80**, 6446-6450.

Temperley, S.M. and Hay, R.T. (1992) Recognition of the adenovirus type 2 origin of DNA replication by the virally encoded DNA polymerase and preterminal proteins. *EMBO J*, **11**, 761-768.

Thomas, G.P. and Mathews, M.B. (1980) DNA replication and the early to late transition in adenovirus infection. *Cell*, **22**, 523-533.

Tibbetts, C. (1977) Viral DNA sequences from incomplete particles of human adenovirus type 7. *Cell*, **12**, 243-249.

Tihanyi, K., Bourbonniere, M., Houde, A., Rancourt, C. and Weber, J.M. (1993) Isolation and properties of adenovirus type 2 proteinase. *J Biol Chem*, **268**, 1780-1785.

Tomko, R.P., Johansson, C.B., Totrov, M., Abagyan, R., Frisen, J. and Philipson, L. (2000) Expression of the adenovirus receptor and its interaction with the fiber knob. *Exp Cell Res*, **255**, 47-55.

Toth, M., Doerfler, W. and Shenk, T. (1992) Adenovirus DNA replication facilitates binding of the MLTF/USF transcription factor to the viral major late promoter within infected cells. *Nucleic Acids Res*, **20**, 5143-5148.

Tremblay, M.L., Dery, C.V., Talbot, B.G. and Weber, J. (1983) In vitro cleavage specificity of the adenovirus type 2 proteinase. *Biochim Biophys Acta*, **743**, 239-245.

Valentine, R.C. and Pereira, H.G. (1965) Antigens and structure of the adenovirus. *J Mol Biol*, **13**, 13-20.

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

van Oostrum, J., Smith, P.R., Mohraz, M. and Burnett, R.M. (1986) Interpretation of electron micrographs of adenovirus hexon arrays using a crystallographic molecular model. *J Ultrastruct Mol Struct Res*, **96**, 77-90.

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

Varga, M.J., Weibull, C. and Everitt, E. (1991) Infectious entry pathway of adenovirus type 2. *J Virol*, **65**, 6061-6070.

Vaghan, O.A., (1997) The Adenovirus Type 2 Protease: The generation and characterisation of monoclonal antibodies and their use in determining the subcellular distribution of the protease within lytically infected cells. **PhD Thesis**, Universtity of St. Andrews.

Velcich, A. and Ziff, E. (1985) Adenovirus E1a proteins repress transcription from the SV40 early promoter. *Cell*, **40**, 705-716.

Vrati, S., Brookes, D.E., Strike, P., Khatri, A., Boyle, D.B. and Both, G.W. (1996) Unique genome arrangement of an ovine adenovirus: identification of new proteins and proteinase cleavage sites. *Virology*, **220**, 186-199.

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

Weber, J.M. and Anderson, C.W. (1988) Identification of the gene coding for the precursor of adenovirus core protein X. *J Virol*, **62**, 1741-1745.

Weber, J.M., Dery, C.V., Mirza, M.A. and Horvath, J. (1985) Adenovirus DNA synthesis is coupled to virus assembly. *Virology*, **140**, 351-359.

Weber, J.M., Talbot, B.G. and Delorme, L. (1989) The orientation of the adenovirus fiber and its anchor domain identified through molecular mimicry. *Virology*, **168**, 180-182.

Webster, A., Hay, R.T. and Kemp, G. (1993) The adenovirus protease is activated by a virus-coded disulphide-linked peptide. *Cell*, **72**, 97-104.

Webster, A. and Kemp, G. (1993) The active adenovirus protease is the intact L3 23K protein. *J Gen Virol*, **74**, 1415-1420.

Webster, A., Leith, I.R. and Hay, R.T. (1994) Activation of adenovirus-coded protease and processing of preterminal protein. *J Virol*, **68**, 7292-7300.

Webster, A., Leith, I.R. and Hay, R.T. (1997a) Domain organization of the adenovirus preterminal protein. *J Virol*, **71**, 539-547.

Webster, A., Leith, I.R., Nicholson, J., Hounsell, J. and Hay, R.T. (1997b) Role of preterminal protein processing in adenovirus replication. *J Virol*, **71**, 6381-6389.

Webster, A., Russell, S., Talbot, P., Russell, W.C. and Kemp, G.D. (1989a) Characterization of the adenovirus proteinase: substrate specificity. *J Gen Virol*, **70**, 3225-3234.

Webster, A., Russell, W.C. and Kemp, G.D. (1989b) Characterization of the adenovirus proteinase: development and use of a specific peptide assay. *J Gen Virol*, **70**, 3215-3223.

White, E. and Cipriani, R. (1989) Specific disruption of intermediate filaments and the nuclear lamina by the 19-kD product of the adenovirus E1b oncogene. *Proc Natl Acad Sci U S A*, **86**, 9886-9890.

White, E. and Cipriani, R. (1990) Role of adenovirus E1b proteins in transformation: altered organization of intermediate filaments in transformed cells that express the 19-kilodalton protein. *Mol Cell Biol*, **10**, 120-130.

Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A. and Harlow, E. (1988) Association between an oncogene and

an anti-oncogene: the adenovirus E1a proteins bind to the retinoblastoma gene product. *Nature*, **334**, 124-129.

Wickham, T.J., Mathias, P., Cheresh, D.A. and Nemerow, G.R. (1993) Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell*, **73**, 309-319.

Wong, H.K. and Ziff, E.B. (1994) Complementary functions of E1a conserved region 1 cooperate with conserved region 3 to activate adenovirus serotype 5 early promoters. *J Virol*, **68**, 4910-4920.

Yeh-Kai, L., Akusjarvi, G., Alestrom, P., Pettersson, U., Tremblay, M. and Weber, J. (1983) Genetic identification of an endoproteinase encoded by the adenovirus genome. *J Mol Biol*, **167**, 217-222.

Yoshinaga, S., Dean, N., Han, M. and Berk, A.J. (1986) Adenovirus stimulation of transcription by RNA polymerase III: evidence for an E1a -dependent increase in transcription factor IIIC concentration. *EMBO J*, **5**, 343-354.

Zantema, A., Fransen, J.A., Davis-Olivier, A., Ramaekers, F.C., Vooijs, G.P., DeLeys, B. and Van der Eb, A.J. (1985) Localization of the E1b proteins of adenovirus 5 in transformed cells, as revealed by interaction with monoclonal antibodies. *Virology*, **142**, 44-58.

Zhang, Y. and Schneider, R.J. (1994) Adenovirus inhibition of cell translation facilitates release of virus particles and enhances degradation of the cytokeratin network. *J Virol*, **68**, 2544-2555.

Zhu, X.G., Hanover, J.A., Hager, G.L. and Cheng, S.Y. (1998) Hormone-induced translocation of thyroid hormone receptors in living cells visualized using a receptor green fluorescent protein chimera. *J Biol Chem*, **273**, 27058-27063.