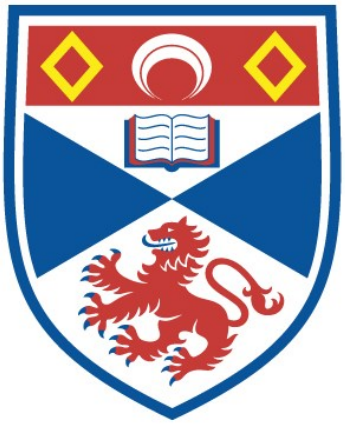
THE ACTIVITY OF HUMAN RHINOVIRUS 14 3C PROTEASE IN ARTIFICIAL POLYPROTEINS

Emma Jane Byrne

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



1999

The Activity of Human Rhinovirus 14 3C protease in Artificial Polyproteins.

Emma Jane Byrne

University of St Andrews

A thesis submitted in partial fulfilment
of the requirements for the degree
of Doctor of Philosophy

March 1999

Declarations

I, Emma Byrne, hereby certify that this thesis, which is approximately 50 000 words in length, has been written by me, and that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 29/3/99 signature of candidate

I was admitted as a research student in October 1994 and as a candidate for the degree of Doctor of Philosophy in October 1994; the higher study for which this is a record was carried out in the University of St Andrews between 1994 and 1999.

date 29/3/99 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 26/3/99 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 a copy of the work may be made and supplied to any *bona fide* library or research worker.

date 26/3/99 signature of supervisor

Acknowledgements

I would like to thank Dr Martin Ryan for his supervision and support over the course of this work. I would also like to thank him for his help in writing this thesis.

I would like to acknowlege the following people; Alex Houston for DNA sequencing, Ian Armitt for synthesis of oligonucleotides, Dr Glyn Stanway for the gift of his HRV14 cDNA clone and Vanessa Cowton for the construction of plasmids pGEM7zfGFP-2A and 3C' pGEM-T.

Special thanks are due to all my friends for their support and encouragement during this PhD especially Claire Dunn, Vanessa Cowton, Michelle Donnelly, Gayle Middleton and Ruth Edgar.
Thanks to everyone in the lab for their help and advice especially Edwin ten Dam for his help in the writing of this thesis.

Thanks also to my family for their support, eternal optimism and faith in my abilities.

Abstract

HRV14 3C acts as a protease and has a role in RNA replication *in vivo*, interacting with a cloverleaf structure in picornaviral genomic RNA. Picornaviral 3C proteases are able to cleave both N- and C-terminally producing $3CD^{pro}$, or, $3C^{pro}$ and $3D^{pol}$, respectively. In order to investigate the mechanisms whereby these alternative processing pathways are adopted an artificial polyprotein system was constructed, composed of viral sequences from the P3 region of the viral polyprotein flanked by reporter genes.

Two antibiotic resistance genes (Kan^R, Tet^R) were cloned to act as reporter genes flanking the viral region of interest. Analysis of the cleavage products in a coupled TnT system showed whether N- or C-terminal cleavage had occurred. $3C^{pro}$ cleaved preferentially at its N-terminus in [$Kan^R3C^{pro}Tet^R$] with a lesser degree of cleavage at its C-terminus. When 3ABC was used as the viral component of the system cleavage at the N-terminus of $3C^{pro}$ was also observed.

The use of $3CD^{pro}$ as the viral component also had a regulatory effect on the site (N- or C-terminal) of cleavage by $3C^{pro}$. With $3CD^{pro}$ as the viral component of the artificial reporter polyprotein cleavage occurred at both the N-and C-termini of $3C^{pro}$ as well as at the C-terminus of $3D^{pol}$. This surprising result has led to comparisons with the proteolytic action of viral proteases in the caliciviruses and some plant viruses and the proposal of possible evolutionary links between these viruses.

The use of the antibiotic resistance genes as reporters allowed investigation into the use of antibiotic resistance phenotypes in *E.coli* for monitoring cleavage of the artificial polyprotein. Preliminary results indicated that the system may be useful as a genetic screen to quantify large numbers of mutants.

Contents

LIST OF FIGURES

3. Cloning And Molecular Characterisation Of The Components Of A 'reporter' Polyprotein System.

4. Proteolytic Activity Of HRV14 Proteases In An Artificial Reporter Polyprotein System

List of tables

Abbreviations

Å	angstroms
ACTH	Adrenocorticotrophic hormone
AEV	avian enterovirus
AiV	Aichi virus
ATP	Adenosine triphosphate
bp	base pairs
BEV	bovine enterovirus
BRV	bovine rhinovirus
CBC	cap binding protein complex
CLIP	Corticotrophin-like intermediate lobe peptide
CPMV	Cowpea mosaic virus
CV	coxsackievirus
DFP	Diisopropyl fluorophosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECHO virus	Enteric cytopathic human orphan virus
EDTA	Ethylene diamine tetra acetic acid
eIF	Eukaryotic initiation factor
EMCV	Encephalomyocarditis virus
Enk	Enkephalin
ERAV	equine rhinitis virus A
ERBV	equine rhinitis virus B
ERV	Equine rhinovirus
EV	enterovirus
E. coli	*Escherichia coli*
FCV	Feline calicivirus
FMDV	Foot-and-mouth disease virus
gag	group specific antigen
HCV	Hepatitis C virus
HEV	human enterovirus
HHAV	human hepatitis A virus
HIV-1	Human immunodeficiency virus 1
HPeV	human parechovirus
HRV	Human rhinovirus
ICAM-1	Intercellular adhesion molecule 1
IRES	Internal ribosome entry site
kDa	kilodaltons

LB	Luria Bertani
LDL	low denisty lipoprotein
LPH	Lipotrophic hormone
MAP-4	Microtubule associated protein 4
MAV	Myeloblastosis virus
MSH	Melanocyte stimulating hormone
NCR	Non-coding region
NIm site	Neutralising Immunogenic site
OEV	ovine enterovirus
ORF	Open reading frame
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEV	porcine enterovirus
POMC	Prepro-opiomelanocortin
PR	retroviral protease
PTV	porcine teschovirus
PV	poliovirus
PVR	poliovirus receptor
RHDV	Rabbit haemorrhagic disease virus
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RRL	Rabbit reticulocyte lysate
RSV	Rous sarcoma virus
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEV	simian enterovirus
SHAV	simian hepatitis A virus
Taq polymerase	*Thermophilus aquaticus* polymerase
TBP	TATA-binding protein
TEV	Tobacco etch virus
TFIIIC	Transcription factor IIIC
TMEV	Theiler's murine encephalitis virus
TnT	Transcription/translation
VHEV	Vilyuisk human encephalitis virus
VPg	Viral genome protein
WGE	Wheatgerm extract

Abbreviations for amino acids

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

1. Introduction

1.1 Picornaviruses

Picornaviruses are small, positive strand RNA viruses which are of great medical and economic importance. Poliovirus is one of the most well-known members of the family and has been documented to afflict man for many hundreds of years. Other members of the family infect animals, such as foot-and-mouth disease virus (FMDV) and the bovine enteroviruses. Phylogenetically the picornaviruses can be divided into five groups; Group I containing the aphthoviruses, Group II containing the cardioviruses, Group III containing the hepatitis A viruses, Group IV containing the rhino- and enteroviruses and Group V containing the parechoviruses. The parechoviruses were formerly ECHO22 (now parechovirus 1; Stanway *et al.*, 1994) and ECHO 23 (now parechovirus 2). Three new genera have also been proposed which contain previously unclassified viruses (see Table 1.1).

The genome organisation of the prototype strain for each class is similar (Figure 1.1) and the capsid structure is also common to all classes. The capsid is a 60 subunit protein shell 20-30 nm in diameter with 5-, 3- and 2-fold axes of symmetry. The particle is composed of 4 non-identical polypeptide chains; VP1 (1D), VP2 (1B), VP3 (1C) and VP4 (1A). The 'VPX' nomenclature is according to the descending molecular weight of the proteins on polyacrylamide gels and the 1A-1D nomenclature signifies the genome order. Each subunit of the capsid (protomer) is made up of one copy of each of the capsid proteins, assembled into pentamers. Twelve pentamers then make up the dodecahedral capsid.

No lipid envelope surrounds the capsid and its infectivity is therefore relatively unaffected by exposure to organic solvents such as ether. A single copy of positive strand RNA is surrounded by the capsid and is released into the cytoplasm of target cells to initiate replication.

The vRNA has a single long open reading frame (ORF) which encodes a single polypeptide known as the polyprotein. This is processed by various proteolytic activities

Table 1.1 Classification of the picornaviruses

Genus	No. of serotypes	Examples
Aphthovirus	7	Foot and mouth disease virus (FMDV)1-7 (serotypes A, C, O, SAT-1, 2, 3, Asia-1)
	1	equine rhinitis A virus (ERAV) (formerly equine rhinovirus 1)
Cardiovirus	1	encephalomyocarditis virus (EMCV)
	1	Theiler's Murine encephalitis virus (TMEV)
	1	Vilyuisk human encephalitis virus (VHEV)
Enterovirus	3	Human polioviruses (PV) 1, 2, 3
	23	Human coxsackieviruses (CV) A1-22, 24
	6	Human coxsackieviruses B1-6
	28	Human echoviruses (E)1-7, 9, 11-21, 24-27, 29-34
	4	Human enteroviruses (EV) 68-71
	18	Simian enteroviruses (SEV) 1-18
	2	Bovine enteroviruses (BEV) 1, 2
	7	Porcine enteroviruses (PEV) 1-8
	1	Ovine enterovirus (OEV)
Hepatovirus	1	Human hepatitis A virus (HHAV)
	1	Simian hepatitis virus (SHAV)
	1	Avian encephalitis virus(AEV) (proposed)
Rhinovirus	102	Human Rhinovirus 1A-100, 1B, Hanks
	3	Bovine Rhinovirus 1, 2, 3.
Parechovirus	2	Human Parechovirus 1 (HPeV-1) (formerly Echovirus-22) Human Parechovirus 2 (HPeV-2) (formerly Echovirus-23)
Unassigned		Equine rhinovirus 3
		cricket paralysis virus
		Drosophila C virus
Proposed Genus		
Erbovirus	1	Equine rhinitis B virus (ERBV) (formerly equine rhinovirus 2)
Aichi-like virus	1	Aichi virus (AiV)
Teschovirus	1	porcine teschovirus 1 (PTV-1) (formerly porcine enterovirus 1)

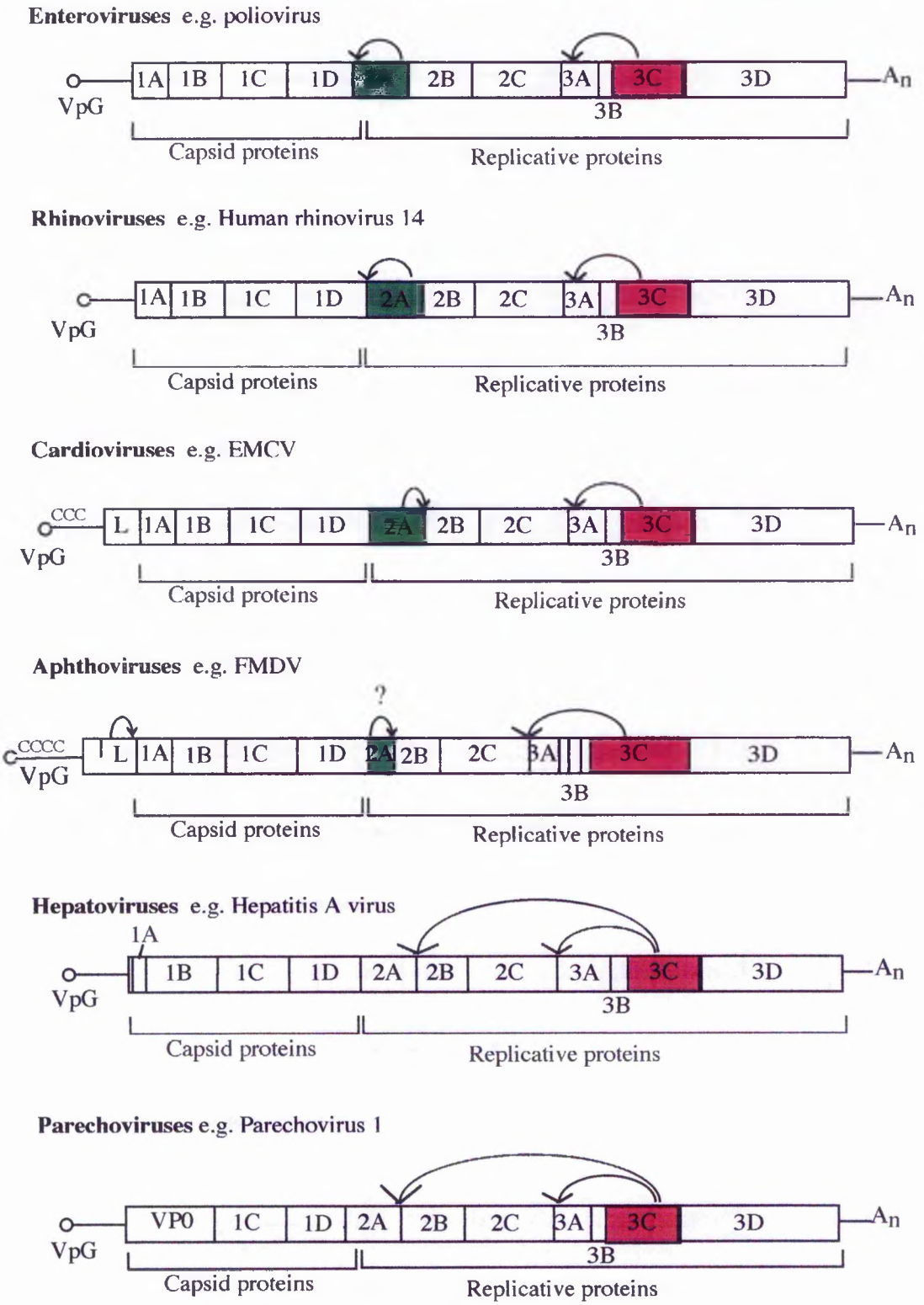
to yield 'mature proteins'. The exceptions to this scheme are ; encephalomyocarditis virus (EMCV), foot-and-mouth disease virus (FMDV) and human hepatitis A virus (HHAV) - all having an additional translation start site separate from the polyprotein initiation codon. This second initiation codon may initiate translation *in vitro* (EMCV and hepatitis A) and *in vitro* and *in vivo* (FMDV). The two polyproteins produced in these cases are of slightly different lengths but in the same ORF. The DA strain of Theiler's murine encephalitis virus (TMEV) produces a 17 kDa protein (L*) out of frame with the polyprotein (Kong and Roos, 1991). This is initiated from an AUG 13 nucleotides downstream from the polyprotein AUG. This protein plays a role in TO subgroup induced demyelinating disease (Chen *et al.*, 1995) .

The length of this genomic RNA varies. 7102 bases in rhinovirus type 2 (HRV 2) is one of the shortest and 8282 bases in FMDV 01K is one of the longest. This is excluding the polyA tail which varies in length from 50 to 150 bases. The polyprotein accounts for 85-90% of the theoretical coding capacity of the genome. The other 10-15% of the genome is distributed between the 5' and 3' non-coding regions (NCR) of the genome.

Initiation of translation in picornaviruses is mediated by the 5' NCR, which contains an internal ribosome entry site (IRES; Alsaadi *et al.*, 1989). A viral coded genome linked protein (VPg) is linked to the 5' terminal pUpUp of the RNA through the phenolic O^4 hydroxyl group of a tyrosine residue (Wimmer, 1982). The function of the VPg is unknown but it is thought to be involved in RNA synthesis (Rothberg *et al.*, 1978). The 3' end of the genome has a polyA tail.

All picornaviruses have a heterogeneous 3' terminal polyA tail whose length is on average from 40 to 60 nucleotides. It forms part of the 3' NCR of the positive strand viral (genomic) RNA. A role has been proposed for the 3' NCR and the polyA tail in replication.

Figure 1.1 Genome organisation of picornaviruses. The proteolytic enzymes encoded by the polyproteins are shown in colour. Green indicates a 2Aprotease and pink a 3C protease. Arrows indicate the primary cleavages of the viral polyprotein.



1.1.2 Replication cycle of picornaviruses

Replication of picornaviruses occurs entirely in the cytoplasm of the infected cell (Figure 1.2). This process has been reviewed by Rueckert (1996) and is summarised here. The initial event in the replication process is the attachment of the virus particle to the host cell. This occurs *via* receptors on the cell surface. These molecules have been identified for many picornaviruses. For instance, the major group of rhinoviruses utilises intercellular adhesion molecule-1 (ICAM-1). The ICAM-1 binds to the viral "canyon" which surrounds the five-fold vertex of the virus capsid. The gene for this molecule maps to human chromosome 19, as does that of the poliovirus receptor (PVR). The role of the viral receptor is two-fold. It positions the virion near to the host cell membrane and triggers a change in the virion. This involves the loss of a protein (VP4) located in the virion and the delivery of the viral RNA across the host cell membrane into the cytosol.

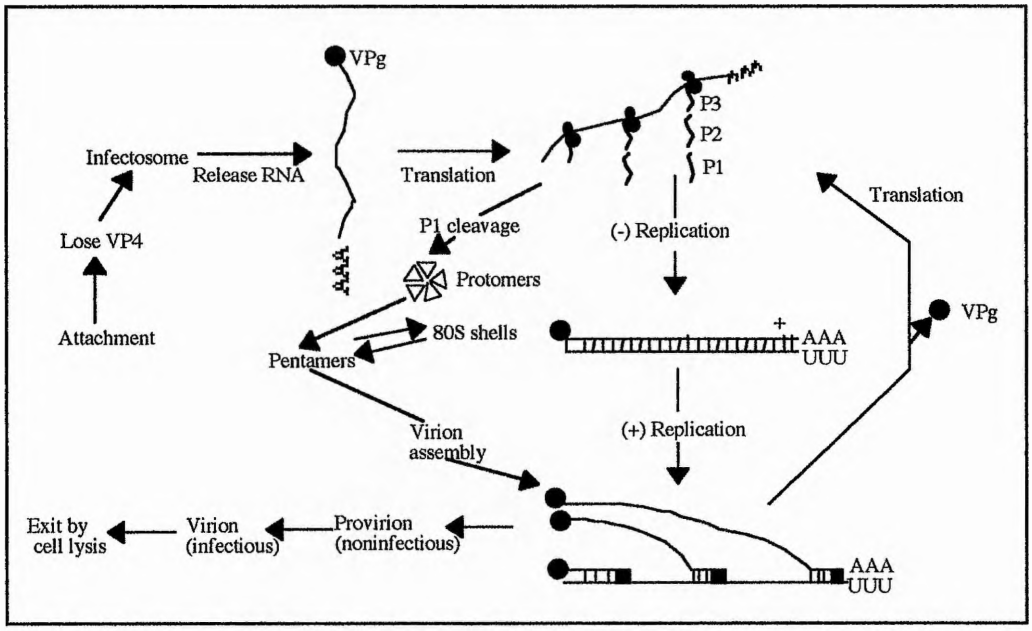


Figure 1.2 Replication of picornaviruses

The process of entry into cells differs among groups of picornaviruses. FMDV and the minor receptor group rhinoviruses are internalised by their receptors into endosomes (Zeichhardt *et al.*, 1985). Virus-induced pores are opened in the vesicular membrane in response to the low pH and it is likely that this allows the RNA to leave the endosome..

The major receptor group rhinoviruses, such as HRV14, enter cells by a different method. Infection with the virus results in endosome rupture which releases subviral particles into the cytoplasm (Schober *et al.*, 1998). The viral receptor, ICAM-1, is internalised either as a result of clustering around the virus capsid or natural turnover. After the virus has been transported into endosomes it becomes modified by the uncoating activity of ICAM-1 and the increase in hydrophobicity causes the endosome to rupture. The subviral particles are transferred to the cytosol ,however, rather than the vRNA which happens in the minor receptor group. This may not be the only infection pathway in these viruses.

The next stage of replication is translation of the input vRNA. This is crucial as RNA synthesising machinery encoded by the virus needs to be produced before replication of the RNA can begin. The virus utilises ribosomes and other protein synthesising components of the host cell to produce proteins. The RNA is translated to produce a polyprotein which undergoes self cleavage whilst protein synthesis is still occurring. Translation can be carried out by many ribosomes at once. Polysomes have been observed in poliovirus with up to 40 ribosomes present at one time on a single vRNA.

The first domain to be released from the polyprotein is P1, the coat precursor protein. The P2 domain is then released and finally P3. These fragments are all released by protease activity encoded by the virus in the polyprotein (see section 1.3.5). P3 can be cleaved either *in cis*, in a monomolecular reaction, or *in trans*.

RNA is transcribed by copying the incoming viral RNA (Figure 1.2) producing a negative strand RNA which, in turn, is used as a template for new positive strands. The synthesis of positive strand RNA occurs on the smooth endoplasmic reticulum and generates multi-stranded replicative intermediates composed of one strand of negative sense RNA and several positive strands. In the early stages of replication the positive strands are recycled as replication centres until the number is so great that they are packaged into virions.

Virion assembly is controlled by the cleavage of P1 which generates the coat proteins VP0, VP1 and VP3 which constitute protomers (Figure 1.3). Cleavage of P1 is

slow in the early stages of replication because of low concentrations of P1 and the virus encoded proteases $3C^{pro}$ and $3CD^{pro}$. Later in the replication cycle rising concentrations trigger assembly into pentamers. These package the positive strand VPg-RNA into provirions.

The provirions are not infectious and the formation of infectious particles requires the cleavage of the capsid protein precursor VP0 to VP2 and VP4. After maturation new virus particles are released by the infection-mediated disintegration of the host cell. The time required for the replication cycle from infection to completion of assembly is from 5 to 10 hours depending on several factors such as pH, the virus and the host cell.

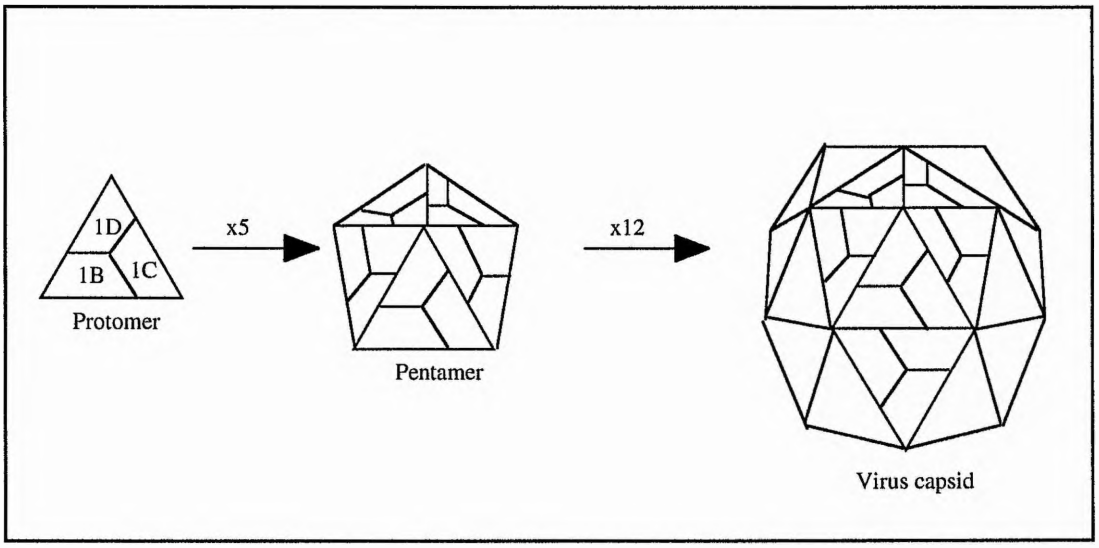


Figure 1.3 Picornavirus capsid assembly.

1.2. Rhinoviruses

1.2.1 Overview

Rhinoviruses are responsible for approximately 50% of cases of the common cold in humans (Sperber and Hayden, 1988). They account for a significant amount of lost man-hours in Western countries and so are of great economic importance. Rhinovirus infection is widespread throughout the world (Brown and Taylor-Robinson, 1966). Many different serotypes (~102) have been identified in humans (Hamparian *et al.*, 1987). These have been classified in two "species", HRV-A and HRV-B. Some serotypes have yet to be assigned to species. Bovine rhinoviruses, of which there are 3 serotypes, are also included in the rhinovirus group. Equine rhinoviruses, however, have been assigned to different genera (see Table 1.1). The rhinovirus group was first described in 1963 by Tyrrell and Channock. The name refers to the primary site of infection in the human airway epithelial cells. They can be distinguished from the enteroviruses by their acid lability (Newman *et al.*, 1973) and their more extreme species specificity. All members of the genus are inactive below pH6 and most rhinoviruses are thermostable, although the extent of this property varies between serotypes.

Although there are many serotypes of human rhinovirus studies show that a limited number are prevalent in communities (Fox *et al.*, 1985;Monto and Cavallaro, 1972). The prevalent serotypes in circulation appear, however, to change from year to year. A small number of the serotypes cause illness in each outbreak. All known serotypes seem to be circulating but some are more often found to cause illness. A change in distribution was seen in a seven year study carried out in Virginia with untypable strains increasing with time (Calhoun *et al.*, 1974).

1.2.2 Symptoms and illness

The primary illness caused by rhinoviruses is the common cold, symptoms of which normally last for two to three days. Complications can result in lower respiratory tract disease and chronic bronchitis. There are no symptoms specific to rhinovirus infection

and it is therefore difficult to diagnose. Transmission of the virus is from person to person *via* virus contaminated respiratory secretions.

Respiratory viruses have been associated with an increase in the severity of asthma. Rhinovirus has been proven to be the virus most often involved (Minor *et al.*, 1974). In this study carried out by Minor and co-workers 14 out of 15 rhinovirus infections resulted in symptomatic respiratory infections and precipitated asthma attacks. The association has been illustrated dramatically in children but asthma patients of all ages can be affected. The mechanisms by which rhinoviruses enhance the severity of asthma have begun to be elucidated. Infection with rhinovirus has been shown to enhance airway responsiveness and late allergic reaction and to promote mast cell release of histamine and eosinophils to the airways. Asthma has also been shown to increase susceptibility to rhinovirus infection (Bianco and Spiteri, 1998).

1.2.3 Classification

There have been many attempts to classify rhinoviruses and some are detailed in the following sections. Today, however, grouping on the basis of sequence similarities between aligned sequences seems to be the most widely used. Conserved regions amongst sequences can predict many features of interest such as binding sites and may be used to predict, *e.g.* protease cleavage sites.

1.2.3.1 Receptor groups

Rhinoviruses have historically been grouped in several ways. There are two groups in the genus according to their cellular receptors. The major group uses ICAM-1 as its receptor (Greve *et al.*, 1989; Staunton *et al.*, 1989; Tomassini *et al.*, 1989) and the other smaller group uses the low density lipoprotein (ldl) receptor (Hofer *et al.*, 1994). ICAM-1 is a cytokine-inducible cell surface receptor which contains five immunoglobulin superfamily (IgSF) domains, a short transmembrane region and a small carboxyl-terminal cytoplasmic domain (Figure 1.4). The binding site for rhinovirus is at the tip of domain 1 and extends about half way down (Bella *et al.*, 1998; Casasnovas *et al.*, 1998; McClelland

McClelland*et al.*, 1991; Register *et al.*, 1991; Staunton *et al.*, 1989). Soluble ICAM-1 inhibits infection by rhinovirus acting as a competitive inhibitor and initiating irreversible disruption of the capsid with release of the viral RNA (Casasnovas and Springer, 1994; Greve *et al.*, 1991; Marlin *et al.*, 1990; Martin *et al.*, 1993). The two N-terminal domains, D1 and D2, bind to HRV3 with the same affinity as the entire extracellular domain (Casasnovas *et al.*, 1998). It interacts with a deep depression (canyon) on the surface of HRV (Figure 1.5).

The position of the receptor binding site in the canyon prevents it from being blocked by antibodies and the virus is therefore more able to avoid antibody neutralisation (Rossmann and Palmenberg, 1988; Rossmann *et al.*, 1985). The major receptor group includes 90 of the 102 known human rhinovirus serotypes and the minor receptor group has 10 members.

One serotype, HRV-87, uses neither the major or minor receptor and may represent a third receptor group (Uncapher *et al.*, 1991). It displays cell binding tropisms similar to those of the minor group viruses but does not use the ldl receptor for binding. The receptor for this virus has yet to be identified but it has been found to require sialic acid for attachment. A similar observation has also been made for encephalomyocarditis virus binding to human erythrocytes (Allaway and Burness, 1986).

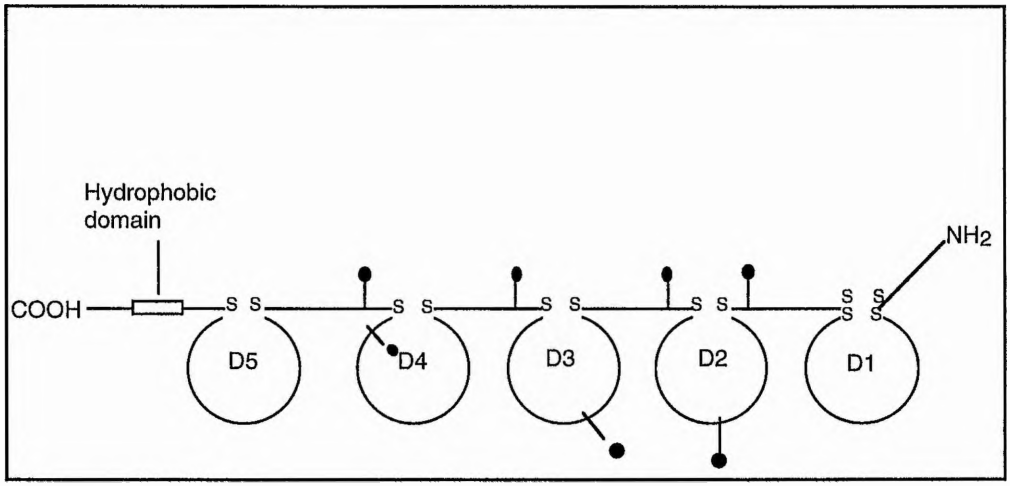
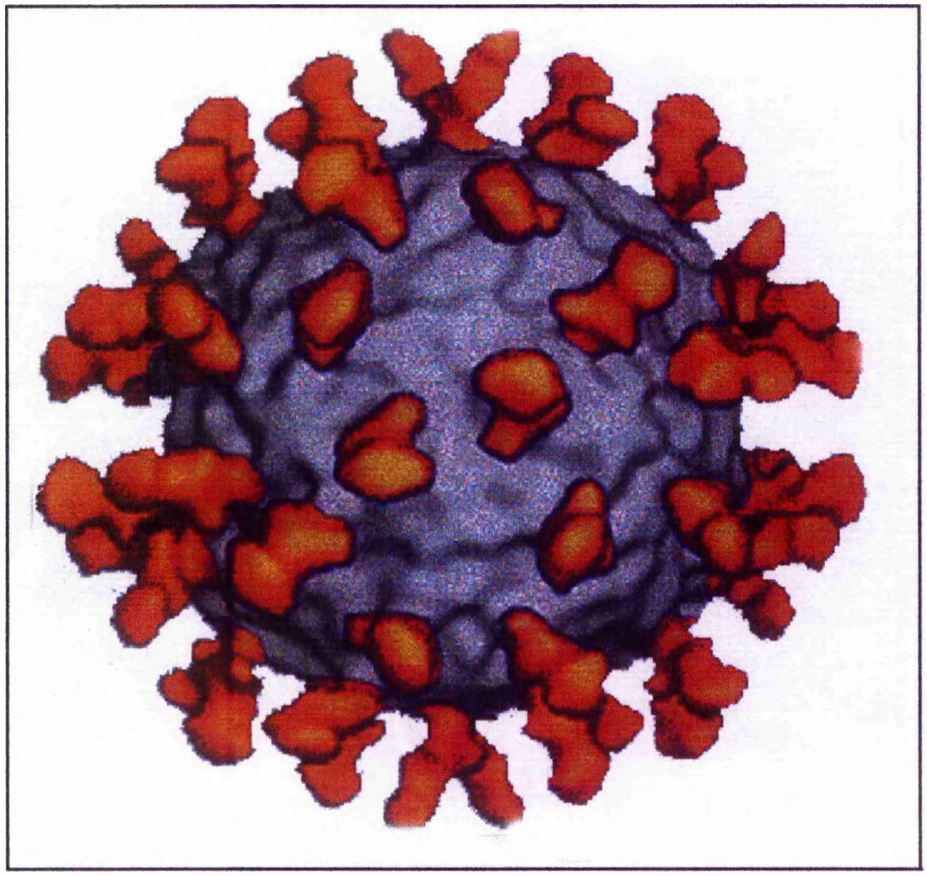


Figure 1.4 Diagram of an ICAM-1 molecule showing sites of glycosylation (lollipop shaped structures).

Figure 1.5 Model of Human Rhinovirus 16 complexed with its cellular receptor ICAM-1. The structure was solved by cryo-electron microscopy and image reconstruction (Baker, 1992). The orange areas show the receptor molecules (ICAM-1) which are bound in the viral canyon.



1.2.3.2 Sensitivity to antiviral agents

Subgroups have also been predicted according to the sensitivity of the virus to antiviral agents, such as WIN compounds (see below). These interact with the capsid and interfere with uncoating or receptor binding (Andries *et al.*, 1990). Two groups have been designated: Group A contains 33% of the known serotypes and Group B 67%. This indicates that the antiviral binding site is dimorphic. Binding of antiviral agents correlates highly with sequence similarities in the antiviral binding site and in VP1. The viruses which are designated as Group B cause twice as many infections as those in Group A. All of the viruses in the minor receptor binding group are found in antiviral binding Group B.

1.2.3.3 Antigenic relationships

Rabbit antisera were raised to 90 known rhinovirus serotypes and neutralisation tests showed that there were significant numbers of cross-reactions between the different rhinovirus serotypes. Sixteen groups have been defined which include 50 serotypes (Cooney *et al.*, 1982). There are a number of antigenically related virus pairs including serotypes 12 and 78 and 36 and 58. HRV67 is related to both serotypes 9 and 52.

1.2.3.4 Sequence groups

Nucleotide and predicted amino acid sequences have also allowed classification. The majority of HRVs appear to be closely related by this method, with HRV14 being more distinct (Stanway, 1990). Genetic groups have also been defined by polymerase chain reaction (PCR) of a 112 amino acid region corresponding to neutralising immunogenic site II (NIm-II), a major immunogenic site in HRVs. Two groups, designated GG1 and GG2, were identified. All of Group A from the antiviral grouping is in GG1 and GG2 contains all of Group B (Horsnell *et al.*, 1995). Some correlation was also seen with groupings according to antiviral drug sensitivity which suggests that there is a fundamental division in HRVs. No correlation was seen with receptor groups as the larger group contains members of both receptor groupings.

1.2.4 Prevention and control

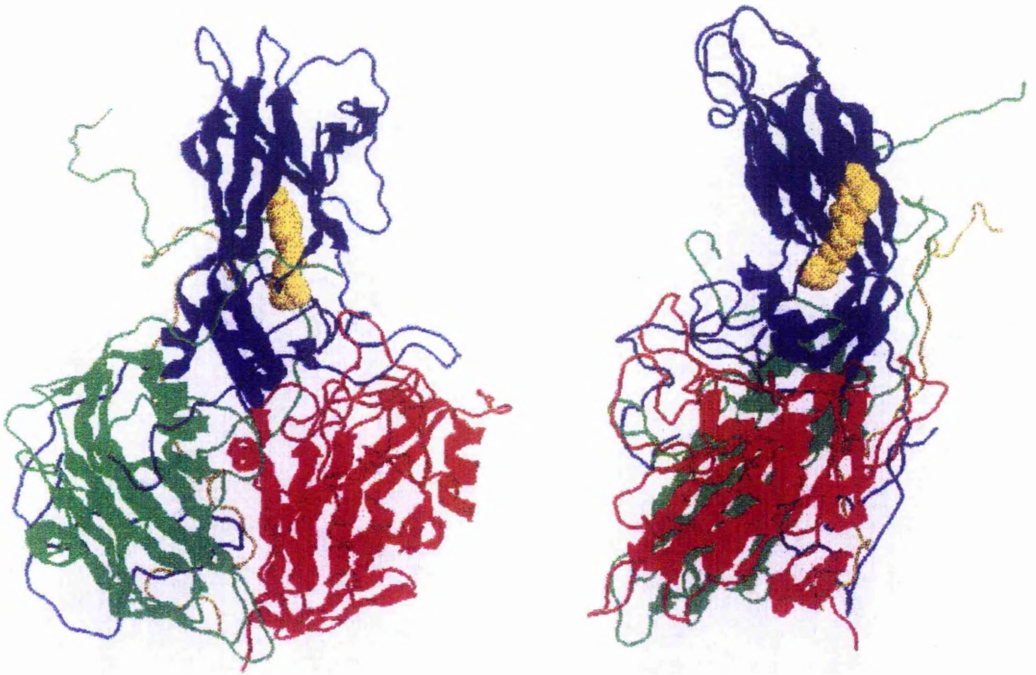
The development of therapeutic agents for the common cold has been hindered by the diversity of rhinovirus serotypes and by the fact that exposure to one serotype affords no immunological protection from other different serotypes. Peptide immunogens have been suggested as possible therapeutic agents for other picornaviruses (Parry *et al.*, 1989). Other antiviral compounds have also been tested (McKinlay *et al.*, 1992) such as inhibitors of attachment and uncoating.

1.2.4.1 Antivirals

Antiviral compounds function by inhibiting different parts of the viral replication cycle. The compound WIN 54954 is an antipicornaviral compound which binds in the hydrophobic pocket of the capsid protein VP1. It induces a conformational change which stabilises the virus and prevents uncoating as well as inhibiting attack of the host cell by binding in the canyon, the receptor binding site. The floor of the canyon is raised by the compound binding in the pocket beneath and displacing the pocket factor. In plaque reduction assays WIN54954 reduced plaque formation in 50 out of 52 rhinovirus serotypes and inhibited 15 enterovirus serotypes (Woods *et al.*, 1989). In humans effectiveness has been low. Although the drug was well absorbed after oral administration, beneficial effect on infection and illness was lacking (Turner *et al.*, 1993). The lack of efficacy was thought to be due to an inability to deliver the drug to the site of infection. The binding site of WIN compounds in relation to the viral protomer can be seen in Figure 1.6.

Crystallographic analysis of poliovirus showed the presence of an area of unidentified electron density in the drug binding pocket of serotypes 1 and 3, this was termed the pocket factor. Its structure was found to be that of a long sphingosine-like molecule but it has yet to be chemically characterised (Hogle *et al.*, 1987). In HRV1A and HRV 16 the structure resembles a fatty acid eight or more carbons long (Oliveira *et al.*, 1993). This is shorter than that found in poliovirus but this has also yet to be

Figure 1.6 Rhinovirus protomer with a WIN compound bound in the hydrophobic pocket of VP1. The WIN compound is shown in yellow and the capsid proteins are shown as follows:VP1-blue, VP2-green and VP3-red.



characterised. The role of the pocket factor is thought to be in stabilising the virion on transit from cell to cell (Mosser and Rueckert, 1993).

Other antiviral compounds with different modes of action, such as the benzimidazide derivative enviroxime (Philpotts *et al.*, 1983), have been identified but they have exhibited similar problems with their use. A reduction in clinical evidence of infection was seen after therapy but in a trial on volunteers who did not have symptoms when trials commenced, no enhancement of effects of enviroxime could be demonstrated.

One of the most recent antiviral compounds to be developed is one which inhibits the action of the 3C protease by binding to it and rendering it inactive (Guterman, 1998). AG7088 has been developed by Agouron Pharmaceuticals. It has been tested *in vitro* on 46 serotypes of rhinovirus and appears to be effective in preventing 90% of infected cells from dying. It also produces a significant decrease in levels of interleukin-6 (IL-6) and IL-8 in cells infected with rhinovirus indicating a potential to diminish the clinical symptoms of infection. It is now in phase 1 clinical trials in England (Press release-Agouron Pharmaceuticals, 1998). AG7088 is also effective in inhibiting replication in other viruses such as some coxsackieviruses, enterovirus 70 and echovirus 11.

1.2.4.2 Zinc

Zinc ions have been shown to be effective in reducing the length of a common cold infection significantly (Alnakib *et al.*, 1987; Farr *et al.*, 1987; Farr and Gwaltney, 1987). Zinc is thought to complex with the ICAM-1 binding sites on the virus surface. This hampers cell binding by the virus and interrupts infection of the cell (Novick*et al.*, 1996). The effect of zinc on the replication of rhinoviruses was first observed in 1974 (Korant *et al.*, 1974). The most dramatic effect was in the inhibition of cleavages of the polypeptides in HRV1A.

1.2.4.3 Vaccines

Monovalent and multivalent vaccines have been evaluated against artificial rhinovirus challenge in volunteers. Reduction in virus shedding and severity of illness was seen but illness and infection were not prevented. Live attenuated vaccine would require the use of multivalent preparations and dual infection has been induced experimentally. A live virus vaccine approach has not been pursued.

The problem with the vaccine approach for rhinovirus is that immunity to one serotype does not afford protection to other serotypes and there are over 100 different rhinovirus serotypes known at the moment. Most of these have epidemiologic significance so all are equally valid candidates for a vaccine approach and several vaccines would be required to provide immunity for all serotypes. In studies of communities it was found that a limited number of serotypes were prevalent at any one time. However the prevalent serotypes did change from year to year. A small number of serotypes cause most illness but all known serotypes are circulating.

1.2.4.4 Interferon

Interferon has been proposed as the primary mediator of recovery from rhinovirus infection and its use in prevention or treatment of infection has been investigated. The production of large quantities of interferon by recombinant DNA technology has allowed the use of higher doses for therapy. A high degree of effectiveness has been seen in prevention of illness. Volunteers were given alpha-2 interferon for 5 days and then challenged with virus (Hayden and Gwaltney, 1984; Samo *et al.*, 1983). Side effects are seen if the therapy is continued for long periods. Interferon α is more effective than interferon β in preventing natural rhinovirus colds but trials *in vitro* have shown comparable activity (Sperber *et al.*, 1989). Therapy of established infection has not been successful even with high doses of interferon.

1.2.4.5 ICAM-1

Soluble ICAM-1 has also been proposed as an alternative therapy. The virus attaches to the soluble ICAM-1 molecules rather than attaching to ICAM-1 molecules on the cell surface. This means that replication cannot continue and the spread of infection is slowed. Trials have been carried out in chimpanzees which appear to have been successful (Huguenel *et al.*, 1997).

1.2.5 Human Rhinovirus 14

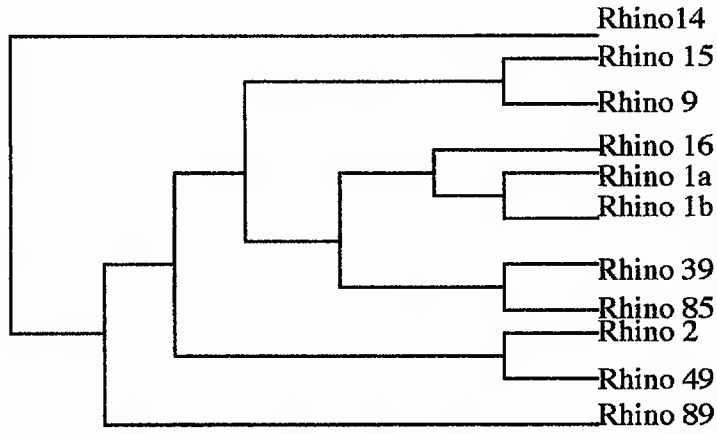
Human Rhinovirus 14 (HRV 14) is a member of the rhinovirus group but does not have much sequence similarity with the other members of the group. A phylogenetic analysis of the group shows the distinct nature of this virus (Figure 1.7). It has been classified in species HRV-B with serotypes 3 and 72. Classification of all serotypes has not yet been completed. The complete sequence of HRV 14 was derived by two groups, Callahan and co-workers in 1985 and Stanway and co-workers (1984). It was shown to have 44-65% nucleotide homology with poliovirus and is adenosine rich. It favours A or U at the third position of its codons. It has an open reading frame of 6537 nucleotides and two non-coding regions. The 5' NCR is 624 nucleotides in length and the 3' NCR is much shorter (47 nucleotides). HRV14 uses ICAM-1 as its cellular receptor and this puts it in the major receptor group of rhinoviruses.

Translation of genome length HRV14 in HeLa cell extract-supplemented rabbit reticulocyte lysate systems is poor but if the 5' NCR is replaced with that of poliovirus, translation efficiency improves (Todd *et al.*, 1997). The recombinant virus retains the growth characteristics of the parental wild type but IRES-mediated translation increases. The inefficiency of the HRV IRES is not a rate-limiting step in rhinovirus assembly but packaging of the viral genome or capsid assembly seems to be limiting.

HRV proteins involved in replication recognise the poliovirus IRES and translation is more efficient but this may be due to the HeLa cell extract being optimised for poliovirus. Poliovirus does not recognise the HRV IRES. The lack of increase in efficiency when the cloverleaf is substituted alone suggests that there are other *cis*-acting signals which affect efficiency. If the whole 5' NCR is substituted these long range interactions are preserved allowing an increase in efficiency.

The secondary structure of the 3' NCR of rhinoviruses was predicted by Pilipenko and colleagues in 1990. It has been shown to have a single stem-loop structural motif (Todd and Semler, 1996) which is conserved amongst all rhinoviruses. The stem-loop structure is not required for infectivity and not absolutely required for RNA replication

Figure 1.7 Phylogenetic analysis of members of the rhinovirus group. The distinct nature of HRV14 can be seen by its position on a separate branch from the rest of the HRVs analysed. Unrooted cladogram of rhinoviruses from aligned genomic RNAs. 11 x 7290 bases. Mininum tree length : 10045. Constructed via PAUP program package. Adapted from A.Palmenberg, Inst. Mol. Virology, UW-Madison, Madison, W.I



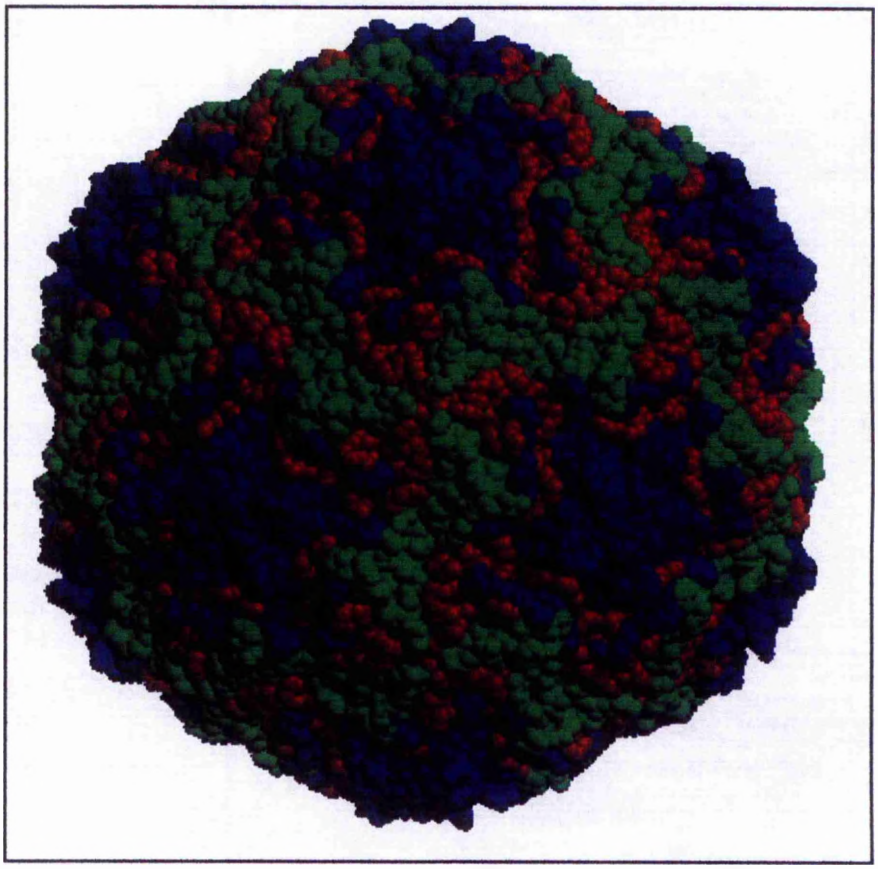
Efficient replication does, however require the 3'NCR in a *cis*-active function. Deletion of all but the first 7 nucleotides of the 3'NCR results in debilitated RNA but does not completely abolish its function. In poliovirus pseudoknots and RNA-RNA interactions are seen in this region (Jacobson *et al.*, 1993) but these do not occur in rhinovirus.

Substitution of most of the P1 region of rhinovirus with reporter genes resulted in replication deficient viruses despite efficient primary cleavage of the polyprotein. Constructs which had deletions involving the 5' nucleotides replicated efficiently but those involving the 3' end did not. Residues 2121-2724 of the virus are required for replication (McKnight and Lemon, 1996). This region has been identified as a *cis*--acting replication element (*cre*). The minimal sequence in this region required for replication is 83-96 nucleotides in length and located between nucleotides 2318-2413 of the genome (McKnight and Lemon, 1998) Translation of the *cre* sequence is not required for replication as it functions as an RNA entity. Analysis of mutant RNAs has shown that only the predicted positive strand structure is essential for replication. The lack of negative strand synthesis from RNAs without the *cre* suggests that it is required for initiation of negative strand synthesis and is able to participate in the long range interactions required for this.

1.2.5.1 Atomic Structure

The first atomic resolution of the structure of the HRV 14 capsid was reported in 1985 by Rossmann and co-workers. The structure of VP1, VP2 and VP3 is very similar, all comprising 8-stranded, anti-parallel β-barrels. The major differences in their conformations are in the external loops that connect the β-strands. Protrusions on VP1, 2 and 3 create a cleft or canyon on the surface of the viral particle. VP1 contributes most of the hydrophobic residues which line the canyon. The canyon is 25 Å deep and between 12 and 30 Å wide. A computer reconstruction of the capsid structure can be seen in Figure 1.8.

Figure 1.8 HRV14 capsid structure coloured by its proteins. Solved by cryo-electron microscopy and image reconstruction. The image is coloured to show the capsid proteins, VP1 (blue), VP2 (green) and VP3 (red). Taken from www.bocklabs.wisc.edu



Several lines of evidence have supported the idea that the canyon is involved in receptor binding. Site directed mutagenesis of residues forming the canyon floor revealed several sites which were important for receptor binding (Colonno *et al.*, 1988). Deformations in the canyon floor caused by antiviral drugs were also shown to eliminate cell receptor attachment (Pevear *et al.*, 1989), further implicating the canyon in receptor binding. Eventually cryomicroscopy of HRV16 with the amino-terminal immunoglobulin-like domains of its receptor ICAM-1 provided direct proof for the role of the canyon (Olson *et al.*, 1993; Olson *et al.*, 1992). A road map analysis of the HRV14 viral protomer (Figure 1.9) shows the position of the canyon on the surface.

Unlike other picornaviruses HRV14 does not have an area of unidentified electron density in the hydrophobic pocket below its receptor binding site, *i.e.* it does not appear to have a pocket factor. The empty pocket allows binding of antiviral compounds which deform the roof of the pocket, which is also the floor of the receptor binding site (the canyon) and prevent receptor attachment. Another member of the antiviral binding group A, HRV3, has also been shown to have an empty pocket (Zhao *et al.*, 1996). These viruses are similar in their amino acid sequence and structure and this suggests that rhinoviruses in the same antiviral group have similar amino acid sequences and structures. Two viruses from antiviral group B, HRV16 and HRV1A, both have pocket factors (see section 1.2.4.1) and the similar amino acid compostition in the pocket region of all other group B viruses suggests that all members of the group have pocket factors.

1.2.5.2 Immunogenic sites

Experiments by Sherry and Rueckert (1985) and Sherry and co-workers (1986) identified four major neutralisation immunogens on the viral surface. Mouse hybridoma cell lines which secreted monoclonal antibodies that neutralised HRV14 were isolated and used to select mutants resistant to neutralisation by that antibody. As the viruses have a high error rate in replicating, every population has a relatively large number of mutants which can escape antibody neutralisation. Every antibody was then assayed for its ability to neutralise the mutants. The results revealed four major immunogenic neutralisation sites

(NIm sites). When the electron density map of the protein became available it was clear that the substitutions that conferred resistance to neutralisation localised into four distinct areas. These were designated NIm-IA, NIm-IB, NIm-II and NIm-III. Most of these sites surround the large depressions that encircle each five-fold axis.

It appears that residues which line the canyon are more conserved amongst rhinoviruses than the residues which occur on the rim. The narrowness of the canyon inhibits antibodies from binding to the canyon floor and this is thought to play a role in the recognition of the cell receptor binding protein.

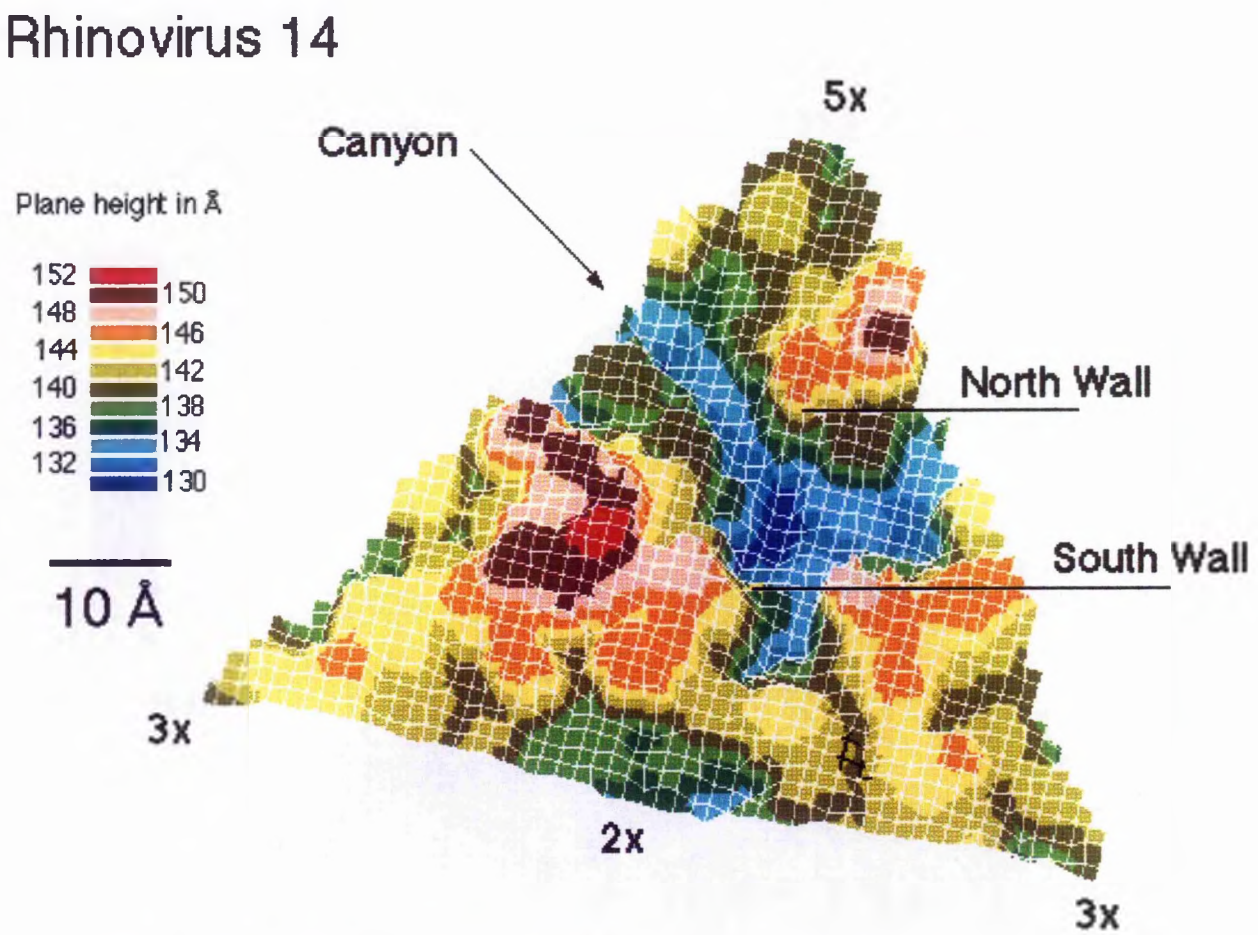


Figure 1.9 Road map of protomer of HRV14 showing canyon binding site. A.C. Palmenberg

1.3 Polyprotein processing

1.3.1 Overview

The 'one gene, one protein' hypothesis of Beadle and Tatum (1941) is still recognised as being correct in most cases. However, there are instances where multiple gene products can be produced from a single mRNA. At a translational level there are several methods in which this is the case:

1) in-phase translation of overlapping genes with read through of termination codons or use of internal in-phase translation start sites.

2)Ribosomal frameshifting.

3)Polyprotein processing.

4) Alternative splicing.

5) Protein splicing

Polyprotein processing allows controlled limited proteolysis to occur in many systems. Encoding proteins in the form of polyproteins is by no means unique to viruses. Examples in eukaryotes include the conversion of zymogens to enzymes and prohormones to hormones. Other examples are the assembly of cytoskeletal components, the control of cellular differentiation, the initiation of cascade mechanisms, such as blood clotting, the removal of signal peptides and recycling of cellular proteins. Polyproteins also occur in prokaryotes although they are not as common as they are in viruses and eukaryotes.

1.3.2 Prokaryotic polyproteins

The first prokaryote polyprotein to be discovered was the *E.coli* penicillin G acylase (Böck *et al.*, 1983). This is composed of two dissimilar subunits, α and β, of 23kDa and 69 kDa respectively. The protein is only functional as a heterodimer but it is encoded by a single gene (pac; Mayer *et al.*, 1979). There is an open reading frame of 840 amino acids which encodes the polyprotein. This is proteolytically processed in the course of maturation. The first step is the removal of the signal peptide and then an N-terminal cleavage of the β subunit. The removal of the spacer peptide generates a free α subunit.

Cleavage is processed by proteases found in the cell or autocatalytically. This was demonstrated by growth in Gram-negative bacterial strains which do not contain the *pac* gene where cleavage still occurred. Similar polyproteins have been found in other bacteria such as *Kluyvera citrophila* (Barbero *et al.*, 1986) where the processing sites are identical, indicating a common ancestor for the *pac* genes. *Pseudomonas* sp. also has a similar system (Matsuda and Komatsu, 1985).

Bradyrhizobium japonicum cytochromes b and c_1 are encoded by the 5' and 3' halves of a single gene, *fbcH* (Thöny-Meyer *et al.*, 1991). The two sections are linked by a spacer region of DNA encoding a signal peptide for translocation across the cytoplasmic membrane. Processing of the polyprotein occurs at a typical signal peptidase recognition site (Ala-Arg-Ala).

1.3.3 Eukaryotic polyproteins

In eukaryotes polyproteins are usually in the form of single proteins which require cleavage of a precursor before they are mature. Some polyproteins are found which produce a number of mature proteins from one polyprotein.

Bioactive peptides can be synthesised in the form of precursors which may be the source of more than one peptide. However these polyproteins do not self-process. One of the first large precursors to be discovered was proopiomelanocortin (POMC; Nakanishi *et al.*, 1977). It is more than 6 times the size of any of the bioactive peptides that it contains and is a precursor for adrenocorticotrophic hormone (ACTH), β-endorphin and α-, β- and γ-melanocyte stimulating hormones (MSH). Cleavage of the polyprotein is thought to be carried out by a trypsin-like enzyme in a distinct order (Figure 1.10).

Polyproteins are also present in many other species including *Neurospora crassa* (Gessert *et al.*, 1994) , *Dictyocaulus viviparus* (a nematode) (Kennedy *et al.*, 1995) and *Bombyx mori* (Sato *et al.*, 1993).

Prepro-α mating factor found in yeast gives another example of multiple polypeptide products from a single polyprotein (Kurjan and Herskowitz, 1982). The yeast α-factor gene has four direct repeats of closely homologous 63 bp sequence. The

gene codes for 13 amino acids with an 8 amino acid spacer region in between, which are α-factors and the α-factor decapeptide. The arrangement may have evolved from the duplication of a unique α-factor and spacer regions. This arrangement is conserved in all species of *Saccharomyces* but there are polymorphisms in length due to differences in the number of spacer units in the coding sequence of the α factor precursor gene. There may be between 3 and 5 copies.

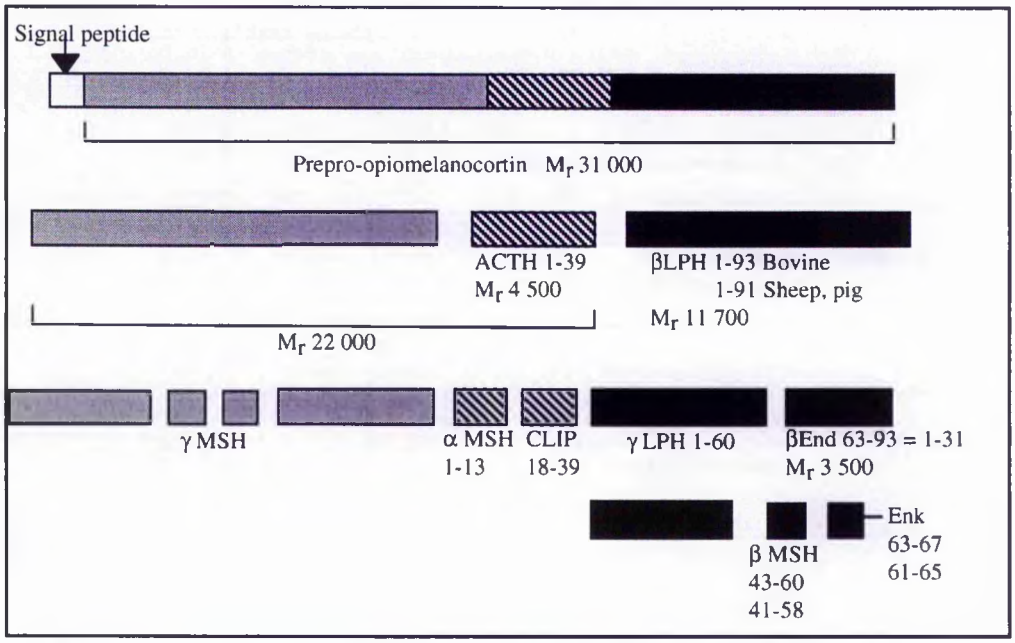


Figure 1.10 Cleavage of prepro-opiomelanocortin (POMC).

1.3.4 Virus polyproteins

The use of polyproteins by viruses is widespread especially among the smaller RNA viruses. Their small capsids need to contain a great deal of genetic material and polyproteins are a convenient way of producing several functional proteins from the minimum of genetic information. This has the disadvantage that all proteins are produced in equimolar amounts. A polyprotein only requires one site of initiation and eliminates the need for intergenic sequences, such as promoters for each gene, which take up valuable space. Polyproteins can encode their own protease eliminating the need for cellular enzymes or separate proteases. Polyproteins are also useful for regulating cleavage if it needs to be performed at a particular site or stage of virion morphogenesis.

1.3.4.1 Polyprotein processing in plant viruses

Picornaviruses have several relatives in the plant virus kingdom. Two of the most widely studied are cowpea mosaic virus (CPMV) and tobacco etch virus (TEV; Figure 1.11) belonging to the comoviruses and potyviruses respectively. There are several similarities between the genomes of these viruses. All genome sequences have a polyA tail at the 3' end. Poliovirus and the potyviruses have a VPg at the 5' end and no cap. They have the same order of non-structural domains - putative helicase, VPg, protease and polymerase. The capsid proteins of poty- and picornaviruses have no sequence similarity. The structures of the capsid proteins of four picornaviruses and two comoviruses (CPMV and bean pod mottle virus) are, however, very similar with all having an 8-stranded antiparallel β-barrel motif (Chen *et al.*, 1989). Although, the capsid of the picornaviruses has three principal proteins and that of the comoviruses two, their virions have the same geometry.

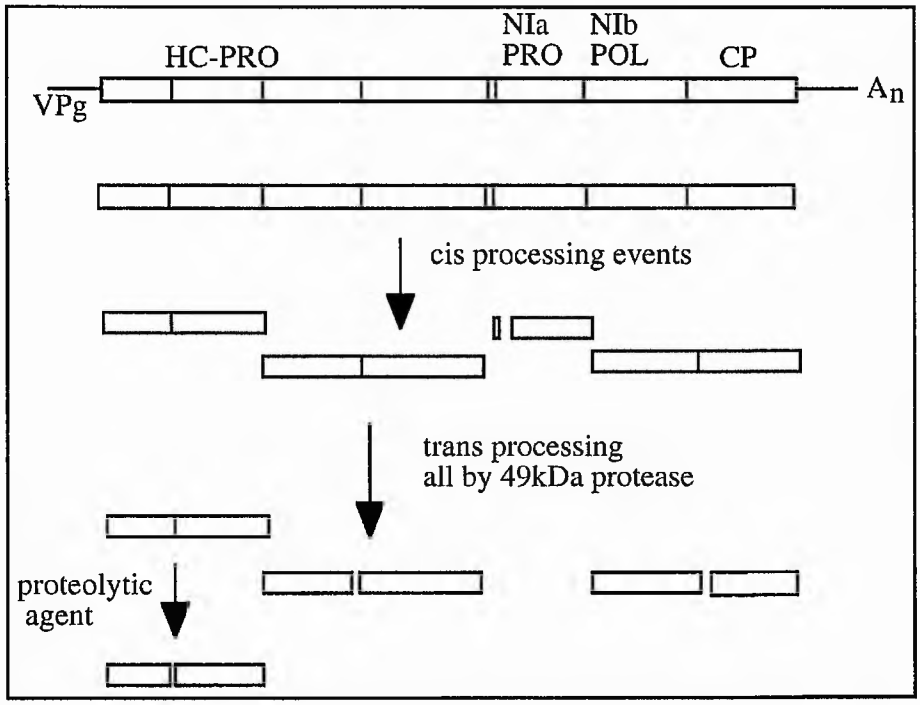


Figure 1.11 Processing in TEV

The primary cleavage between encapsidative and replicative domains of picornaviruses is mimicked in the como- and nepoviruses by the division of the genome

separating these two functions. The potyviruses have a rapid self-cleavage on the C-terminal side of the HC proteinase which performs the same function. The division occurs co-translationally and by an intramolecular mechanism. Subsequent steps in the polyprotein processing pathway are carried out by enzymes with homology to the picornavirus 3C proteinase. There are cleavage events in both como- and potyviruses for which no proteolytic activity has been defined. These are similar to the maturation cleavage of 1AB in poliovirus.

The 24 kDa proteinase in CPMV is the major proteolytic enzyme but in order to process the capsid precursor it has to act as the 32 kDa precursor protein. This is similar to the action of $3CD^{pro}$ in the processing of the capsid proteins in poliovirus. Comoviruses have a bipartite positive strand RNA genome, each strand encoding 1 long open reading frame. One of these encodes for the capsid proteins and the other for the non-structural proteins. The cleavage sites are similar to those in picornaviruses, with residues at the P4 and P1 positions having particular influence (see Figure 1.12 for nomenclature). The potyviruses have a monopartite genome with one long open reading frame which encodes 8 mature proteins (Dougherty *et al.*, 1988).

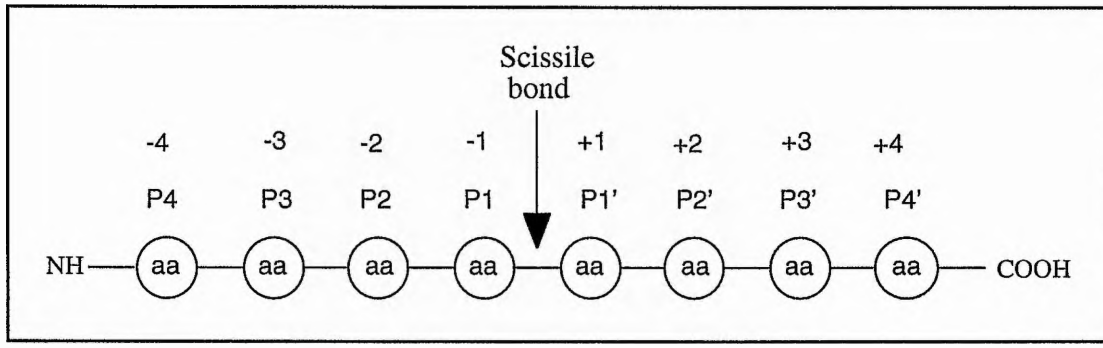


Figure 1.12 Substrate cleavage site nomenclature after (Schechter and Berger, 1967).

1.3.4.2 Polyprotein processing in retroviruses

Retroviruses are a class of viruses that replicate *via* an obligatory DNA intermediate. All have the same gene order, 5'-*gag-pol-env*- 3'. The *gag* (**g**roup specific **a**ntigen) region encodes up to 6 structural proteins. The *pol* region encodes the viral replication enzymes. The protease cleaves at the gag-pol junction. This separates the

structural and non-structural proteins as does the primary cleavage in picornaviruses. Protein processing is necessary for a productive infectious cycle. This was demonstrated in an avian retrovirus (Hayman *et al.*, 1979). Viruses which were lacking the viral protease (PR) did not have any processing of gag. The protease is an aspartyl protease (see section 1.4.2.1).

Polyprotein processing in Rous sarcoma virus (RSV; Figure 1.13) involves a single frameshift for the expression of gag and pol from a single reading frame.

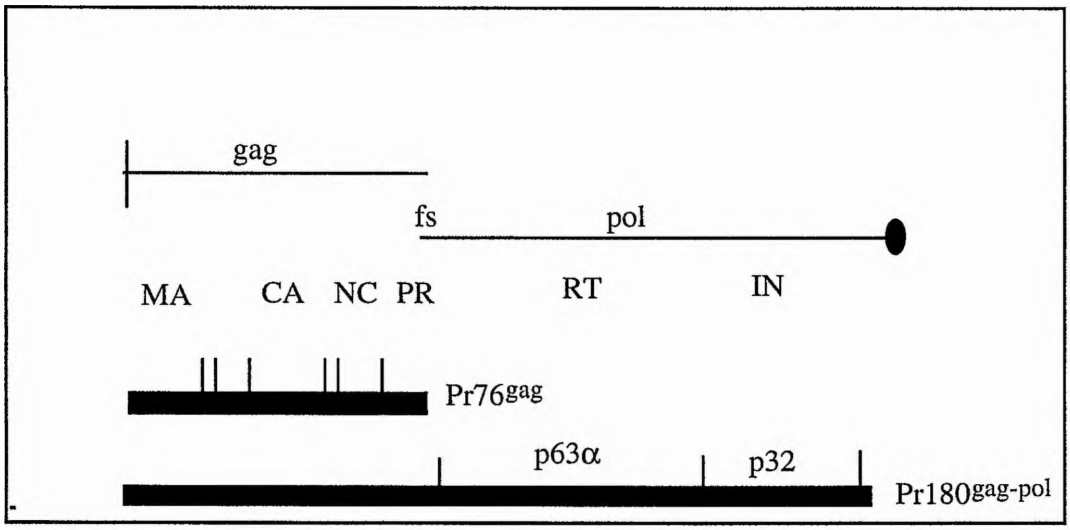


Figure 1.13 Processing in RSV. RSV utilises a single frameshift in its expression of *pro* and *pol* from a single *gag -pro -pol* translational unit. The bold lines represent the primary translation products which are further processed to give the products indicated.

1.3.4.3 Polyprotein processing in DNA viruses

It was shown that some proteins (p150, p37 and p34) in African Swine Fever Virus (ASFV) are synthesised as precursors of high molecular weight (López-Otìn *et al.*, 1989). The proteins are mature after proteolytic processing. Cleavage occurs at a consensus Gly-Gly-X site with cleavage occurring after the second glycine residue. This site is also seen in adenovirus structural proteins and some cellular proteins such as polyubiquitin. There are 16 other Gly-Gly-X sites besides those that are cleaved. There is an ordered cleavage cascade (Fig 1.14). All the mature proteins are structural and their production may be

linked to virion morphogenesis. The polyprotein (pp220) is encoded by a single ORF of 2475 amino acids (Simòn-Mateo *et al.*, 1993).

Another polyprotein has also been identified in ASFV (Simòn-Mateo *et al.*, 1997). pp62 is cleaved to give the structural proteins p35 and p15. The cleavage occurs at the same Gly-Gly-X site as in pp220 and there is an ordered cascade. As in pp220 only some of the Gly-Gly-X sites are cleaved suggesting that other residues may have an effect on the conformation of the cleavage site. However, there is no homology between the surrounding sequences of the cleavage sites in pp220 and pp62.

Production of structural proteins by processing of a polyprotein may be a way of ensuring a 1:1 ratio of structural proteins. Equimolar amounts of the proteins from pp220 are required in the virus particle and they all occupy similar locations in the virus shell (Simòn-Mateo *et al.*, 1997).

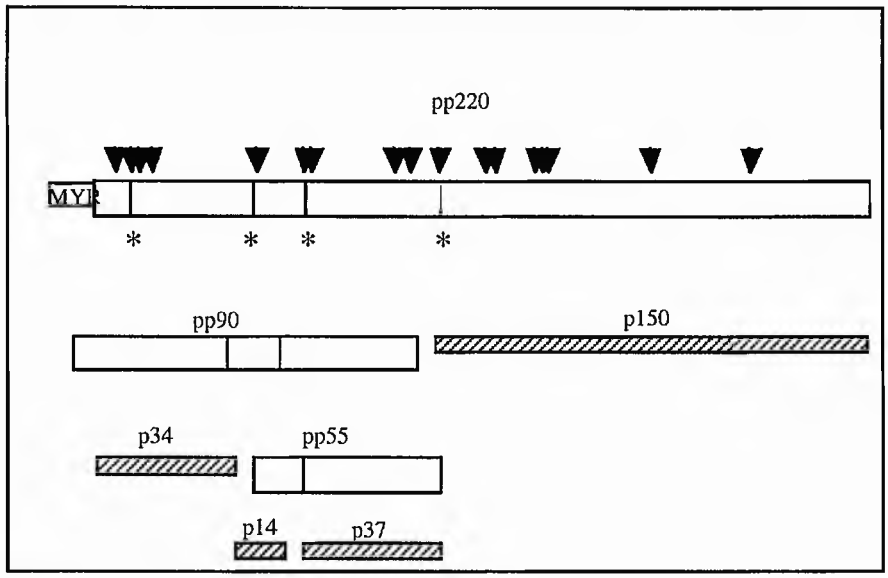


Figure 1.14 Cleavage of ASFV polyprotein pp220. Triangles indicate the position of Gly-Gly-X sites and cleavage sites are marked with an asterisk. Myristoylation is denoted by the region marked MYR. Precursor proteins are denoted by open boxes and mature proteins by shaded boxes.

1.3.5 Picornaviruses

In order to produce mature viral proteins the polyprotein in picornaviruses is cleaved autoproteolytically (Figure 1.15). The full length polyprotein is not seen because of cleavage beginning before translation of the polyprotein is complete. In rhinoviruses three different proteolytic activities are seen in the processing pathway. These are: cleavage by $3C^{pro}$, cleavage by $2A^{pro}$ and the maturation cleavage. These events take place both cotranslationally and posttranslationally. They result in the production of both structural and non-structural mature virus proteins.

1.3.5.1 Primary processing

The primary cleavages in the polyprotein yield protein precursors P1, P2 and P3. Cleavage between P1 and P2 in rhino- and enteroviruses is carried out by $2A^{pro}$ (Figure 1.15; Sommergruber *et al.*, 1989; Toyoda *et al.*, 1986). The 2A protease does not need to be released from the polyprotein to carry out this cleavage since it is acting *in cis*. The reaction takes place before $3C^{pro}$ has been translated and it is therefore impossible for $2A^{pro}$ to be released as this is a cleavage performed by $3C^{pro}$. The 2A protease is also responsible for the alternative cleavage pathway that yields 3C' and 3D' (Toyoda *et al.*, 1986). It has similarity with the small, bacterial, serine-type proteases *e.g.* subtilisin (Bazan and Fletterick, 1988). Apart from effecting the P1/P2 cleavage it has been implicated in host-cell shut-off. Infection by HRV14 results in cleavage of the p220 subunit of the cap binding protein complex (CBC; Etchison and Fout, 1985), a protein that is involved in the recognition of capped mRNAs during the initiation of protein biosynthesis. This correlates with inhibition of host cell protein synthesis by HRV14 infection and the inability of initiation factors to function on capped globin mRNA *in vitro*.

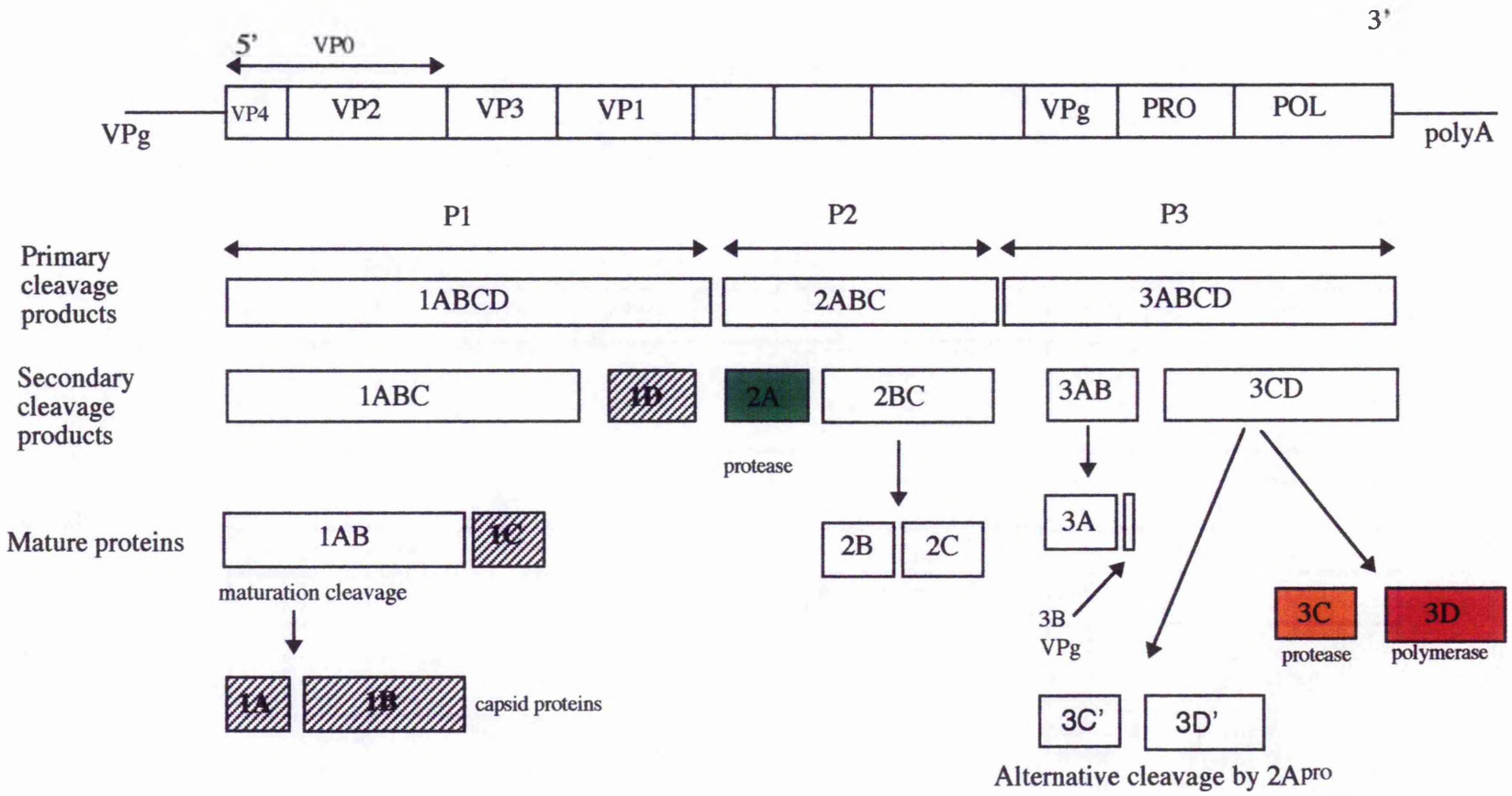


Figure 1.15 Proteolytic processing in HRV14. The 2A protease is shown in green, the 3C protease in orange and the viral polymerase, $3D^{pol}$, in red. Capsid proteins are shown as grey shaded boxes.

Leader proteins which precede the P1 region are found in some picornaviruses including the aphtho- and cardioviruses. In EMCV and TMEV (cardioviruses) cleavage between the L protein and P1 is mediated by $3C^{pro}$ as the first step in L-P1-2A processing (Parks*et al.*, 1989;Roos*et al.*, 1989). In FMDV (an aphthovirus) the L protease (L^{pro}) carries out its own cleavage from P1 in an undefined reaction.

Primary cleavage in aphtho- and cardioviruses occurs at a different location to that in rhino- and enteroviruses (see sections 1.5.2.1. and 1.5.2.2). Cleavage is still mediated by 2A but there are a number of significant differences, 2A in aphthoviruses is a great deal smaller than in other picornaviruses. The remaining 18 amino acids of the aphthoviral 2A show similarity to the carboxyl portions of cardiovirus $2A^{pro}$ but there is no similarity to the catalytic triad of rhino and entero $2A^{pro}$s. The primary cleavage occurs between 2A and 2B generating a larger primary cleavage product, P1-2A. In FMDV this cleavage occurs at a glycyl -prolyl (G-P) amino acid pair.

1.3.5.2 Secondary processing

Cleavage between P2 and P3 is mediated by $3C^{pro}$. All subsequent cleavages (except the maturation cleavage) are also carried out by $3C^{pro}$. $3C^{pro}$ has been characterised as a cysteine protease with a serine protease type fold (Allaire *et al.*, 1994). Cleavages process P3 to smaller products which are more stable and are the non-structural proteins of the virus. These reactions are dilution insensitive (Palmenberg and Rueckert, 1982) which implies a monomolecular mechanism. Cleavage may also take place *in trans* with exogenous $3C^{pro}$ if the cleavage sites have been destroyed in a binary system (Parks *et al.*, 1986; Ypma-Wong and Semler, 1987).

Cleavage between 3A and 3B is not observed *in vitro* unless all other processing routes are blocked (Parks *et al.*, 1986). This occurs in both mono- and bimolecular assays. It is thought that these reactions take place with the addition of VPg to the 5' end of the RNA during the initiation of replication. The reaction may therefore be facilitated by the presence of other proteins such as $3D^{pol}$ or configurations of reaction components.

P1 is also processed by $3C^{pro}$. In EMCV all P3 3C-containing proteins can process P1 (Parks *et al.*, 1986). P1 cleavage sites are processed sequentially in the order; L-VP1, VP1-2A, VP3-VP1 and VP0-VP3. In rhino- and enteroviruses different parts of P3 have different enzyme specificities. VP3-VP1 may be processed by $3C^{pro}$ but VP0-VP3 requires $3CD^{pro}$ (Jore *et al.*, 1988; Ypma-Wong *et al.*, 1988).

Cleavage by $3C^{pro}$ only occurs at specific points in the polyprotein of the wild type poliovirus. It was originally thought that cleavage occurred only at glutamine - glycine (Q-G) pairs in the polyprotein (Hanecak *et al.*, 1984; Kitamura *et al.*, 1980). However as more virus sequences were discovered it became clear that this was not always the case. Most cleavage sites for $3C^{pro}$ are now known to be of the pattern (E, Q), (G,S, A) but there are exceptions to this. There are influences from the surrounding sequences and the structural conformation of the cleavage site is also important. A minimum substrate length of 6 amino acids has been determined (Cordingley *et al.*, 1990) for peptides designed to mimic the 2C-3A cleavage site in HRV14. It was found that residues at positions P4, P1' and P2' could not be substituted but the residue at position P5 was not critical. Small peptides made to mimic the 3A-3B cleavage site made poor substrates. This suggested that surrounding sequences were important for susceptibility to cleavage but as the 3AB cleavage is rarely seen *in vitro* (Parks *et al.*,) the low level of cleavage may be the level seen *in vitro*. HRV14 $3C^{pro}$ was also capable of cleaving peptides representing cleavage sites of coxsackie B virus and poliovirus (Cordingley *et al.*, 1990).

1.3.5.3 Products of polyprotein processing

The mature protein products of the cleavage cascade are either structural or non-structural proteins. Proteins 1A-1D are the capsid proteins VP1-VP4. $2A^{pro}$ is responsible for the P1/P2 cleavage and host cell shut-off, 2C is a putative NTPase. Enterovirus protein 2B has recently been found to enhance membrane permeability in bacterial and mammalian cells when transiently expressed as well as blocking the protein secretory traffic (Doedens and Kirkegaard, 1995; Lama and Carrasco, 1992; van

Kuppeveld *et al.*, 1997a, 1997b) and localising to the outer surface of vesicles derived from endoplasmic reticulum which are sites of viral genome replication (Bienz *et al.*, 1987; Schlegel *et al.*, 1996; van Kuppeveld *et al.*, 1997a). 3B is the VPg, $3C^{pro}$ is the protease, $3D^{pol}$ is the polymerase and $3CD^{pro}$ is a protease. $3CD^{pro}$ is also the viral component of the RNP complex formed with the 5' end of positive strand RNA (Colonno *et al.*, 1988) This may act *in trans* to catalyse the initiation of synthesis of new positive strand RNA. $3CD^{pro}$ has also been shown to be involved in the cleavage of the P1 capsid precursor in poliovirus (Ypma-Wong and Semler, 1987).

3AB has been shown to have multiple functions in the replication of poliovirus. It was shown to have a dual function as both the precursor for VPg and a co-factor for $3D^{pol}$ (Lama *et al.*, 1994). The soluble form of 3AB was found to stimulate poly(U) RNA synthesis catalysed by $3D^{pol}$. Amino acid changes in the hydrophobic region of 3A result in poliovirus mutants with an impaired RNA replication phenotype (Giachetti 1991).

The action of 3AB may be *in vivo* by binding to the $3D^{pol}$ part of $3CD^{pro}$ which is bound to the initiation complex at the 5' end of the poliovirus genomic RNA (Andino *et al.*, 1993; Harris *et al.*, 1994). It enhances the binding of $3D^{pol}$ to RNA and this is supported by experiments showing the binding of $3CD^{pro}$ and $3D^{pol}$ to the pseudoknot at the 3' end of the genomic RNA which requires 3AB.

When in the initiation complex at the 5' end of the genome $3CD^{pro}$ may be able to catalyse the uridylation of 3AB or VPg. Therefore the primer for transcription by $3D^{pol}$ will be generated either by autocleavage of $3CD^{pro}$ or may be present *in trans*.

1.4 Proteases

Proteases are of four main types categorised by their nucleophilic residue: serine, cysteine (or thiol), acid (or aspartyl) and metallo-proteases. A fifth type has also been identified with a threonine residue as the nucleophile but little has been elucidated about them.

All proteases share the property of going through an intermediate or transition state. In serine and cysteine proteases the trigonal carbon of the peptide bond becomes tetrahedral due to the (temporary) covalent bond formed between the enzyme and the subtstrate. In serine and cysteine proteases the nucleophile is the hydroxyl group of the serine or the thiol group of the cysteine respectively. In the other groups the nucleophile is a water molecule. Hydrolysis of the bond occurs in two steps in the serine and cysteine proteases whereas in the other classes it is a single step process.

1.4.1 Cellular

Both endo- and exo-proteases are found in cells. Exo-proteases remove residues from either the N- or C-terminus of the protein and usually continue in a progressive fashion. Endo-proteases cleave peptide bonds between residues internal to the substrate sequence. Examples of the five classes of cellular proteases are shown below.

Table 1.2. Classes of proteases

Class	Examples	Inhibitor
Aspartic	pepsin, gastricin, renin, cathepsin D, cathepsin E	pepstatin
Cysteine	papain, cathepsin L	E-64, cystatins
	cathepsin B	leupeptin
	cathepsin H	p-chloromercuribenzoate
Metallo-	collagenases, meprin	EDTA, phrenanthroline
	thermolysin, EC24.11	phosphoramidon
Serine	trypsin-like	DFP, leupeptin
	chymotrypsin-like	DFP, chymostatin
	elastase-like	DFP, elastinal
Threonine	*Thermoplamsma acidophilum* 20S proteasome	

1.4.1.1 Serine proteases

Serine proteases may be divided into two sub-groups. The members of the' large ' sub-group are similar to trypsin and the members of the ' small ' sub-group have similarities to subtilisin. The serine residue, from which the group derives its name, is unreactive with the hydroxyl group at the end of a catalytic "triad". The catalytic triad of these proteases is usually composed of a histidine, a serine and an aspartate residue (Figure 1.16). The order of the residues (from the N- to the C-terminus) in the catalytic triad is dependent on which sub-group the protease belongs to.

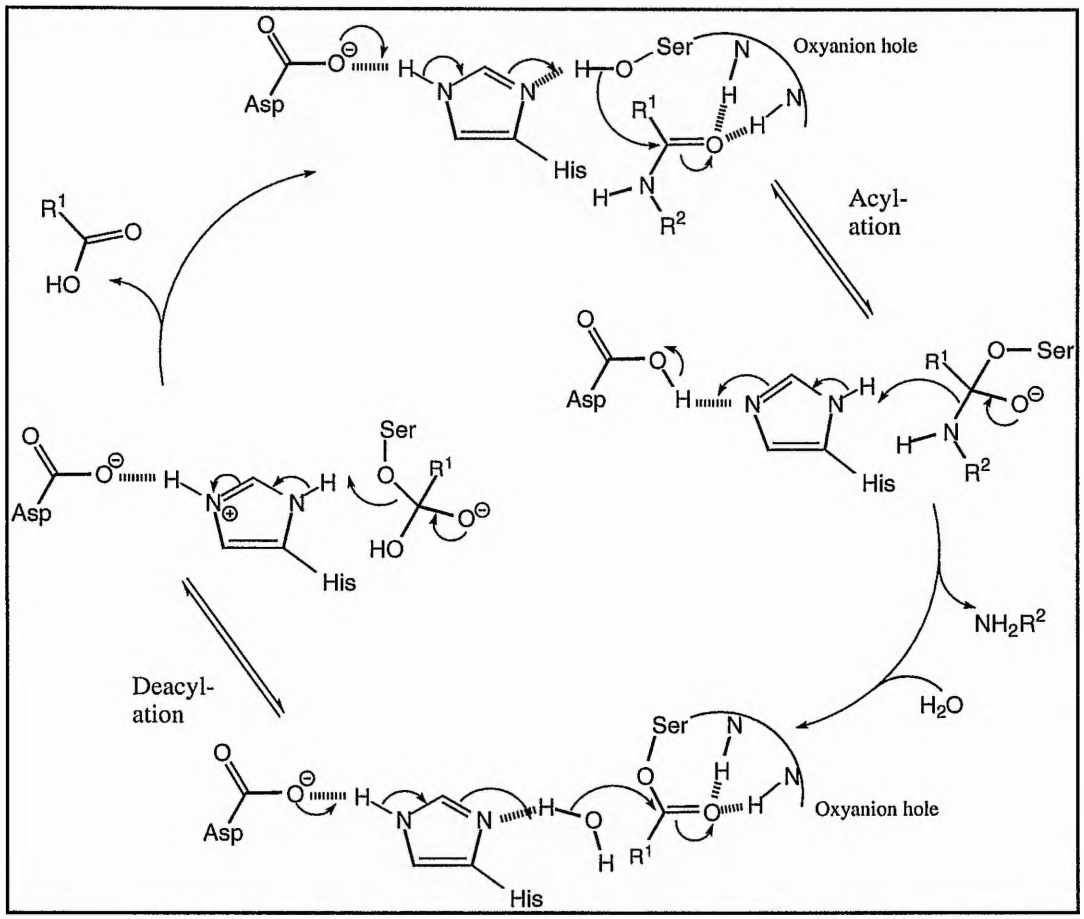


Figure 1.16 Mechanism of action of a serine protease.

In the ' large ' sub-group the order of residues in the primary sequence is His, Asp, Ser as opposed to Asp, His, Ser in the ' small ' sub-group. This makes it unlikely that they are derived from a common ancestor. The two classes also differ in their

preference for amino acids at the P_1 position. Trypsin cleaves only after lysine or arginine residues but chymotrypsin cleaves after large hydrophobic residues.

The oxyanion hole is a site containing dipoles which complement the changes in charge distribution during catalysis. In trypsin the dipoles are formed by the serine and glycine residues. The substrate binding pocket is composed of substrate sites which bind other residues. These can be up to four residues on either side of the scissile bond.

These proteases have been reviewed by Lesk and Fordham (1996). The catalytic mechanism was elucidated by Steitz and Schulman in 1982 and is common to all serine proteases in both sub-groups (Warshel *et al.*, 1989).

1.4.1.2 Cysteine (Thiol) proteases

Papain and actinidin are cysteine proteases whose structure is known in detail. Their catalytic mechanism is similar to that of the serine proteases; formation of a tetrahedral intermediate and an acylenzyme complex. The enzymes have two structural domains with approximately 210 residues in each. These each form half of the polypeptide chain. The active site is in the cleft between the two domains and consists of seven sub-sites, S_4, S_3, S_2, S_1, S_1', S_2' and S_3'. The substrate specificity arises from interactions at S_1 and S_1'. The catalytic activity is pH dependent.

The sulphydryl group of the active site nucleophile of these proteases plays an important part in catalysis. The sulphur atom forms a complex with the histidine residue at neutral pH to give a thiolate/imidazolium ion.

In papain the amide group of the Asn^{175} side chain is hydrogen bonded to the histidine forming an Asn, His, Cys catalytic triad. There is some rotation of the residue without disturbing the bond. It has been suggested that the function of the Asn residue (sometimes suggested as the third member of the catalytic triad) is to limit rotation.

The oxyanion hole has less effect than in the serine proteases (Menard*et al.*, 1995). Dipoles are formed from the Cys^{25} and Gln^{19}.

1.4.1.3 Aspartyl (acid) proteases

Some acid proteases such as pepsin and gastricin are involved in digestion and only have limited substrate specificities. Others cleave at single sites. Renin, for example, removes the decapeptide angiotensin 1 (which regulates blood pressure) from the amino terminus of angiotensinogen.

The structures of known acid proteases are similar, having a polypeptide chain of approximately 330 residues. This is folded into a bi-lobed structure with two domains of similar structure. This suggests that an ancestral protein was a dimer with two identical chains. The two domains in the protease are related by two-fold rotational symmetry. These proteases have a catalytic diad formed from two aspartate residues (Figure 1.17). Catalysis is of the basic acid-base mechanism. A water molecule attacks the carbonyl group of the peptide bond and aspartate side chains assist in the protonation of the carbonyl oxygen and the protonation of the amine leaving group to form cleaved products. At no stage is the substrate covalently bound to the enzyme.

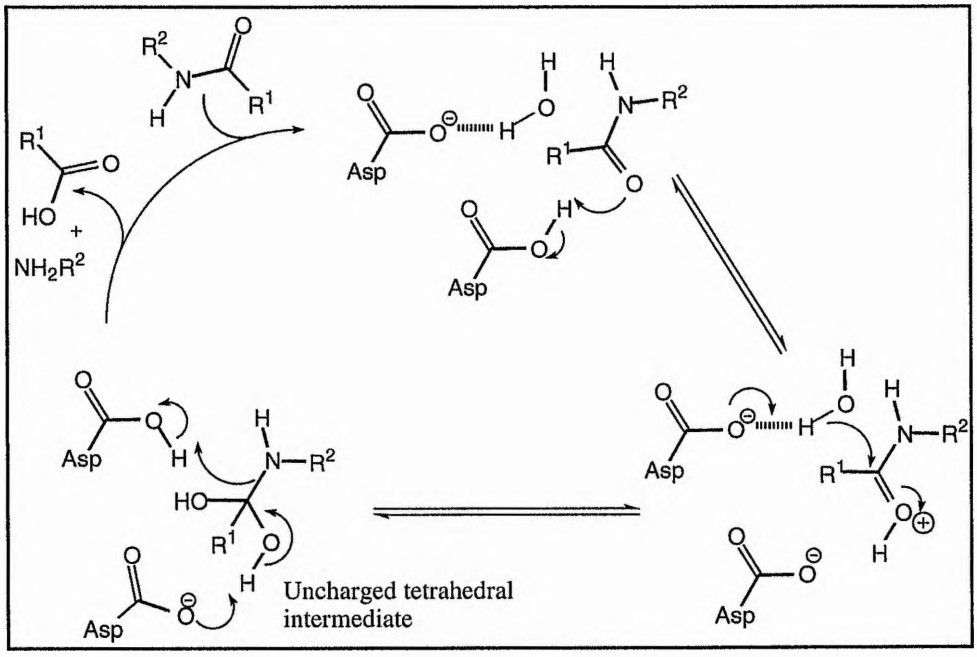


Figure 1.17 Mechanism of action of an aspartyl protease.

1.4.1.4 Metallo -proteases

Most of these proteases are thought to rely on zinc bound in their active sites, but in some cases they are still active when the zinc is replaced with cobalt. This class includes some medically important enzymes including angiotensin-converting enzyme, enkephalinase and collagenase.

Carboxypeptidases A and B are the most well studied of the class. They catalyse the removal of single residues from the carboxyl terminus of polypeptide substrates making them exopeptidases. They show specificity for the P_1' residue. In Carboxypeptidase A the zinc ion is complexed to two histidine residues and a glutamate residue (Figure 1.18). A water molecule is also present. The nucleophilic attack of the Zn^{2+} bound water molecule on the carbonyl bond causes cleavage. The positively charged arginine residue stabilises the oxyanion hole and the glutamate residue deprotonates the zinc bound water. At no stage is the substrate covalently bound to the enzyme.

Thermolysin is another member of the class which acts as an endopeptidase. It catalyses the hydrolysis of non-terminal hydrophobic P_1' residues. It has an active site similar to that of the carboxypeptidases which indicates convergent evolution. The Zn^{2+} ion is bound in both types by interactions with the imidazole side chains of two histidine residues and the carboxyl side chain of a glutamate residue. The water molecule in a co-ordination position for the zinc residue is crucial in the catalytic action of the enzyme.

The enzymes in both groups have similar functions and sequence similarities. The glutamate residue at position 127 in carboxypeptidase corresponds to the glutamate at position 143 in thermolysin and Arg^{127} in carboxypeptidase corresponds to His^{231} in thermolysin. They also have similar electrostatic interactions with the zinc residue.

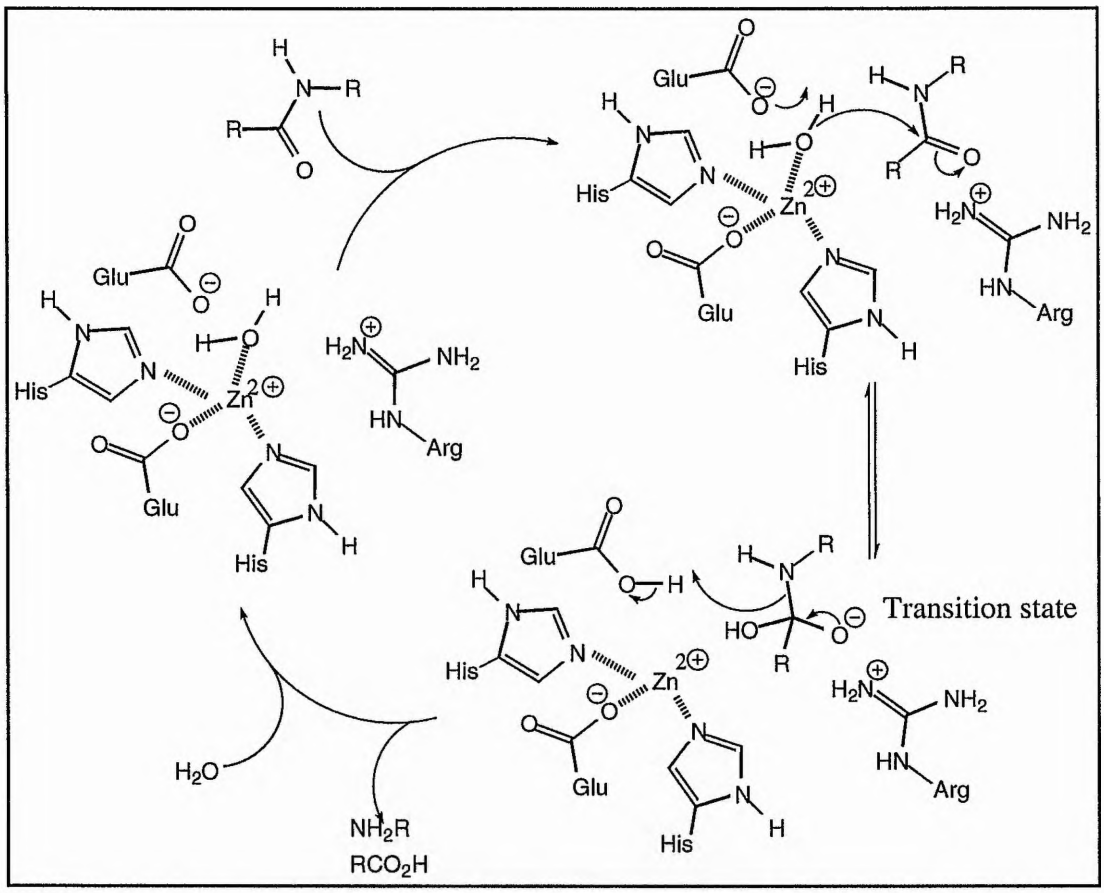


Figure 1.18 Mechanism of action of Carboxypeptidase A

1.4.1.5 Threonine proteases

Study of the 20S proteasome, found in both eukaryotes and the archaebacterium *Thermoplasma acidophilum*, has revealed proteolytic activity. The 20S proteasome forms the core of the 26S proteasome which is part of the non-lysosomal degradation pathway. Inhibitor studies indicated that the protease may be an unusual type of serine or cysteine protease. Mutagenesis of the cysteine residue, the two histidine residues and the conserved aspartate residues has shown this not to be the case (Seemüller *et al.*, 1995). The structure of the β subunit showed that the proteasome had the activity of a threonine protease (Wlodawer, 1995) and has 14 active sites. The catalytic mechanism was analysed by site-directed mutagenesis and inhibitor studies. Deletion of the N-terminal threonine resulted in inactivation of the protease as did substitution with an alanine. Substitution with a serine residue resulted in a fully active enzyme

(Seemüller*et al.*, 1995). A crystal structure of the enzyme complexed to an inhibitor that showed that proteolytic attack is mediated by the N-terminal threonine of processed β subunits.

Further studies indicate that autocatalytic processing takes place in the proteasome and the action is intermolecular (Seemüller*et al.*, 1996)

1.4.2 Viral proteases

Viral enzymes tend to be smaller than the polypeptides of their cellular counterparts. They have a lower turnover number and an unusually high degree of specificity.

Table 1.3. Examples of viral proteases

Virus	Genetic material	Name/protein	Protease type
Adeno-	dsDNA	Adeno 22K	cysteine
Alpha-	+ss RNA	Sindbis nsP2	cysteine
Picorna-	+ssRNA	Polio 2A and 3C	cysteine
	+ssRNA	Rhino 2A and 3C	cysteine
Flavi-	+ssRNA	Hep C NS3	serine
Retro-	+ssRNA	HIV-1 PR	aspartic
	+ssRNA	RSV PR	aspartic
Caulimo-	dsDNA	Cauliflower mosaic	aspartic?
Hepadna-	dsDNA	Hepatitis B	aspartic?

1.4.2.1 Retroviral proteases

Retroviral proteases may be encoded in different regions of the genome. In Rous Sarcoma virus (RSV) and myeloblastosis associated virus (MAV) the protease is encoded wholly in the *gag* gene but in HIV-1 it is encoded by the *pol* gene. The catalytic site of the retroviral protease is homologous to that of the aspartic protease (Toh *et al.*, 1985)

with the residues Asp, Thr (Ser) and Gly forming the site. These must be preceded by two hydrophobic residues. Another region 60 residues downstream from the catalytic site, with a Ile-Ile-Gly motif, has also been shown to have an effect on catalytic activity (Pearl and Taylor, 1987). The protease has been shown to be a dimer with a fold similar to that of the aspartic acid proteases. The connecting loops in the β-sheet that forms the core of the protease are shorter so the molecules are smaller than cellular aspartate proteases (Pearl and Taylor, 1987). In HIV the protease is composed of two chains of 99 residues in length. The active site is between the two domains. The two aspartate residues in the two domains are connected by a network of hydrogen bonds. Substrates bind with the C=O bond to be cleaved in between catalytic carboxyl groups. The oxygen group is close to the position of the displaced water molecule which normally bridges the gap.

Retroviral proteases are highly substrate specific but there is no consistent pattern to their cleavage sites. The residues that form the cleavage sites may determine the order of the cleavage cascade as a variety of bonds must be cleaved by the same enzyme.

1.4.2.2 Adenoviral cysteine proteases

Proteolysis in adenoviruses occurs in a limited fashion. The six virus structural proteins are proteolytically cleaved to facilitate the production of mature virus particles. The L3 23K protein has been identified as the protease responsible for the cleavages by mapping of mutations in a temperature sensitive mutant which was defective in protease activity at certain temperatures (Yeh-Kai *et al.*, 1983). It has been shown to be a cysteine protease but not of the conventional papain-like type. It has also been shown that the protease is activated by a disulphide linked peptide (Webster*et al.*, 1993).

1.4.2.3 Flaviviral serine proteases

Flaviviruses include the flavi-, pesti- and hepatitis C viruses. For a review of the proteinases of the *Flaviviridae* see (Ryan *et al.*, 1998) and references therein. All three classes have been shown to have virus encoded trypsin-like serine proteases. This was

first shown by alignment and then by experimentation. Hepatitis C virus (HCV) has two identified virus-encoded proteases. One of these has been characterised as a zinc-containing serine protease and the other is a trypsin-like serine protease, $NS3^{pro}$. The crystal structure of the NS3 protease has been determined and this reveals a trypsin-like fold and a structural zinc binding site (Love*et al.*, 1996). NS3 had previously been found to have the catalytic domain of a trypsin-like serine protease (Miller and Purcell, 1990) . The protease is 180 amino acids in length and the catalytic triad of His-57, Asp-75 and Ser-135 is strictly conserved in all HCV sequences. Two six-stranded β- barrels were also identified which make up the chymotrypsin fold. Novel features include a long N-terminus that interacts with neighbouring molecules and a structural zinc binding site. The protease activity is encoded by the N-terminal one third of the protein and the remaining C-terminal two thirds encode a helicase activity.

NS3 is multifunctional and forms a heterodimer with NS4A. It cleaves at the NS3/NS4A junction *in cis* and at the 5A/B junction *in trans*. NS3 requires an accessory viral protein and a structural zinc ion to function. The zinc binding site is on the opposite side of the protein to the active site and it is composed of residues Cys-97, Cys-99, Cys-145 and His-149. The zinc is thought to maintain the structural stability of the enzyme and guide the folding of the NS3 domain. The sequence of the metal binding site of NS3 is more conserved than that of the catalytic site.

NS3 and NS4 form a stable complex. It anchors NS3 to membranes and stabilises NS3 *in vitro*. The N-terminal 28 amino acids of NS3 are required for formation of a stable complex.

1.4.2.4 Flaviviral metalloproteases ?

Some experiments have indicated that HCV NS2/3 protease may be a zinc dependent enzyme which performs a single cleavage to release the N-terminus of NS3 REF. The zinc residue bound to NS3 has no role in the serine protease activity of the enzyme. The zinc residue is bound by three cysteine residues and an activated water as

in proteins which act as catalysts. This releases the structural protein core and the envelope glycoproteins.

1.5. Polyprotein processing in picornaviruses

For a review of the virus-encoded proteinases of the picornaviruses see (Ryan and Flint, 1997) and references therein.

1.5.1 L protease

The L protease is found at the N-terminus of the aphtho- and Equine rhinoviral (ERV 1 and 2) polyproteins. It cleaves co-translationally at its own C-terminus (Strebel and Beck, 1986). Two forms of the protease, Lab^{pro} and Lb^{pro}, are found due to initiation at either of two AUG codons 84 nucleotides apart (Clarke *et al.*, 1985; Sangar*et al.*, 1987). Lb^{pro} (and possibly Lab^{pro}) undergoes posttranslational modification by a carboxypeptidase B-like activity which gives rise to Lb' or Lab' (Sangar *et al.*, 1988). Lab^{pro} and Lb^{pro} cleave at the L/P1 junction *in cis* or *in trans.* (Devaney *et al.*, 1988).

The L proteinase has been shown to have similarities to thiol proteases (Schneemann *et al.*, 1994). The active site was shown to be composed of Cys^{51} and His^{148} (Piccone *et al.*, 1995; Roberts and Belsham, 1995). Modelling the sequence onto the 3-dimensional structure of papain showed that Glu^{71} and Asp^{174} were also part of the catalytic site. Cys^{51} is shown to be conserved in FMDV and ERV1 and 2. His^{148} is not conserved in ERV2. The only candidate histidine (His^{75}) is misaligned by 9 residues.

The L proteinase also has a role in cleaving eIF-4G *in trans* , the same role fulfilled by $2A^{pro}$ in entero- and rhinoviruses. However L^{pro} is a thiol protease and $2A^{pro}$ is thought to belong to the ' small ' sub-group of serine proteases. The cleavage site differs between L^{pro}, which cleaves at the junction between Gly^{479} and Arg^{480}, and $2A^{pro}$ which cleaves between Arg^{486} and Gly^{487} (Kirchweger *et al.*, 1994).

1.5.2 2A protease

1.5.2.1 $2A^{pro}$ activity in entero and rhinoviruses

In entero- and rhinoviruses $2A^{pro}$ mediates the primary cleavage between P1 and P2P3, this occurs at a tyrosine-glycine (Y-G) pair in poliovirus (Hanecak *et al.*, 1984). The 2A proteinase, which has a molecular weight of 17kDa, cleaves at its own N-terminus (Sommergruber *et al.*, 1989; Toyoda *et al.*, 1986). It may also have a role in cleaving the P3 precursor to give 3C' and 3D' at a tyrosine-glycine pair in 3D, as demonstrated in poliovirus (Hanecak *et al.*, 1982; McLean *et al.*, 1976). This cleavage is strain specific and the significance of the cleavage is not clear. The Y-G pair is conserved in all entero- and rhinoviruses sequenced to date but in aphtho- and cardioviruses the $2A^{pro}$ sequence and function is different (see below). Mutation in the sequence of $2A^{pro}$ in polioviruses does not affect growth of the virus in tissue culture and produces a parental phenotype.

1.5.2.2 Classification of the 2A protease

The 2A protease has been found to be a zinc containing enzyme (Sommergruber *et al.*, 1994) and the zinc ion to play a structural role in the enzyme (Voss *et al.*, 1995). The catalytic triad was thought to consist of His^{20}, Asp^{38} and Cys^{109}, with Cys^{109} as the active site nucleophile (Bazan and Fletterick, 1988). Inhibitor studies have shown $2A^{pro}$ to be sensitive to the thiol protease inhibitors iodoacetamide and N-ethylmaleamide (Chen *et al.*, 1996). The active site and other properties were confirmed by site-directed mutant analysis (Dinakarpandian *et al.*, 1997; James, 1994; Lumry, 1997; Sommergruber *et al.*, 1989). It has preferred requirements for the residues near to the cleavage site but no absolute requirements.

1.5.2.3 Other functions of 2A

The 2A proteinase also inhibits host-cell protein synthesis *in trans* (Markland *et al.*, 1997). This correlates with inhibition mediated by cleavage of the 220kDa polypeptide associated with the cap binding protein (Olson *et al.*, 1994). Cleavage of

eIF-4G directed by 2A (Sommergruber *et al.*, 1994) inhibits cap-dependent translation by the host cell but does not affect cap-independent translation. This results in host-cell ' shut-off ' and selective viral RNA translation.

The 2A proteinase is the transactivator of translation by the poliovirus internal ribosome entry site (IRES) when host-cell (cap-dependent) translation is not inhibited (Qiu *et al.*, 1997).

It was found that mutations in the 5' non-coding region of viral RNA could be compensated by second site mutations in the $2A^{pro}$ region (Macadam *et al.*, 1994). Point mutations in $2A^{pro}$ show loss of cleavage *in trans* but not *in cis* and RNA replication is inhibited (Yu and Lloyd, 1991).

1.5.2.4 2A protease activity in aphtho- and cardioviruses

The 2A protease in aphtho- and cardioviruses cleaves at its own C-terminus between $2A^{pro}$ and 2B, releasing LP12A in cardioviruses and P1-2A in aphthoviruses. Precursors to 2A/2B are not detected *in vitro*. The 2A protease shows no known protease motifs and no similarities with the $2A^{pro}$ region of entero-and rhinoviruses but is comparable in size. The 2A protease region is conserved among the TME and EMC viruses. Only the C-terminus of the region is conserved in cardioviruses but it is similar to the 2A region in FMDV. The final three residues at the C-terminus (NPG) are conserved in cardio-and aphthovirus 2As and the N-terminus of the 2B region is completely conserved.

The cleavage at the 2AB junction in FMDV and TMEV does not require L^{pro} or $3C^{pro}$. The 18 residues of $2A^{pro}$ and the 2 residues of 2B mediate co-translational cleavage (Ryan and Drew, 1994).

1.5.3 3C protease

The 3C protease in picornaviruses has been shown to have homology with the 'large' serine proteases such as trypsin and chymotrypsin. The protease is inhibited by chymostatin indicating its similarity to chymotrypsin (Orr *et al.*, 1989). It has a catalytic triad composed of Cys^{147}, His^{40} and Glu^{71}. The active site nucleophile is a cysteine residue instead of the serine normally found in this type of protease. The active site nucleophile is conserved in all rhinoviruses. An alignment of the $3C^{pro}$ amino acid sequences from two polio serotypes (PV1 Mahoney and PV1 Sabin) and HRV14 shows that there is a great deal of similarity between the two proteases (Figure 1.19).

```
                                                                    34
PV1-MAH.3C      ----------GPGFD---YAVA-MAKRNIVTATTSKGEFTM--LGVHDNV
PV1-SAB.3C      ----------GPGFD---YAVA-MAKRNIVTATTSKGEFTM--LGVHDNV
HRV-14..3C      ----------GPNTE---FALS-LLRKNIMTITTSKGEFTG--LGIHDRV

                                                                    68
PV1-MAH.3C      AILPTHA---------------SPGESIVIDGKEVEILD-AKALEDQAGT
PV1-SAB.3C      AILPTHA---------------SPGESIVIDGKEVEILD-AKALEDQAGT
HRV-14..3C      CVIPTHA---------------QPGDDVLVNGQKIRVKD-KYKLVDPENI

                                                                   102
PV1-MAH.3C      NLEITIITLKRN--EKFRDIRPHIPTQITETNDGVLIVNTSK--YPNMYV
PV1-SAB.3C      NLEITIITLKRN--EKFRDIRPHIPTQITETNDGVLIVNTSK--YPNMYV
HRV-14..3C      NLELTVLTLDRN--EKFRDIRGFISEDLEG-VDATLVVHSNN--FTNTIL

                                                                   136
PV1-MAH.3C      PVGAVTEQGYLNLGGRQTARTL----MYNFPTRAGQCGGVITCTG-----
PV1-SAB.3C      PVGAVTEQGYLNLGGRQTARTL----MYNFPTRAGQCGGVITCTG-----
HRV-14..3C      EVGPVTMAGLINLSSTPTNRMI----RYDYATKTGQCGGVLCATG-----

                                                                   170
PV1-MAH.3C      PVGAVTEQGYLNLGGRQTARTL----MYNFPTRAGQCGGVITCTG-----
PV1-SAB.3C      KVIGMHVGGNGSHGFAAALKRSYFTQS----------Q
HRV-14..3C      KIFGIHVGGNGRQGFSAQLKKQYFVEK----------Q
```

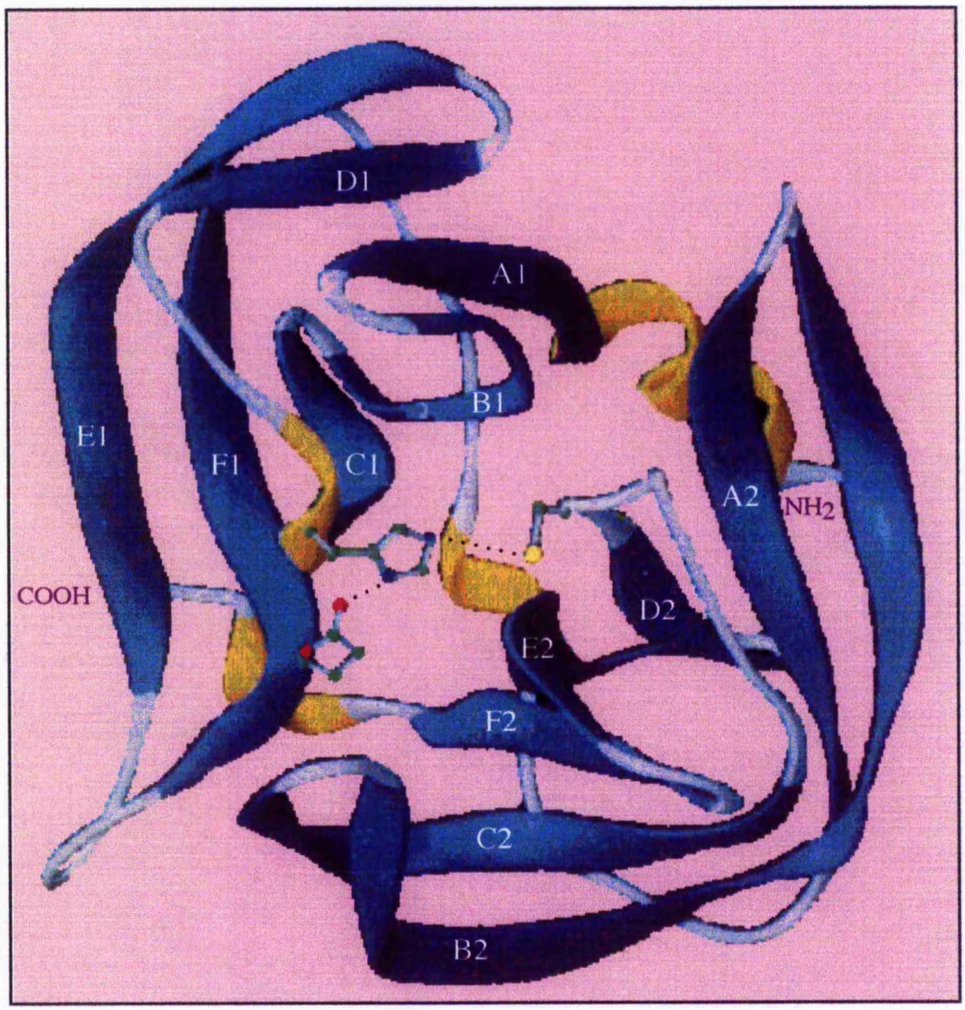
Figure 1.19 Alignment of HRV14 $3C^{pro}$ and polio $3C^{pro}$ sequences. The residues forming the catalytic triad are shown in red. These are conserved in the polio virus and rhinovirus sequences

The three dimensional structure of HRV14 $3C^{pro}$ (Figure 1.20) has been shown to consist of two topologically equivalent 6 stranded β-barrels with a long shallow groove

for binding between the two (Matthews *et al.*, 1994). Secondary cleavages are carried out by the viral protease $3C^{pro}$.

In HRV 14 $3C^{pro}$ was originally thought to cleave only at Q-G pairs. It has now been shown that cleavage occurs mostly at (E,Q), (G,S,A) pairs, however, there are exceptions. Specificity appears to depend on the surrounding sequences (Palmenberg, 1990).

Figure 1.20 Structure of HRV14 $3C^{pro}$ Matthews *et al..*, (1994).
The two six-stranded ß -barrels can be seen in blue numbered A1-F1 and A2-F2. The helical parts of the structure are shown in yellow. The residues of the catalytic triad are also marked on the diagram; His-40 (blue), Glu-71 (red) and Cys-146 (yellow).



1.5.3.1 Cleavage by $3C^{pro}$ *in trans*

$3C^{pro}$ has also been shown to cleave cellular sequences. Poliovirus $3C^{pro}$ has been shown to cleave active transcription factor IIIC (TFIIIC) to its inactive form (Clark*et al.*, 1991). It proteolyses TFIIIC *in vivo* and *in vitro* to give complex III. If a point mutation is inserted in $3C^{pro}$ then no complex formation is observed. TFIIIC complexes I and II are cleaved to give complex III. This accounts for the inhibition of transcription in poliovirus infected cells. $3CD^{pro}$ is known to enter the nucleus after infection and $3C^{pro}$ is small enough, at 20 kDa, to enter the nucleus by diffusion. FMDV $3C^{pro}$ does not have the same effect on TFIIC and the same effect is not seen in FMDV-infected cells.

Inhibition of host cell RNA polymerase II (Pol II) by poliovirus is correlated to a decrease in activity of a fraction containing transcription factor TFIID. Analysis of a component of TFIID, the TATA-binding protein (TBP), shows that direct cleavage by $3C^{pro}$ occurs *in vitro* and *in vivo* (Clark *et al.*, 1993). As cleavage of TBP occurs at the same time as Pol II transcription is inhibited $3C^{pro}$ is, therefore, seen to have a role in host cell Pol II-mediated transcription shutoff.

Another cellular sequence which is cleaved by both poliovirus and HRV14 $3C^{pro}$ is microtubule-associated protein 4 (MAP-4; Joachims *et al.*, 1995). No cleavage is seen with the other viral protease from poliovirus, $2A^{pro}$. Cleavage with $3C^{pro}$ results in cleavage products which are identical to those seen *in vivo*. No cleavage is seen with a mutant $3C^{pro}$. Cleavage also occurs with $3CD^{pro}$ and has been correlated with microtubule collapse.

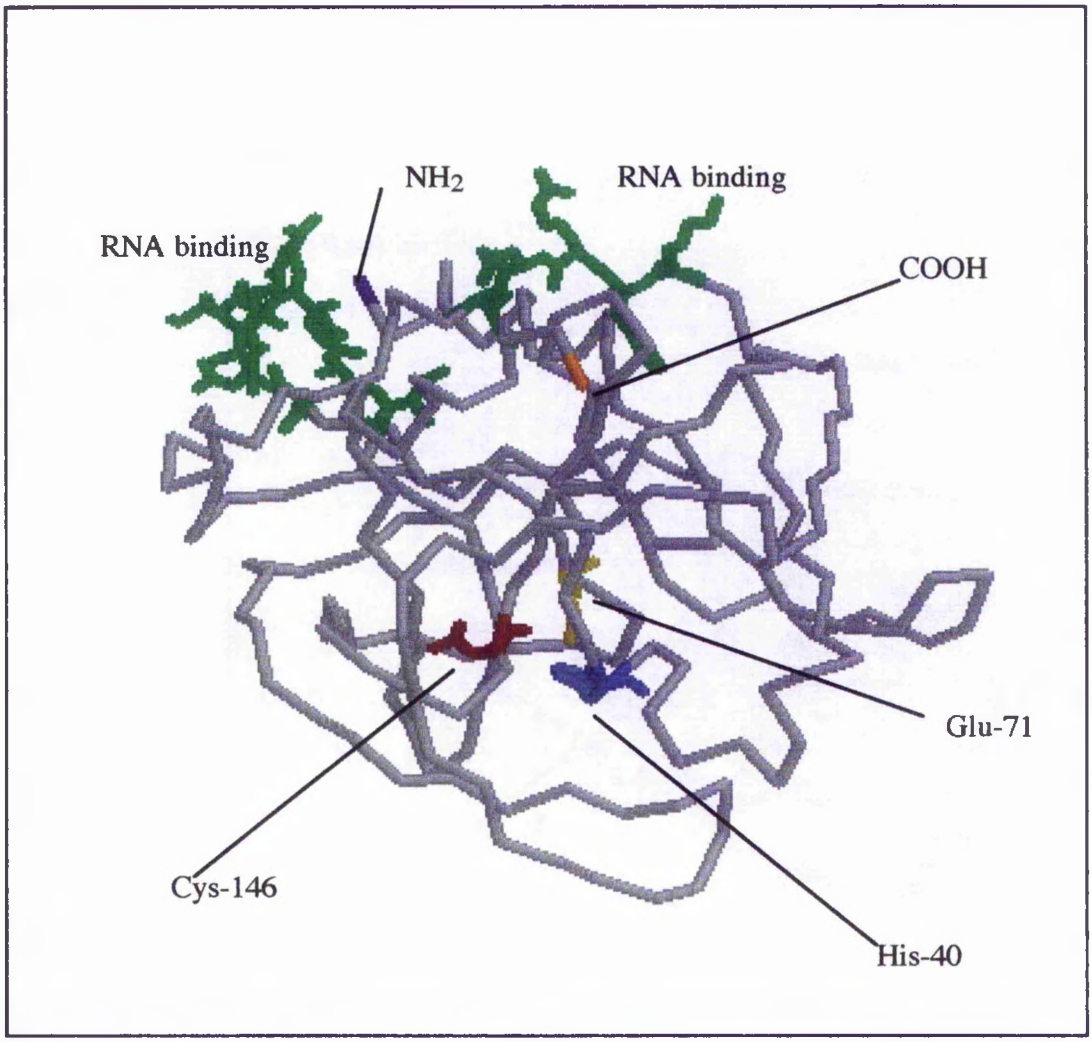
1.5.3.2 RNA binding by $3C^{pro}$

Mutations in poliovirus $3C^{pro}$ have been shown to suppress mutations in the 5' end of the genome (Andino *et al.*, 1990). A four base insertion in loop d of the 5' NCR rendered the virus deficient in RNA synthesis but revertants showed a partial recovery of wild type synthesis. Analysis of these showed point mutations in the $3C^{pro}$ region of the genome which were *cis*-active and implied that $3C^{pro}$ interacted with the 5' end of the genome.

Further studies showed that the cloverleaf structure at the 5' end of the poliovirus genome was able to bind $3C^{pro}$ and $3D^{pol}$ (Andino *et al.*, 1990). This forms part of a ribonucleoprotein (RNP) complex which is crucial to the virus life cycle. The 5' most 90 nucleotides of the genome were shown to fold into a cloverleaf structure bound by $3C^{pro}$ and $3D^{pol}$. This interaction was detected by immunological methods (Andino *et al.*, 1990). A cellular component was also found in the RNP complex, p36 (Andino *et al.*, 1993). $3C^{pro}$ was found to bind RNA on its own but with much lower affinity than $3CD^{pro}$. Mutants in $3C^{pro}$ were identified which affected the formation of the RNP complex. They were found to localise to a region on the opposite face of the protein from those which affect proteolytic processing.

Experiments by Leong and co-workers (1993) show that rhinovirus 14 $3C^{pro}$ binds to the 5' 126 nucleotides of the genome. Point mutations in $3C^{pro}$ indicated amino acids which are important for this non-proteolytic function of $3C^{pro}$ (Figure 1.21). Conservative amino acid substitutions at Asp^{85} destroyed the ability of the protein to bind RNA showing that this is required for specific binding to the 5' 126 nucleotides of the genomic RNA. Amino acid substitutions at the residues of the catalytic triad (His^{40}, Glu^{71} and Cys^{146}) resulted in proteolytically inactive mutants which were, however, able to bind RNA. It was also shown that substitution of Asp^{85} with a serine residue results in a loss of P1-3CD interaction (Hämmerle *et al.*, 1992). This may be due to distortion of the structure of the loop or entire structure of $3CD^{pro}$. The size of the residue at position 85 is probably an important factor in this interaction.

Figure 1.21 Residues involved in RNA binding in HRV14. Modelled on the structure of HAV 3C protease. The residues shown in green are those which bind RNA. The residues of the catalytic triad of HRV14 are also shown: His-40 (pale blue), Glu-71 (red) and Cys-146 (yellow). The N-terminus is shown in purple and the C-terminus in orange.



1.5.4 3CD

In an alternative cleavage pathway in P3 the cleavage between $3C^{pro}$ and $3D^{pol}$ does not take place. The resulting polyprotein $3CD^{pro}$ has protease activity but not the polymerase activity found in $3D^{pol}$. The lack of detectable polymerase activity in poliovirus $3CD^{pro}$ was demonstrated by analysis of a bacterially expressed protein (Harris *et al.*, 1992). The 3C/D cleavage site had been mutated to prevent proteolysis. The resulting protein was, however, capable of cleaving P1 and peptides mimicking the 2B/C cleavage site. It is responsible for the cleavage of P1 into the capsid proteins, an activity which can be carried out by $3C^{pro}$ but not as efficiently (Jore *et al.*, 1988; Ypma-Wong *et al.*, 1988). It has been shown that 3AB interacts with $3CD^{pro}$ in this role to perform an RNA binding function at the 5' end of the genome (Xiang *et al.*, 1995). It has been implicated in RNA replication by genetic analysis (Harris *et al.*, 1994). The 3CD protease also stimulates the cleavage of 3AB to 3A and VPg (Lama *et al.*, 1994). 3AB also stimulates cleavage of $3CD^{pro}$ to $3C^{pro}$ and $3D^{pol}$ (Molla *et al.*, 1994).

In an alternative cleavage pathway, mediated by $2A^{pro}$, $3CD^{pro}$ is cleaved, at a Tyr-Gly pair, to produce 3C' and 3D'. These proteins are not required for replication (Lee and Wimmer, 1988). The cleavage site can be mutated and cleavage will still occur (Tyr-Gly to Phe-Gly). Mutations in the surrounding sequence affect the cleavage with threonine in the P2 position of the cleavage site being particularly important.

1.5.4.1 Alternative roles of 3D : polymerase and protease.

In vivo in poliovirus the 3C/D cleavage is seen to be less efficient than most other Q-G cleavages. The 3C protease and $3D^{pol}$ are present in much smaller molar quantities than other viral processing products (Dewalt and Semler, 1987). This may possibly maintain a pool of $3CD^{pro}$ which is thought to be responsible for all the Q-G cleavages in the viral genome and especially those in P1. It has been shown that purified recombinant 3AB interacts with $3CD^{pro}$ and accelerates processing. This yields the products $3C^{pro}$ and $3D^{pol}$ (see above). Initiation of replication by $3D^{pol}$ has

been shown to be both template and primer dependent poly-U synthesis. VPg is uridylated by $3D^{pol}$ to VPgpU(pU), the primer for $3D^{pol}$ (Paul *et al.*, 1998). 3AB also inhibits the processing of P1 by $3CD^{pro}$.

The structure of poliovirus $3D^{pol}$ has been determined by X-ray crystallography (Hansen*et al.*, 1997) and it is the first RNA dependent RNA polymerase to have its structure determined in this way. It has the same shape as other polymerases with differences in the structure of the subdomains. The core sequence of the polymerase is similar to that of other polymerases and has four of the sequence motifs that have been described for RNA-dependent polymerases.

1.5.5 Maturation cleavage

The final cleavage of picornaviral processing is observed *in vivo* during the final stages of virion morphogenesis. The cleavage occurs concomitantly with RNA association with the large capsid assembly structures (12S-14S pentameric assembly unit; Jacobson and Baltimore, 1968). The mechanism has not yet been determined. One possible mechanism is similar to that of the viral proteases, but involving bases of the encapsidated RNA. This RNA may act in place of the histidine and accept protons. This would then catalyse the VP0 cleavage (Rossmann *et al.*, 1985). Conservation of the serine residue at the carboxyl end of VP4 and the amino end of VP2 suggests that the serine is the nucleophile for the cleavage of VP0 to VP4 and VP2. The reaction also requires a proton abstracting base to activate the nucleophile and this role may be filled by an RNA base or by polyamines (Altman, 1984; Fout *et al.*, 1984).

1.5.6 Other products of P3

3AB has been shown to have multiple functions in the replication of poliovirus. It was shown to have a dual function as both the precursor for VPg and a co-factor for $3D^{pol}$ (Lama *et al.*, 1994). The soluble form of 3AB was found to stimulate poly(U) RNA synthesis catalysed by $3D^{pol}$. Amino acid changes in the hydrophobic region of 3A result

in polio mutants with an impaired RNA replication phenotype (Giachetti and Semler, 1991).

The action of 3AB *in vivo* may be stimulated by binding to the $3D^{pol}$ part of $3CD^{pro}$ which is bound to the initiation complex at the 5' end of the poliovirus genomic RNA (Colonno *et al.*, 1988; Harris *et al.*, 1994) . It enhances the binding of $3D^{pol}$ to RNA and this is supported by experiments showing the binding of $3CD^{pro}$ and $3D^{pol}$ to the pseudoknot at the 3' end of the genomic RNA which requires 3AB.

When in the initiation complex at the 5' end of the genome $3CD^{pro}$ may be able to catalyse the uridylation of 3AB or VPg. Therefore the primer for transcription by $3D^{pol}$ will be generated either by autocleavage of $3CD^{pro}$ or may be present *in trans*.

1.5.7 Proteolytic activities in picornaviruses (*cis* and *trans*)

Processing occurs both *in cis* and *in trans* in the polyprotein processing cascade of picornaviruses. In the early stages of RNA replication processing is carried out by proteases acting *in cis* as this is more kinetically favourable when there are few RNA molecules to be processed. In the later stages of replication there are more protease molecules and more polyproteins to be processed so *trans* processing also occurs. Reactions performed *in trans* are, by their nature, slower than those *in cis* and also more sensitive to the sequences surrounding their cleavage sites.

Polyprotein processing eliminates the need for genetic regulation of RNA replication. Utilisation of both *cis* and *trans* processing activities allows an ordered cascade and the production of mature products when they are required. Reactions *in trans* cannot take place until the protease has been released from the polyprotein, meaning that they occur later in the processing cascade. Primary cleavages of the polyprotein occur *in cis* as they are performed rapidly and whilst translation of the polyprotein is still occurring. The structural proteins are processed *in trans* by $3CD^{pro}$ later in the replication cycles, they are not required for capsid formation until later in the replication cycle

1.6 Mutation of $3C^{pro}$

Work has been done on the effects of substitution mutations on the residues of the catalytic triad of $3C^{pro}$ in poliovirus (Kean *et al.*, 1993). Substitutions at the Glu^{71} and Cys^{147} sites of the poliovirus $3C^{pro}$ did not abolish cleavage altogether *in vitro*. A hierarchy of activity was found, with the double mutation G71Y, C147S resulting in a complete lack of processing by $3C^{pro}$. Cleavage was affected more easily at some sites than others. In infectivity and RNA synthesis assays after transfection all $3C^{pro}$ mutants gave negative results. This indicates that mutation of these two residues almost abolishes viral replication.

Studies have also been carried out which result in "overcleaving" at the C-terminus of $3C^{pro}$ (Kean *et al.*, 1988). Replacement of the isoleucine residue at position 74 of poliovirus $3C^{pro}$ with a threonine results in impaired cleavage at the -COOH terminus. The low level of $3D^{pol}$ indicated that the processing was impaired rather than the stability of $3C^{pro}$. Insertion of a serine residue at position -3 relative to $3C^{pro}$ also impaired cleavage at the -COOH terminus (Semler *et al.*, 1987).

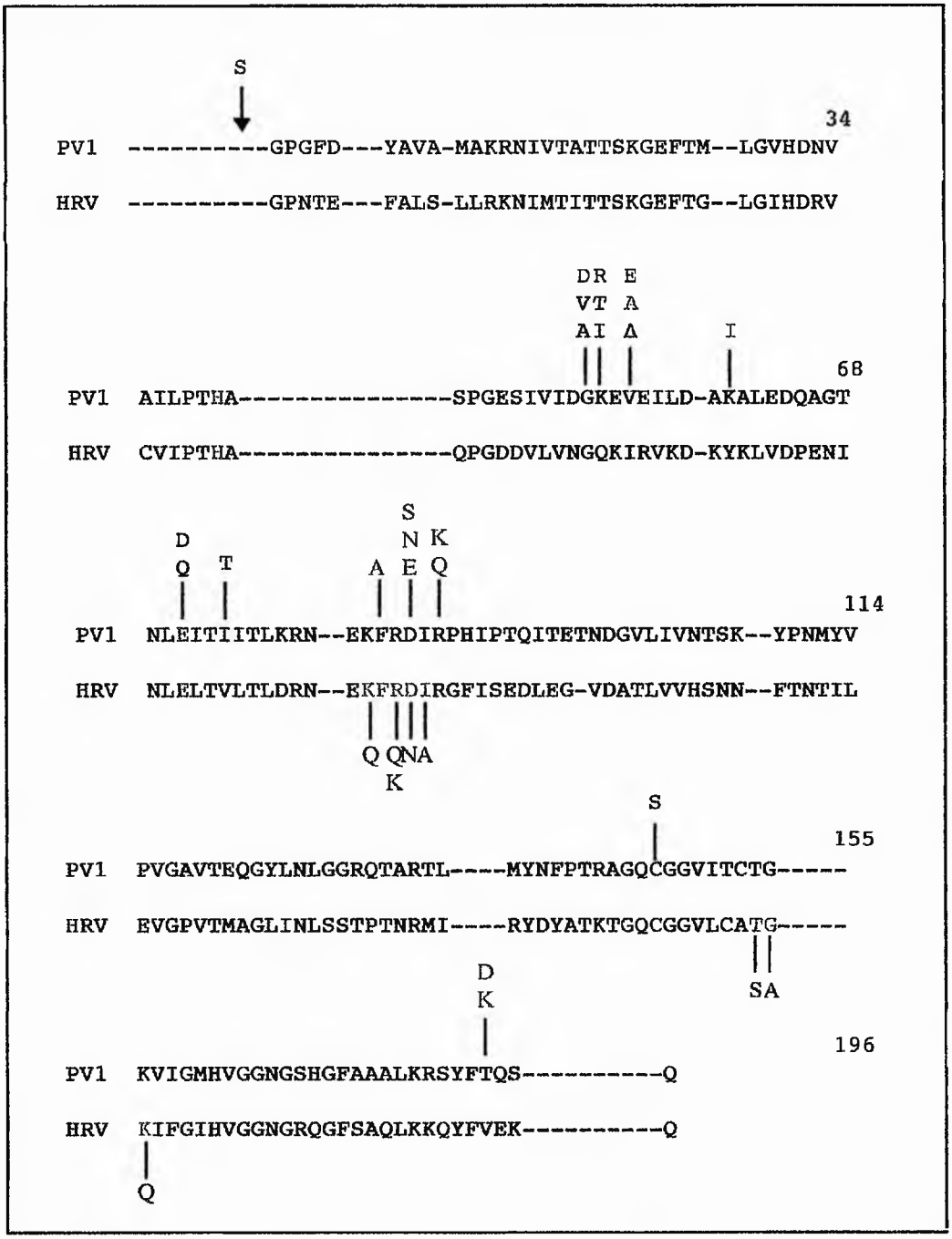
The substitution of the valine residue at position 54 in the polio $3C^{pro}$ sequence for an alanine residue resulted in defective processing at the 3B/C junction (Dewalt and Semler, 1987). Low levels of $3C^{pro}$ and $3D^{pol}$ were seen but normal levels of 3C' and 3D' were found, indicating that processing by $2A^{pro}$ was not affected. This mutation was also characterised by secondary effects on the replication of the mutant virus (Dewalt *et al.*, 1989).

It was also found that substitution of the lysine residue at position 60 in wild type poliovirus with an isoleucine residue resulted in overproduction of $3D^{pol}$. The increase in processing at the carboxy terminus was three to five times that of the wild type virus (Dewalt *et al.*, 1990).

A summary of residues which have been mutated to study their role in cleavage activity can be seen in Figure 1.22.

Figure 1.22 Alignment of HRV14 and poliovirus 3C protease sequences showing residues which have been mutated to elucidate their role in cleavage activity. Residues shown above the wild type sequence aremutant forms. The arrow indicates that an amino acid was inserted at this position in the sequence. Numbers at ends of lines are residue numbers for the poliovirus sequence.

I = overproduction of $3D^{pol}$
S,T = Cleavage at COOH impaired
Q, K, N, A, S = No RNA binding activity
H, E, C = residues of catalytic triad
K, F, R, D, I, T, G, K = residues involved in RNA binding



1.7 Regulation of cleavage of 3CD

The regulation of cleavage by $3C^{pro}$ in HRV 14 has yet to be determined. There are several possible viral components which act with $3CD^{pro}$ or $3C^{pro}$ which may be responsible for determining the cleavage pathway. The virus requires both $3C^{pro}$ and $3CD^{pro}$ in its replication cycle so there must be some mechanism for determining which protein is produced. 3AB is known to act with $3CD^{pro}$ in the formation of a ribonucleoprotein complex at the 5' end of the genome but is also known to catalyse cleavage of $3CD^{pro}$ to $3C^{pro}$ and $3D^{pol}$. The 5' end of the genome binds to $3CD^{pro}$ so this may also have a role in determining the cleavage. The 3C protease has been determined to act as an RNA binding protein which may mean that RNA has a role in determining the cleavage site.

The mutations mentioned above (section 1.6) which cause disruption to the cleavage sites may also indicate residues which have a role in directing cleavage of $3C^{pro}$.

1.8 Aims of this project

In vivo Human Rhinovirus 14 $3C^{pro}$ acts in two different forms. Cleavage can either generate $3C^{pro}$ (which is the viral protease) and $3D^{pol}$ (the viral polymerase) or $3CD^{pro}$ which is the protease. The virus requires all of these activities for efficient replication. The method by which the virus switches from one to the other has yet to be determined. This project was therefore designed to try and determine the regulatory switch.

A reporter gene system was designed using antibiotic resistance genes to flank the viral sequence of interest. Several 3C-containing genes were cloned between the genes to study the action of $3C^{pro}$ away from other viral components and with other parts of P3. Other viral components and RNA were then added *in trans* to see whether they had an effect on processing by 3C in the reporter gene system. Cleavage in the reporter system was studied *in cis* and *in trans*. Mutations in the nucleotide sequence of $3C^{pro}$ in poliovirus have been reported which affect cleavage at the N- and C-termini of the protease (see section 1.6). Mutations in HRV14 $3C^{pro}$ will also be studied to see if they have an effect on cleavage in the reporter gene system

2. Experimental

2.1 Materials

T4 DNA ligase, T4 DNA polymerase, *Taq* polymerase, 1kb DNA ladder, restriction enzymes and their buffer solutions were obtained from Promega. Protein size markers were obtained from Sigma and BDH. Oligonucleotides were synthesised using phosphoramidite chemistry on an Applied Biosystems 381A by I. Armitt or obtained from Oswel Ltd. DNA sequencing was carried out by Alex Houston on a Perkin-Elmer ABI Prism™ 377 DNA sequencer using multicolor ABI dRhodamine 'BigDye' terminators. Transformations, mini-preparations and maxi-preparations were carried out using the *E. coli* strains JM109 or SURE.

Solutions:

TE	10 mM Tris.HCl 1 mM EDTA (pH 8.0)
TAE	0.04 M Tris-acetate 0.001 M EDTA
TBE:	0.9 M Tris-borate, 0.002 M EDTA.
Agarose gel loading buffer	2 x TAE 50 % [v/v] glycerol 0.005 % [w/v] bromophenol blue 0.004 % [v/v] ethidium bromide
fmol Sequencing buffer (5 x):	250 mM Tris.HCl (pH 9.0), 10 mM $MgCl_2$.
Destain:	20 % [v/v] methanol, 10 % [v/v] glacial acetic acid.
Tris glycine buffer:	0.1 % [w/v] SDS, 25 mM Tris, 250 mM glycine.
2 x SDS-PAGE loading buffer: (containing β-mercaptoethanol)	124 mM Tris.HCl (pH 6.8), 4 % SDS, 0.2 % bromophenol blue, 20 % glycerol, 10 % β-mercaptoethanol.

LB	1 % [w/v] bacto tryptone 0.5 % [w/v] bacto-yeast extract 1 % [w/v] NaCl pH 7.0.
SOB	2 % [w/v] bacto-tryptone, 0.5 % [w/v] bacto-yeast extract 0.05 % [w/v] NaCl 0.25 M KCl 10 mM $MgCl_2$
Phenol/chloroform	50 % phenol 50 % chloroform overlaid with 0.01 M Tris.Cl
TFB	10 mM MES (2-[*N*-morpholino]ethansulfonic acid) 45 mM $MnCl_2$.4H2O 10 mM $CaCl_2$.2H2O 100 mM KCl
KMES	1 M MES pH to 6.3 with 5 M potassium hydroxide

Solutions for minipreparation of DNA by alkaline lysis

Solution 1	50 mM glucose 25 mM Tris.Cl (pH 8.0) 10 mM EDTA
Solution 2	1 % SDS 0.2 M NaOH
Solution 3	3 M KAc 5 M acetic acid

Solutions for immunoprecipitation

NET/GEL	0.15 M NaCl 0.05 M Tris.HCl (pH 7.4) 0.005 M EDTA 0.02 % NaN_3 0.05 % NP-40 0.25 % gelatin
NET/GEL/BSA	NET/GEL with 0.2 % BSA
NET/GEL/NaCl	NET/GEL with a final concentration of 0.5 M NaCl.
NET/GEL/SDS	NET/GEL with a final concentration of 0.1 % SDS.

2.2 Methods

2.2.1 Cloning techniques

Restriction enzyme digestions. All plasmid DNA restriction enzyme digestions were carried out under the conditions specified by the supplier. In general, 1 μg of DNA was digested with 1 unit of enzyme in a total volume of 20 μl, containing 2 μl of 10X restriction buffer at 37 °C unless otherwise specified.

Agarose-gel electrophoresis. Flat bed agarose gels, of concentration 0.7 - 2 % w/v, were prepared with 1 x TAE which contained ethidium bromide at a final concentration of 0.5 μg/ml. DNA samples were applied to the gel in agarose gel loading buffer. Electrophoresis was carried out at 100 - 180 V in 1 x TAE containing 0.5 μg/ ml ethidium bromide. Following electrophoresis the DNA bands were visualised by illumination from a UV transilluminator (UVP).

Agarose Gel Purification of DNA fragments. DNA was isolated from low melting point agarose using one of the two following methods:

(i) Phenol/acetate method. The DNA band was run into 1 % low melting point agarose (LMP) from where it was excised and placed in a 1.5 ml microcentrifuge tube. This was heated to 70 °C to melt the agarose, then filled with phenol/acetate (phenol equilibrated with 0.3 M sodium acetate), vortexed and incubated on ice for 10 minutes before being spun in a microfuge at 13000 rpm for 5 minutes . The top aqueous layer was removed to a fresh tube. The extraction was repeated and 400 μl of the top layer removed to a fresh tube. To precipitate the DNA, 1 ml of ethanol and 20 μl of 2 M sodium acetate were added and mixed. The tube was kept at -20 °C overnight, spun at 14000 rpm at 4 °C in a microcentrifuge for 20 minutes and the supernatant discarded. The DNA was resuspended in sterile distilled water.

(ii) Wizard method. DNA was excised from LMP agarose as above and isolated using the Wizard Prep DNA Purification system (Promega), according to the manufacturer's instructions.

Ligations. Ligation reactions were routinely carried out in a final volume of 20 µl. The ligation reaction mix generally consisted of 1 µl (4 Weiss units) of T4 DNA ligase , 0.5 µg of vector DNA, and insert concentrations at 2 fold, 5 fold or 10 fold molar ratios in T4 DNA ligase buffer (30 mM Tris.HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP). The reaction mixes were incubated overnight at 16 °C .

Transformation of *E. coli.* Competent *E. coli* were prepared and transformed with plasmid DNA by one of the following methods:

(i) Calcium chloride method. For transformation of ligation reactions, where a medium efficiency of transformation was required the calcium chloride method was used. 10 ml of LB were inoculated with 1 ml from an overnight culture of LB inoculated with *E. coli* , and incubated on a shaker until exponential phase (OD_{600} = 0.4) was reached, then the cells were cooled to 0 °C for 10 minutes in centrifuge tubes, pelleted at 4 °C and the supernatant discarded. Each pellet was resuspended in 10 ml of ice cold 0.1 M calcium chloride and stored on ice for 15 minutes. The cells were pelleted, the supernatant discarded and then each pellet resuspended in 1 ml ice cold 0.1 M calcium chloride and stored on ice for 15 minutes (or until required). The transformation efficiency of cells prepared by this method is at an optimum between 12 and 24 hours after preparation. To 200 µl of cells 0.5 µg of DNA, or a ligation reaction, is added and incubated on ice for 30 minutes. The cells were then heat shocked by heating to 42 °C for 90 seconds and rapidly transferring them to an ice bath for 2 minutes, before plating out on LB agar plates containing the appropriate antibiotics.

(iii)Modified Hanahan method. $MgCl_2$ (5 mM) and $MgSO_4$ (5 mM) were added to 25 ml SOB and this was inoculated with 250 µl bacterial culture (SURE or

JM109). This was incubated on a shaker at 37 °C overnight. SOB (25 ml) plus 5 mM $MgCl_2$ and 5 mM $MgSO_4$ was inoculated with 250 µl of the overnight culture and incubated at 37 °C until in log phase. The cells were cooled to 0 °C for 10 minutes and then pelleted at 4 °C and the supernatant discarded. The cells were resuspended in 2 ml cold TFB and incubated on ice for 30 minutes. Cells were pelleted at 4 °C and the supernatant discarded. Cells were resuspended in 2 ml cold TFB and incubated on ice for 5 minutes. Dimethylformamide (DMF; 70 µl) was added and incubated on ice for 5 minutes. A mixture of β-mercaptoethanol and KMES (70 µl) was added, mixed by swirling and incubated on ice for 10 minutes. DMF was added (70 µl), the cells were mixed by swirling and incubated on ice for 5 minutes. The DNA to be transformed was put into pre-cooled tubes and 200 µl of cells were added. The cells were incubated on ice for 30 minutes . The cells were heat shocked at 42 °C for 2 minutes. The whole mix was then plated onto LB agar plates containing appropriate antibiotics.

DNA Preparations. Mini-, midi-, and maxi-preparations of plasmid DNA were used depending on the quantity and quality of DNA required. Midi- or maxi-preparations were required for transcription and translation.

(a) Mini-preparation of plasmid DNA. Minipreparation of DNA was carried out *via* 2 different methods; the latter being used for preparation of DNA for automated sequencing.

(i)Alkaline lysis method. LB broth (10 ml) containing ampicillin (150 µg/ml final concentration) was inoculated with a single colony and incubated in an orbital incubator overnight. The cells were pelleted, the supernatant removed, and the cells resuspended in 200 µl of solution 1. Solution 2 was added (400 µl) and mixed, followed by 300 µl of solution 3. The sample was mixed and incubated on ice for 20 minutes, the solution was centrifuged and the supernatant transferred to a fresh eppendorf. The sample was then extracted by adding an equal volume of phenol/chloroform, vortexing, centrifuging and removing the upper aqueous layer

to a fresh eppendorf. Two volumes of cold ethanol and 1/20 volumes of 2 M sodium acetate were then added, mixed, and the sample incubated at -20 °C for 1 hour. The sample was centrifuged and the DNA precipitate was washed with 70 % (v/v) ethanol, dried, and then resuspended in 50 µl of distilled water with 2 µl RNaseA (1 mg/ml).

(ii)Promega Wizard SV DNA miniprep kit. Minipreparation of DNA for automated sequencing was carried out according to the manufacturers instructions.

(b) Maxi-preparation and Midi-preparation of plasmid DNA. Maxi-preparation of plasmid DNA was carried out *via* two different methods; the latter of these two methods was also used for midi-preparation of plasmid DNA:

(i) Caesium chloride gradient method. LB broth (200 ml) containing 100 µg/ml ampicillin was inoculated with a single colony and incubated in an orbital incubator overnight. The cells were pelleted and the supernatant decanted. The cells were then resuspended in 10 ml of solution 1. Solution 2 was added (20 ml) and mixed. Solution 3 was added (15 ml) and mixed. The mixture was then kept at 0 °C for 5 minutes, centrifuged for 15 minutes and the supernatant poured through muslin into fresh Sorval tubes. Isopropanol (0.6 volumes) was added, left at room temperature for 30 minutes and then centrifuged for 5 minutes. The supernatant was pipetted off and the nucleic acid resuspended in 8.5 ml of TE. 10 g of caesium chloride and 1 ml of 5 mg/ml ethidium bromide were added to the DNA solution and the sample was centrifuged for 20 hours at 55Krpm at 20 °C. The plasmid DNA bands were visualised under UV light and the lower band, containing closed circular supercoiled DNA, was removed. The gradient fraction was extracted with 1 ml of caesium chloride saturated isopropanol, the sample vortexed and the upper phase removed. This was repeated until all ethidium bromide was removed. Distilled water (2 ml), 300 µl of 2 M sodium acetate and 7.5 ml of ethanol were then added and the sample incubated at room temperature for 30 minutes to precipitate the DNA. This was then pelleted by centrifugation for 10 minutes, washed with 70 % ethanol and resuspended in 150 µl of TE.

(ii) Qiagen method. Maxi-preparations and midi-preparations of plasmid DNA were carried out according to the Qiagen Plasmid Maxi or Midi Kit protocol, as per manufacturer's instructions (Qiagen).

Polymerase chain reaction (PCR). The PCR amplification of DNA was used to amplify genes for cloning purposes and for the production of gene fragments with site directed mutation within the HRV14 $3C^{pro}$ region. A typical PCR reaction was carried out using 20 ng of template DNA with 200 pmol of each primer, 0.25 mmol of each dNTP, 10 µl of *Taq* DNA polymerase buffer (50 mM KCl, 10 mM Tris.HCl (pH 9.0), 0.1 % (v/v) Triton X-100, 1.5 mM $MgCl_2$) and 2.5mM $MgCl_2$ in a total volume of 100 µl. The reactions were overlaid with mineral oil to prevent evaporation of the reaction mixture during thermal cycling. Reactions were heated to 94 °C for 5 minutes then held at 85 °C whilst 2 units of *Taq* DNA polymerase (Promega) were added to the aqueous phase of each reaction. Amplification was carried out on a thermal cycler using the following parameters: 94 °C for 2 minutes, to denature the DNA; 50 °C for 1 minute, to allow primers to anneal to the DNA template; 72 °C for 1 minute for every thousand base pairs to be amplified to allow *Taq* DNA polymerase to extend from each primer. The amplification was carried out for 25 or 30 cycles, with the final 72 °C step being increased to 5 minutes, to ensure that the majority of final product was full length double stranded DNA. The annealing temperature was varied according to the base composition and, therefore, annealing temperature of the primer. Oligonucleotide primer sequences can be seen in Tables 2.1a and b.

Site directed mutagenesis by overlap PCR. PCR was carried out as above and products of the primary reactions were purified using Wizard PCR Preps (Promega). The purified products were then used as templates in the secondary reactions at a concentration of

Table 2.1a Oligonucleotide primers for construction of control plasmids.

OB4; 30mer. HRV14 3C Forward Primer.	M CGCGCGGACGTCATGCGCCCAGTTGTTGTGCAA AatII
OB6; 30mer. HRV14 3C Reverse Primer.	* GCGCGCCTGCAGTCATCTAGCTATTACTTGGCC PstI
OB21; 30 mer HRV14 3D reverse primer	* TTTTTTCTGCAGTCACTAAAAGAGGTCCAACCA PstI
OB23; 30 mer. HRV 14 3B reverse primer	* TTTTTTCTGCAGTCATTGCACAACGGCTGGGCG PstI
OB30; 30 mer. HRV 14 3D forward primer	M TTTTTTGACGTCATGTATTTTGTAGAGAAA AatII
OB31; 30 mer HRV 14 3A forward primer	M TTTTTTGACGTCATGCTAGAAACACTGTTT AatII
Z1554; 33mer HRV 14 5'NCR forward primer	CGCGCGGGATCCTTAAAACAGCGGATGGGTATC BamHI
Z1555; 33mer HRV14 5'NCR reverse primer	CGCGCGCTGCAGCCATGATCACAGTATATGTAT PstI
KAN 10; 31 mer. KAN forward primer	M CCCGGGGGGCCCATGAGCCATATTCAACGGG ApaI
OB8; 30mer. KAN Reverse Primer.	* CGCGCGGACGTCTCAGAAAAACTCATCGAGCAT AatII
OB9; 42mer. TET forward Primer.	M CGCGCGCTGCAGATGAAATCTAACAATGCG PstI
OB10; 30mer. TET Reverse Primer.	* CGCGCGGAGCTCTCAGGTCGAGGTGGCCCGGCT SacI
OB14 ; 36 mer 3C forward primer	AAATTTGGTACCAAGGACCAAACACAGAATTTG KpnI
OB15 ; 36 mer 3C reverse primer	CGCGCGAAGCTTCTATTGTTTCTCTACAAAATA HindIII
OB16; 33mer. 3D reverse primer.	CGCGCGAAGCTTCTAAAAGAGGTCCAACCAGCG HindIII
OB12; 35mer. P1 Forward.	GGGGGTACCGCCGCCACCATGGGCGCTCAGGTT KpnI
OB13; 33mer. P1 Reverse.	CCACTGTAATTTAGGATAATCTGATCACCCCCC SpeI

50ng and the conditions and DNA concentrations were as above. Oligonucleotide primer sequences for the primary reaction may be seen in Table 2.2.

TABLE 2.1b Oligonucleotide primers for the construction of artificial reporter plasmids.

KAN 10; 31 mer. KAN forward primer	M CCCGGGGGGCCCATGAGCCATATTCAACGGG ApaI
OB8.1; 30mer. KAN Reverse Primer.	CGCGCGGACGTCGAAAAACTCATCGAGCAT AatII
OB4.1; 30mer. HRV14 3C Forward Primer.	CGCGCGGACGTCCGCCCAGTTGTTGTGCAA AatII
OB6.1; 30mer. HRV14 3C Reverse Primer.	GCGCGCCTGCAGTCTAGCTATTACTTGGCC PstI
OB9.1; 42mer. TET forward Primer.	CGCGCGCTGCAGAAATCTAACAATGCG PstI
OB10.; 30mer. TET Reverse Primer.	* CGCGCGGAGCTCTCAGGTCGAGGTGGCCCGGCT SacI
OB31.1; 30 mer HRV 14 3A forward primer	TTTTTTGACGTCGGACCAGTGTATAAA AatII
OB21.1; 30 mer HRV14 3D reverse primer	TTTTTTCTGCAGCTAAAAGAGGTCCAACCA PstI

Table 2.2 Oligonucleotide primers for overlap PCR

OB19; 21mer. HRV14 3C Cys^{146}->Ala Forward	ACTGGGCAGGCTGGAGGTGTG
OB20.2; 21mer. HRV14 3C Cys^{146}->Ala Reverse	CACACCTCCTGCCTGCCCAGT
OB4; 30mer. HRV14 3C Forward Primer.	CGCGCGGACGTCCGCCCAGTTGTTGTGCAA AatII
OB6; 30mer. HRV14 3C Reverse Primer.	GCGCGCCTGCAGTCTAGCTATTACTTGGCC PstI

Purification of PCR products. PCR products were purified using the Wizard PCR preps (Promega). Either the direct purification method was used or products were run into 1 % LMP agarose and purified from there.

2.2.2 Nucleotide dideoxy sequencing of recombinant DNA clones

Cycle sequencing with [^{32}P]-end-labelled primer. Nucleotide dideoxy sequencing was carried out *via* a method based on the fmol protocol (Promega), using cycle sequencing with a [^{32}P]-end-labelled primer. Plasmids pEB1, pEB2 and pEB3 were sequenced using this method.

Automated sequencing. Automated sequencing was carried out by Alex Houston on a Perkin Elmer ABI Prism™ DNA sequencer. Sequencing primers were designed to the T7 and SP6 polymerase promoters and a region within the kanamycin resistance gene (Table 2.3). Data was viewed using EditView software from Applied Biosystems. Plasmid DNA (500 ng)was submitted for sequencing with 3-5 pmol of the appropriate oligonucleotide primer.

T7 promoter primer 20 mer	5'-d(TAATACGACTCACTATAGGG)-3'
SP6 promoter primer 19 mer	5'-d(TATTTAGGTGACACTATAG)-3'
Kan sequencing primer 18 mer	5'-d(ATCGCAGACCGATACCAG)-3'

Table 2.3 Nucleotide sequences of oligonucleotide primers used for automated sequencing.

2.2.3 Translation *in vitro*

Coupled transcription/translation (TnT) reactions. Proteins were expressed *in vitro* using coupled transcription/translation kits, for either wheatgerm extract (WGE) or rabbit reticulocyte lysate (RRL) systems, according to the manufacturers instructions (Promega).

Proteins were radiolabelled with [^{35}S]-methionine unless otherwise stated. Reactions were incubated at 30 °C for 90 minutes, unless otherwise stated, then stopped by the addition of SDS-PAGE loading buffer. Aliquots of translation reactions (3 µl) were analysed by denaturing PAGE, or 10 µl aliquots were immunoprecipitated.

Processing *in trans*. To monitor processing *in trans* in the reporter polyprotein system TnT reactions were carried out as detailed above. The rabbit reticulocyte lysate (RRL) system was used for these reactions. Reactions were incubated at 30 °C for 90 minutes. The reactions were stopped by the addition of RNaseA (to a final concentration of 0.5µg/ml) and cycloheximide (to a final concentration of 0.8µg/ml), overlaid with mineral oil and incubated overnight at 30 °C. The results were analysed by SDS-PAGE.

To monitor processing *in trans* by 2A^{pro}, 3CD^{pro} and 3CD' the protease source was synthesised in the TnT wheatgerm extract system without radiolabelling. A mixture of amino acids minus leucine (1 mM) was used to provide a source of methionine in the unlabelled reactions. Labelled substrate was synthesised as detailed above. Both reactions were stopped by the addition of RNaseA and cycloheximide. Unlabelled protease (5µl) and 5 µl of the labelled substrate were then mixed together, overlaid with mineral oil and incubated together overnight at 30 °C.

The *trans* processing reaction was also carried out with 7.5 µl of the labelled substrate and 5 µl of the unlabelled protease source to try and create conditions which were more favourable for *trans* processing.

Co-translation. Equal amounts of two plasmid DNAs were used to program a TnT reaction, one acting as a substrate and the other as a protease source. They were co-expressed in the TnT system for 90 minutes at 30 °C. The reaction was then stopped by the addition of RNaseA and cycloheximide, overlaid with mineral oil and incubated overnight at 30 °C.

Addition of DTT. DTT was added to the *trans* processing reactions to enhance the efficiency of *trans* processing. It was added when the substrate and protease were mixed, before the overnight incubations, to a final concentration of 2 mM.

2.2.4 Protein expression in *E. coli*

Luria Broth with 150 μg/ml of ampicillin was inoculated with *E. coli* strain BL21(DE3) and grown overnight on an orbital incubator at 37 °C. Overnight bacterial culture (1 ml) was added to 9 ml of LB with 150 μg/ml of ampicillin and grown , with shaking at 37 °C for 3 hours. The culture was divided into 2 x 5 ml and to one of these IPTG was added to a final concentration of 1 mM, to induce expression of the fusion protein. Incubation was continued for a further 3 hours. The cells were then pelleted by spinning for 5 minutes at 3000 rpm, the supernatant discarded, and the cells resuspended in 250 ml of ice-cold PBS. The cells were then lysed using a probe sonicator for 3 x 10 seconds with 10 seconds on ice between sonicating. From each sample of whole sonicate, induced and uninduced, 100 ml was removed and spun in a microfuge. The supernatant was removed and the pellet resuspended in 100 ml. For both the induced and uninduced samples, 10 ml of supernatant and insoluble material were analysed by SDS-PAGE.

2.2.5 Expression of antibiotic resistance in *E.coli*

The genes coding for the artificial polyprotein [KanR3C^{pro}Tet^{R}] were released from pEB4 by restriction digest with ApaI and SacI and ligated into pBSIIKS+. This was to allow expression of antibiotic resistance by the reporter genes. Control plasmids were also made. The ligation reactions were transformed into *E.coli* and plated onto LB agar plates. Transformants were screened by restriction digest and those found to be correct were used for a large scale plasmid preparation.

The plasmid DNA was then transformed and plated onto LB agar plates containing combinations of antibiotics (see Table 2.4). When expression of the tetracycline

resistance gene product was being monitored the cultures were incubated at 37 °C for 1 hour before plating to allow expression of antibiotic resistance by Tet^R. Tet^R encodes a 399 amino acid membrane associated protein which prevents tetracycline from entering the cell. The bacteria need time to accumulate this protein product before they show resistance to tetracycline, hence the 1 hour incubation. This was not necessary with ampicillin and kanamycin as these work by different mechanisms. Ampicillin inhibits enzymes in the bacterial membrane that are involved in cell wall synthesis and kanamycin binds to ribosomal components and inhibits protein synthesis. The resistance gene products act on these antibiotics work by breaking down the drugs and therefore do not need to accumulate before they can be active.

	Ampicillin 100µg/ml	Kanamycin 10µg/ml	Tetracycline 25µg/ml
1	√	X	X
2	√	√	X
3	√	X	√
4	√	√	√
5	X	√	√

Table 2.4 Antibiotic components of LB agar plates used to monitor expression of antibiotic resistance.

2.2.6 Immunoprecipitation reactions

The products of *in vitro* transcription/translation reactions were immunoprecipitated with an antibody to neomycin phosphotransferase (NPTII) (5Prime->3Prime). Neomycin is an analogue of kanamycin and the antibody should detect the translation products of the kanamycin resistance gene.

Preparation of immunoprecipitin. Immunoprecipitin (Gibco-BRL), i.e heat-killed, formalin-fixed *Staphylococcus aureus* cells, 1 g in 10 ml PBS, was centrifuged for 20 minutes at 3000 rpm and the supernatant discarded. The cell pellet was resuspended in

PBS containing 10 % β-mercaptoethanol and 3 % SDS, and the sample incubated at 95 °C for 30 minutes. The cells were centrifuged at 3000 rpm for 20 minutes, the supernatant discarded and the cells resuspended in 900µl of NET/GEL/BSA.

Pre-clearing the sample. This removed proteins which would have bound non-specifically directly (*i.e.* not *via* the antibody) to the immunoprecipitin. 50µl of NET/BSA was added to 10 µl of coupled TnT reaction in a 0.5 ml microcentrifuge tube. 1.5 µl of immunoprecipitin solution was added and incubated at room temperature for 15 minutes, then spun in a microcentrifuge at 13000 rpm for 2 minutes and the pellet discarded.

Immunoprecipitation reaction. 5 µl of the anti-NPTII antibody was added to 50 µl of the supernatant of the pre-cleared sample and incubated overnight at 4 °C. 6 µl of immunoprecipitin were added and incubated for 15 minutes at room temperature, spun in a microcentrifuge for 2 minutes at 13000 rpm, and the supernatant discarded. The pellet was resuspended in 100 µl of NET/GEL/NaCl, spun in a microfuge for 2 minutes at 13000 rpm, and the supernatant discarded. The pellet was suspended in 100 µl of NET/GEL/SDS, spun in a microfuge for 2 minutes at 13000 rpm, and the supernatant discarded. The pellet was resuspended in 100 µl of 10 mM Tris.HCl (pH 7.5) and 0.1 % NP-40, spun in a microfuge for 2 minutes at 13000 rpm, and the supernatant discarded. The pellet was resuspended in 1x SDS-PAGE loading buffer containing 100 mM DTT, and the sample boiled for 2 minutes and spun in a microfuge for 2 minutes at 13000 rpm. The results were analysed by SDS-PAGE.

2.2.7 Protein analysis

Denaturing Polyacrylamide gel electrophoresis (SDS-PAGE). The discontinuous buffer gel system based on that of Laemmli (1970) was used for denaturing polyacrylamide gel electrophoresis (SDS-PAGE). Unless otherwise noted all gels were constructed with a 4 % polyacrylamide stacking gel and a 10 % polyacrylamide resolving gel. Typically electrophoresis of each gel was carried out at a constant current of 15 - 20 mA throughout the stacking gel and 30 - 40 mA through the resolving gel.

3. Cloning and molecular characterisation of the components of a 'Reporter' polyprotein system.

In order to observe the cleavage activity of $3C^{pro}$ without the influence of other viral sequences a reporter system was constructed. HRV14 $3C^{pro}$ was cloned into a transcription vector, with two flanking reporter genes to create a single ORF. Two antibiotic resistance genes were chosen as the non-viral components of the reporter system to investigate the possibility of selection of cleavage products by antibiotic resistance in *E.coli*, in addition to the identification of translation products in cell-free translation systems. The antibiotic resistance genes were ligated into a transcription vector flanking the viral sequence of interest. The system was also designed to allow the substitution of components by digestion with the appropriate restriction enzymes at sites which flank each gene.

The predicted molecular weights of the components of the reporter polyprotein system were calculated (Figure 3). This was to ensure that they were of significantly different sizes to allow reasonable separation on an SDS-PAGE gel. Similarly the molecular weights of the predicted primary cleavage products were calculated. It can be seen from Figure 3 that the products all have at least a 10 kDa difference in size which is sufficient to separate them on a 10% SDS-PAGE gel. Once it had been established that the size differences were substantial enough to allow identification of products on an SDS-PAGE gel plasmid construction was started.

The values shown in Figure 3 are for the construct encoding the artificial polyprotein [$Kan^R3C^{pro}Tet^R$]. Other artificial polyproteins were also to be constructed but they would utilise the same reporter genes and have 3C-containing sequences as the viral component. Therefore, molecular weights of products would be almost the same as in this construct with a few additional viral protein cleavage products.

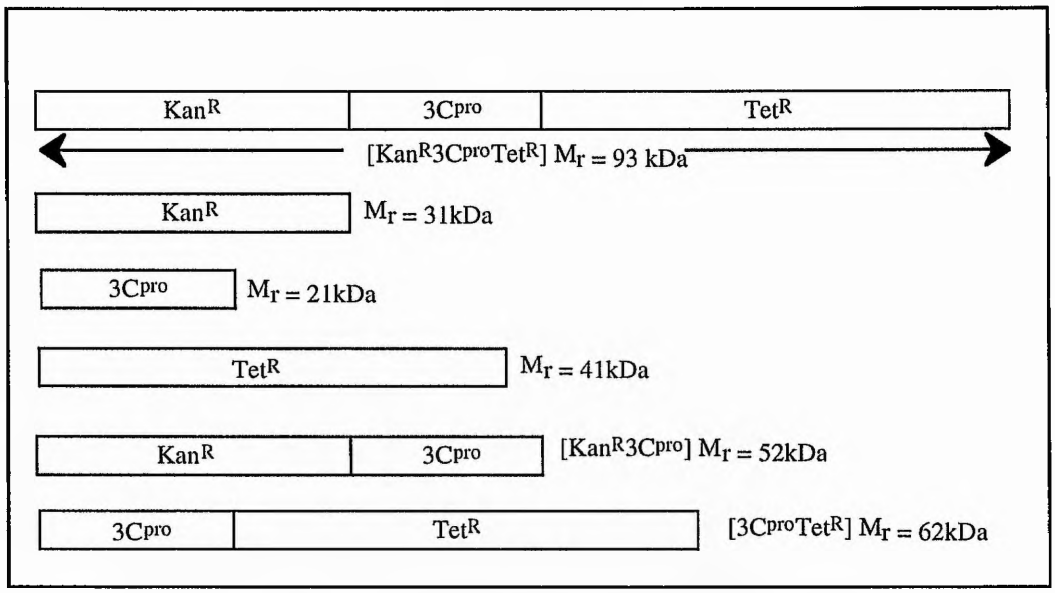


Figure 3 Predicted molecular weights of the components of the reporter polyprotein system [Kan^{R}3C^{pro}Tet^R] and the possible cleavage products.

3.1 Addition of restriction enzyme sites to gene sequences by the PCR

In order to facilitate ligation of the artificial polyprotein components into a transcription vector restriction enzyme cleavage sites were added to the ends of the PCR products. Oligonucleotide primers were designed to the 5' and 3' sequences of each gene. The restriction enzyme site nucleotide sequence was incorporated upstream of the start codon in the forward (5') primer. It was followed by the initiation codon (ATG) in primers for the control plasmids and then by codons from the 5' end of the gene which would anneal to the template (Figure 3.1).

The reverse primer was designed in a similar way. The appropriate restriction site was used but the sequence was reversed and complemented. A stop codon (TGA) followed the restriction site sequence in primers for the genes for the control plasmids. The reverse complement of the six codons from the 3' end of the gene made up the remainder of the primer sequence. Primer sequences can be seen in tables 2.1 a and b and 2.2 in section 2.

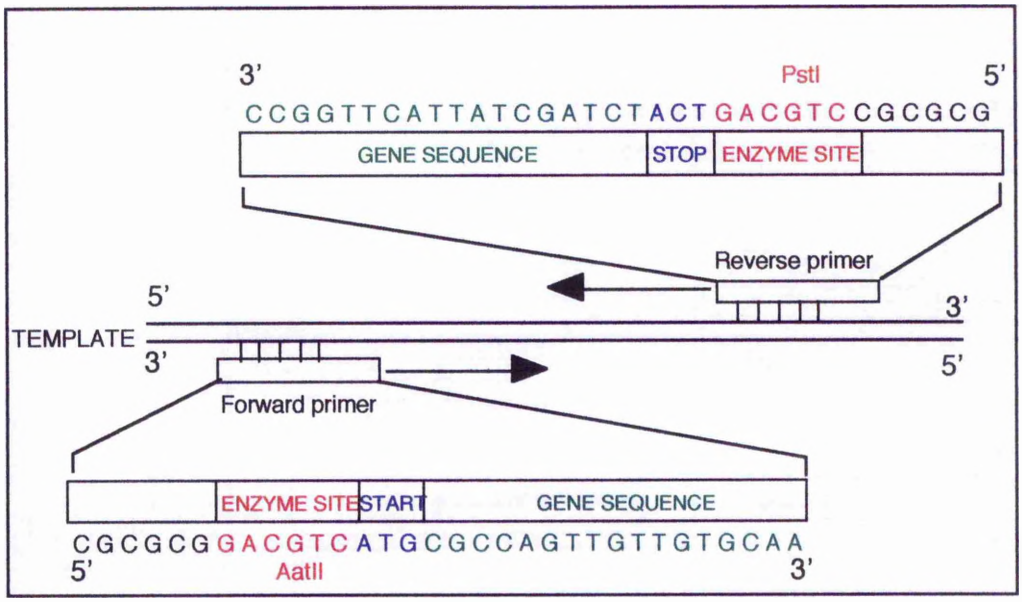


Figure 3.1 Addition of restriction sites to genes by the PCR. The oligonucleotide primers shown are for the amplification of $3C^{pro}$ from the full length HRV14 cDNA.

3.2 Construction of control plasmid pEB1.

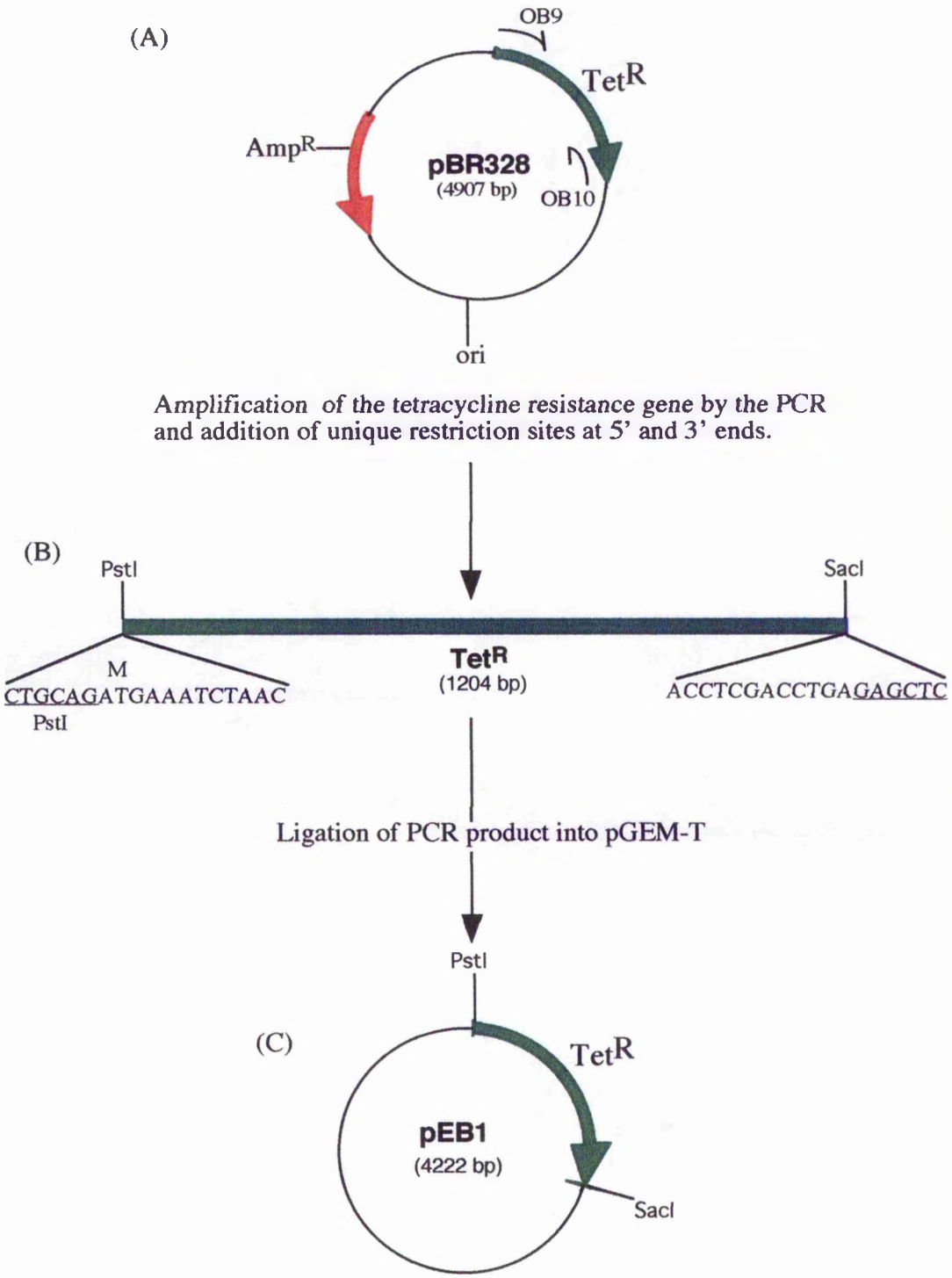
The tetracycline resistance gene (Tet^R) was amplified by the PCR for use as a reporter gene in the system. Oligonucleotide primers were designed to the 5' and 3' ends of the gene (see Table 2.1). Restriction enzyme sites were incorporated into the sequence of the primers so that they would be added to the ends of the PCR products. These were to allow ligation into the transcription vector pGEM5zf(+).

The plasmid pEB1 was constructed as a control for the size of the translation product of the tetracycline resistance gene which would be one of the reporter genes in the artificial polyprotein system (Figure 3.2). Tet^R has its own initiation codon which will allow transcription in the TnT system. This will not be strictly necessary in the reporter system as the polyprotein will be translated as a polypeptide from the initiation codon at the 5' end of the Kan^R gene so different primers were used to amplify the sequence for inclusion in the reporter system.

In order to construct a control plasmid for the translation of the Tet^R gene the gene was amplified using the PCR from the plasmid pBR328. Unique PstI and SacI restriction sites were added to the 5' and 3' ends of the gene respectively using the PCR to facilitate ligation into pGEM5zf(+) which had been similarly restricted. For convenience the PCR product was cloned into pGEM-T (a vector which utilises the adenine added to the ends of amplified products by the PCR) for easy cloning. Small scale DNA preparation (miniprep) was carried out on putative clones and restriction digests were performed to identify clones where the insert had been ligated into the vector. The clone was digested with PstI and SacI and a 1200 bp fragment was released corresponding to the Tet^R gene. The plasmid was sequenced using a T7 promoter primer on an ABI Prism™ automated sequencer. This gave the correct nucleotide sequence and the construct was used to program a wheatgerm extract coupled transcription translation system.

On translation the plasmid did not give a product of the expected size and any translation products which were produced were at very low levels. The plasmid was re-sequenced using T7 and SP6 promoter primers. The sequences from both ends of the Tet^R gene were the same leading to the conclusion that ligation into the pGEM-T vector

Figure 3.2 Construction of pEB1. The Tet^R coding sequences were amplified by the PCR using oligos OB9.1 and OB10.1 (A). The purified PCR product (B) was ligated into the vector pGEM-T to form plasmid pEB1 (C).



had allowed two copies of the PCR product to be cloned into the vector end-to-end. On transcription both positive and negative strand RNA was produced, which then paired and prevented any translation. Therefore no translation product was seen in the TnT reaction (Figure 3.2.1).

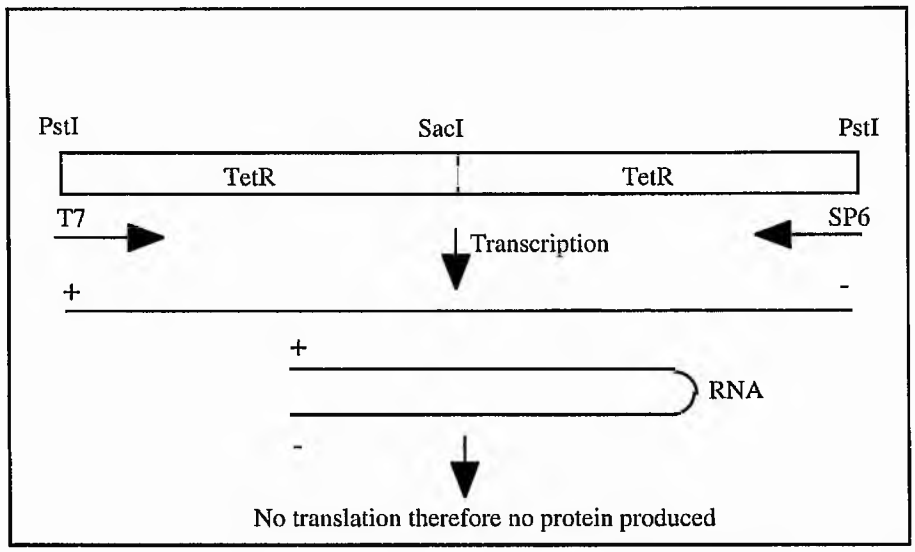
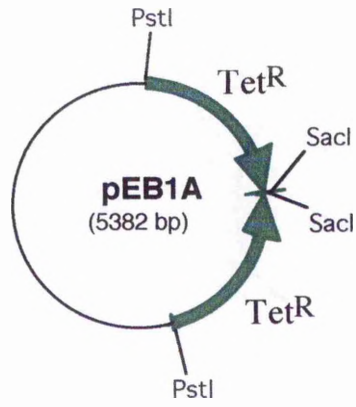


Figure 3.2.1 Product of transcription and translation of pEB1 [Tet^{R}pGEM-T].

In order to produce a control plasmid which would translate effectively, the tetracycline resistance gene PCR product was excised from the pGEM-T vector using the unique restriction sites added at either end and ligated into pGEM5zf(+). pGEM5zf(+) had been restricted with the same enzymes, PstI and SacI (Figure 3.2.2). This plasmid was then used to program a WGE TnT reaction and produced a translation product of the predicted size (Figure 3.2.3).

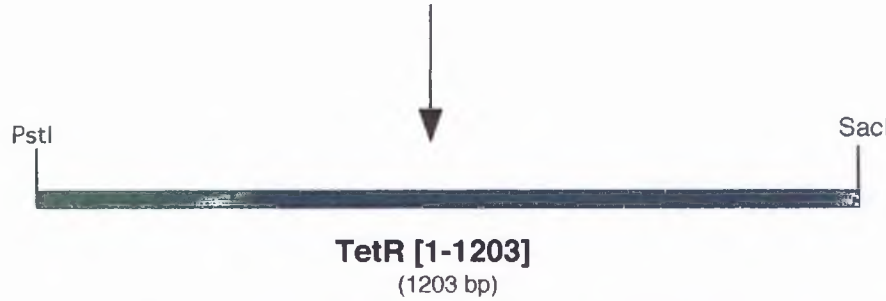
Figure 3.2.2 Construction of pEB1.1. The Tet^R coding sequence was released from the plasmid pEB1A (A) by restriction digest with PstI and SacI (B). The fragment was then ligated into pGEM5zf(+) which had been similarly digested to form the plasmid pEB1.1 (C).

(A)



Tet^R gene released from vector by digestion with PstI and SacI.

(B)



Tet^R gene ligated into pGEM5zf(+) restricted with PstI and SacI.

(C)

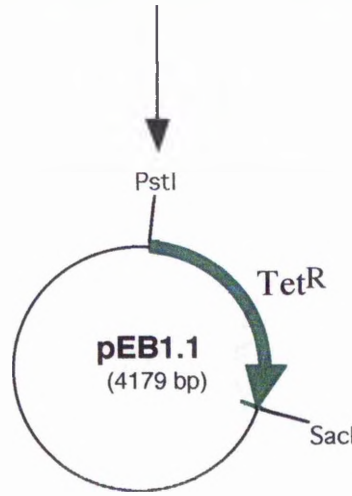
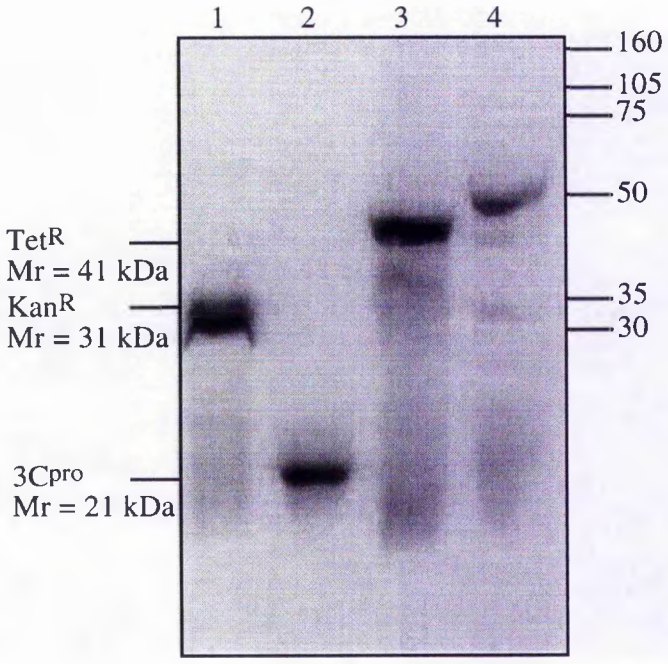


Figure 3.2.3 Translation product of pEB1.1. Two putative clones (Lanes 3 &4) were translated to see whether they produced translation products of the correct size for Tet^R (41 kDa). Using SDS-PAGE to analyse the products of the TnT reaction the clone in Lane 3 was found to produce the correct product and named pEB1.1. Sizes of molecular weight standards are shown in kDa.



Lane 1 pEB2 Kan^R in pGEM5zf(+)

Lane 2 pEB3 $3C^{pro}$ in pGEM5zf(+)

Lane 3 pTet.1 Tet^R in pGEM5zf(+)

Lane 4 pTet.2 Tet^R in pGEM5zf(+)

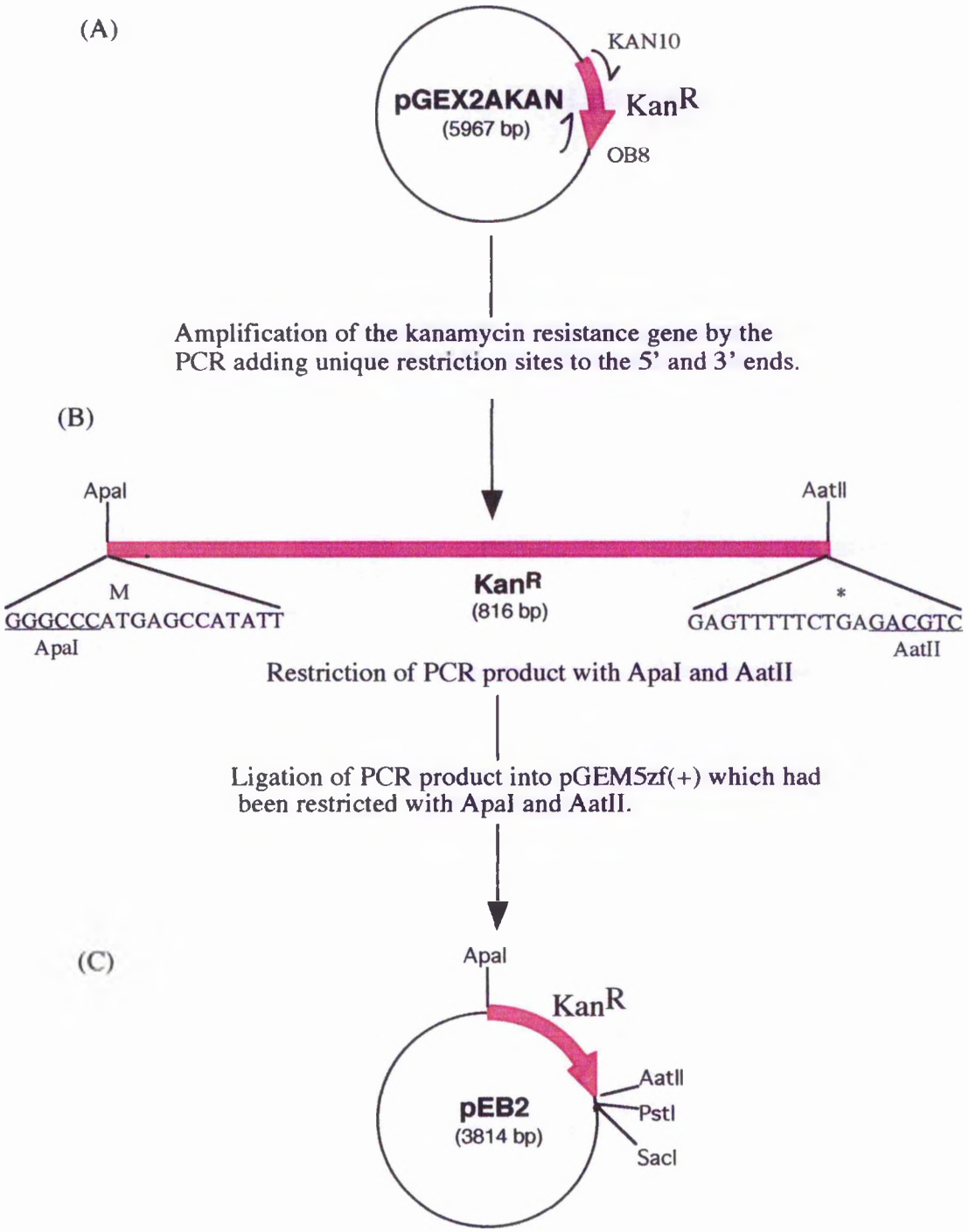
3.3 Construction of the control plasmid pEB2

The kanamycin resistance gene (Kan^R) was amplified by the PCR from the plasmid pGEX2AKAN (constructed previously by Dr Martin Ryan). Oligonucleotide primers were designed to anneal to the 5' and 3' ends of the gene and to add restriction enzyme sites to the termini of the PCR product. An ApaI site was added to the 5' end and an AatII site to the 3' end (for strategy see section 3.1). Both these sites are unique in the multiple cloning site of the chosen transcription vector, pGEM5zf(+) and are not present in the wild type sequence of the kanamycin resistance gene.

The PCR product was digested with AatII and ApaI to remove any extraneous 5' bases added by the PCR and to prepare for ligation into pGEM5zf(+) similarly restricted.. The PCR product was then purified by agarose gel extraction to remove any contaminants. pGEM5zf(+) was digested with ApaI and AatII and purified. The purified, restricted, PCR product (Kan^R) was then ligated into the plasmid as described in the methods section (Figure 3.3). This plasmid was constructed as a control for the reporter polyprotein system and also a precursor in the construction of the reporter system.

The gene has its own methionine codon at the 5' end for the initiation of translation. These bases were retained in the amplified gene as these will function as the initiation codon for the reporter system polyprotein (see below).

Figure 3.3 Construction of pEB2. The Kan^R coding sequence was amplified from pGEX2AKAN by the PCR using primers KAN10 and OB8 (A). The PCR product was restricted withApaI and AatII (B) and ligated into pGEM5zf(+), which had been similarly restricted, to form the plasmid pEB2 (C).



3.4 Construction of the control plasmid pEB3.

The HRV14 3C protease coding sequence was amplified by the PCR. The template used was the full length HRV14 cDNA coding sequence cloned into a Bluescript vector under the control of T7 polymerase promoter sequence. The sequences of the oligonucleotide primers designed for the PCR can be seen below. The region which anneals to the HRV14 template is shown in red (Figure 3.4). The oligos shown below do not include a methionine start codon or a stop codon since this PCR product will be used in the reporter system and the reporter genes flanking the protease sequence will have these codons.

5' oligonucleotide PCR primer

Oligo sequence AatII M 3B->
OB4 CGCGCGGACGTCATGCGCCCAGTTGTTGTGCAA
3C->
Nucleotide sequence CGCCCAGTTGTTGTGCAAGGACCAAAC
of HRV14 3C cleavage site

Nucleotide sequence GACGTCATGCGCCCAGTTGTGCAAGGACCAAAC ->
of PCR product

3' oligonucleotide PCR primer

Oligo sequence PstI 3D->
OB6 GCGCGCCTGCAGTCTAGCTATTACTTGGCC
3C->
Reverse complement TCTAGCTATTACTTGGCCTTGTTTCT
of nucleotide sequence 3C cleavage site

Nucleotide sequence CTGCAGTCTAGCTATTACTTGGCCTTGTTTCT ->
of PCR product

Figure 3.4 Sequences of region of HRV14 template annealed to by $3C^{pro}$ PCR primers OB4 and OB6. Regions in red are those of the primer and template which anneal to each other. All other sequences are shown in black.

Six amino acids of the C-terminal end of HRV14 3B were included at the 5' end of the $3C^{pro}$ gene construct in order to maintain the Glu-Gly $3C^{pro}$ 3B/C cleavage site. The 3C/D cleavage site at the 3' end of the gene was maintained by the inclusion of the first 6 amino acids of $3D^{pol}$ (see Figure 3.4.1). This will create some discrepancy in the sizes of

translation products from restriction enzyme cleaved plasmids and protease cleaved polyprotein products but the differences will be very small. The protease cleavage sites and the restriction enzyme sites are separated by 6 amino acids at either end of the gene.

The PCR product was digested with appropriate restriction enzymes (AatII and PstI) to prepare for ligation and then purified to remove any contaminants. Ligation into pGEM5zf(+), digested with AatII and PstI, was carried out as described in the methods section (Figure 3.4.2). The putative clones resulting from transformation of *E.coli* with the ligation reactions were analysed by restriction digest and found to be correct.

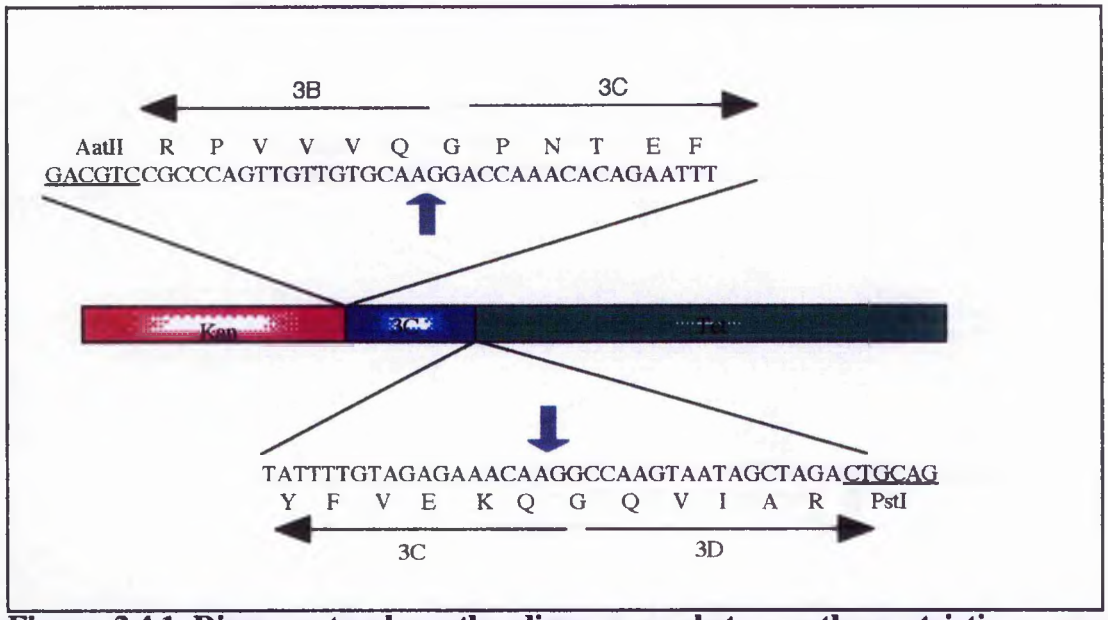
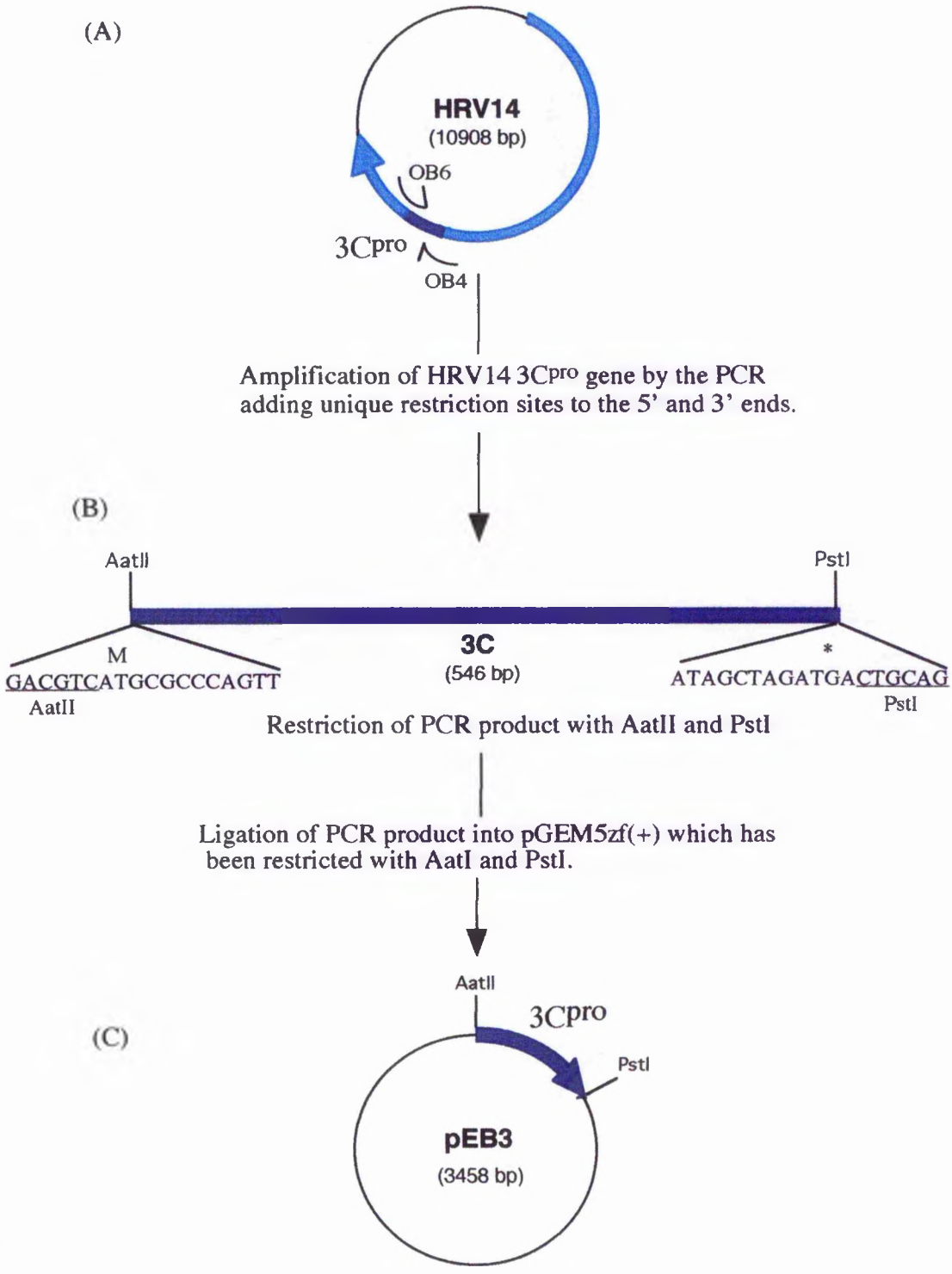


Figure 3.4.1 Diagram to show the discrepancy between the restriction enzyme cleavage sites and $3C^{pro}$ proteolytic cleavage sites.

Slightly different oligos were used to amplify the $3C^{pro}$ sequence for the control plasmid. These included a start codon in the 5' primer and a stop codon in the 3' primer.

Figure 3.4.2 Construction of pEB3. The HRV14 3C^{pro} coding sequence was amplified by the PCR using primers OB4 and OB6 (A). The PCR product was digested with AatII and PstI (B) and ligated into pGEM5zf(+), which had been similarly restricted, to form the plasmid pEB3 (C).



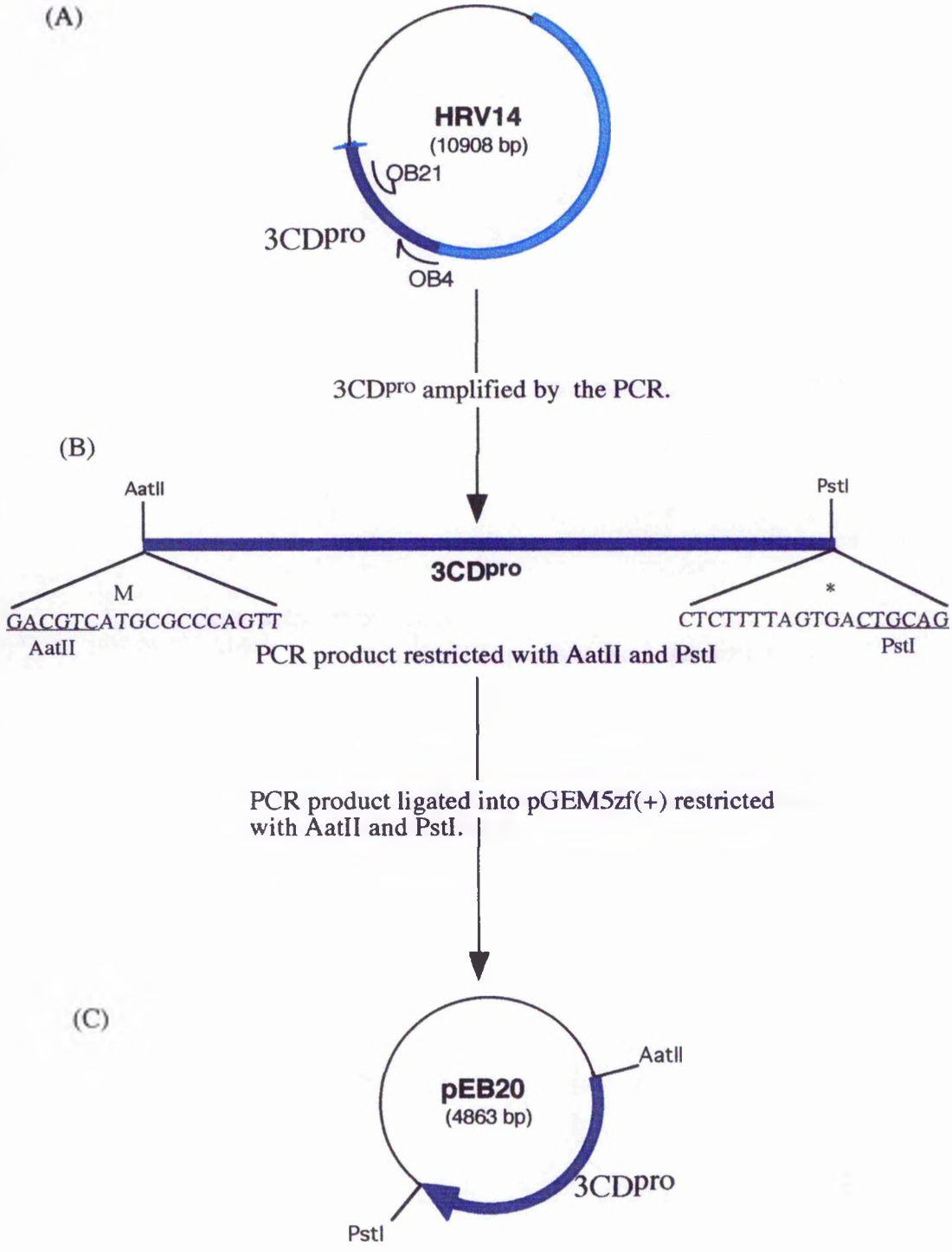
3.5 Construction of the control plasmid pEB20

HRV14 3CD^{pro} was amplified using the PCR from a HRV14 cDNA clone, ligated into a transcription vector similarly restricted. The PCR product was designed with restriction sites at the 5' and 3' ends so that it could be substituted into the reporter polyprotein system (see section 3.1). An AatII site was added to the 5' end and a PstI site to the 3' end. These are the same as the restriction sites on the 3C^{pro} PCR product so that these viral components of the reporter system could be swapped easily and the effect of the different 3C^{pro} containing sequences could be observed.

After digestion with AatII and PstI, the PCR product was ligated into pGEM5zf(+) to provide a control for the reporter polyprotein system. Slightly different oligos were used to amplify the gene for substitution into the reporter polyprotein system to replace 3C^{pro} (see section 4.1.3; Figure 3.5).

In order to maintain the 3C^{pro} cleavage sites in the gene, additional sequences were included in the PCR product. The six amino acids at the 3' end of 3B were included at the 5' end of the PCR product. This is the same strategy as employed in pEB3 to maintain the cleavage sites in 3C^{pro}. It was not necessary to include additional sequences at the 3' end of the gene as there is no 3C^{pro} cleavage site to be maintained.

Figure 3.5 Construction of pEB20. The coding sequence for HRV14 3CD^{pro} was amplified by the PCR using oligos OB4 and OB21 (A). The PCR product was restricted with AatII and PstI (B) and then ligated into pGEM5zf(+) which had been similarly restricted. This formed plasmid pEB20 (C).



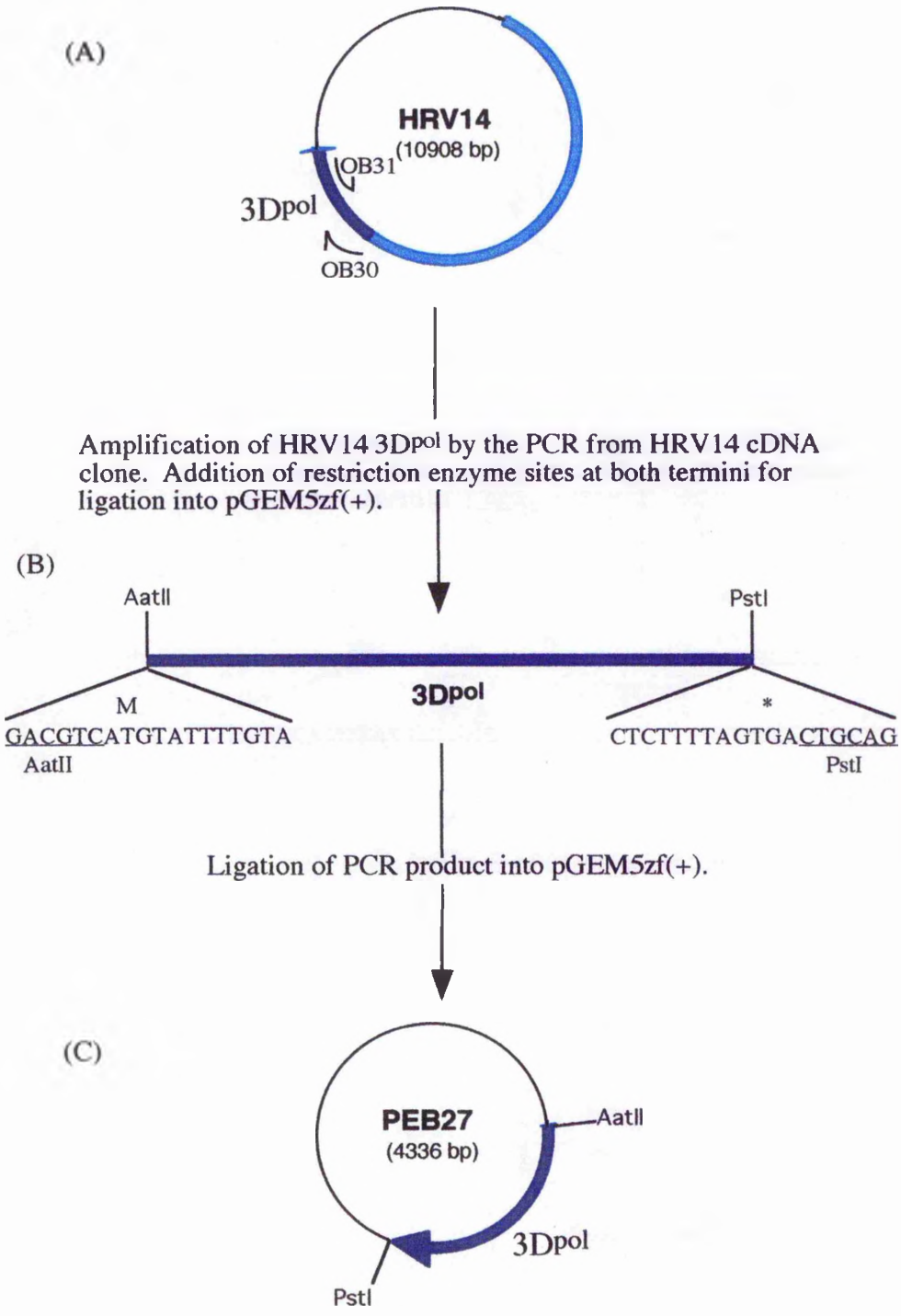
3.6 Construction of control plasmid pEB27

In order to provide a control for the cleavage of $3CD^{pro}$, both in the reporter system pEB25 and in pEB20, $3D^{pol}$ was amplified from HRV14 by the PCR. Six amino acids from the 3' end of $3C^{pro}$ were incorporated in the PCR product. The additional amino acids were not strictly necessary as there was no need to maintain the $3C^{pro}$ 3C/D cleavage site in this plasmid. This gene was intended only as a control and not as part of the reporter system and it was therefore not necessary to make an alternative PCR product without the initiation and termination sequences. The same primer used for the amplification of $3CD^{pro}$ was used at the 3' end and therefore a PstI site was incorporated.

The 5' primer was designed with an AatII site at the 5' end of the gene to allow ligation into pGEM5zf(+) (Figure 3.6) and an initiation codon was included. The enzyme sites were the same as those on the viral components of the reporter system for convenience. A stop codon was also included at the 3' end. The PCR product was then restricted and ligated into pGEM5zf(+) which had been restricted with the same enzymes.

Putative clones were checked by restriction digest after transformation in *E.coli* and nucleotide sequences of those which appeared to be correct were confirmed by DNA sequencing. The plasmid DNA of one correct clone was amplified with a Qiagen maxiprep kit to provide DNA of pure quality for translation and also to increase the quantity to allow for repeated translation reactions.

Figure 3.6 Construction of pEB27. The HRV14 3Dpol coding sequence was amplified by the PCR using oligos OB30 and OB31 (A). The PCR product was restricted with AatII and PstI (B) and ligated into pGEM5zf(+) which had been similarly restricted. The resulting plasmid was named pEB27 (C).



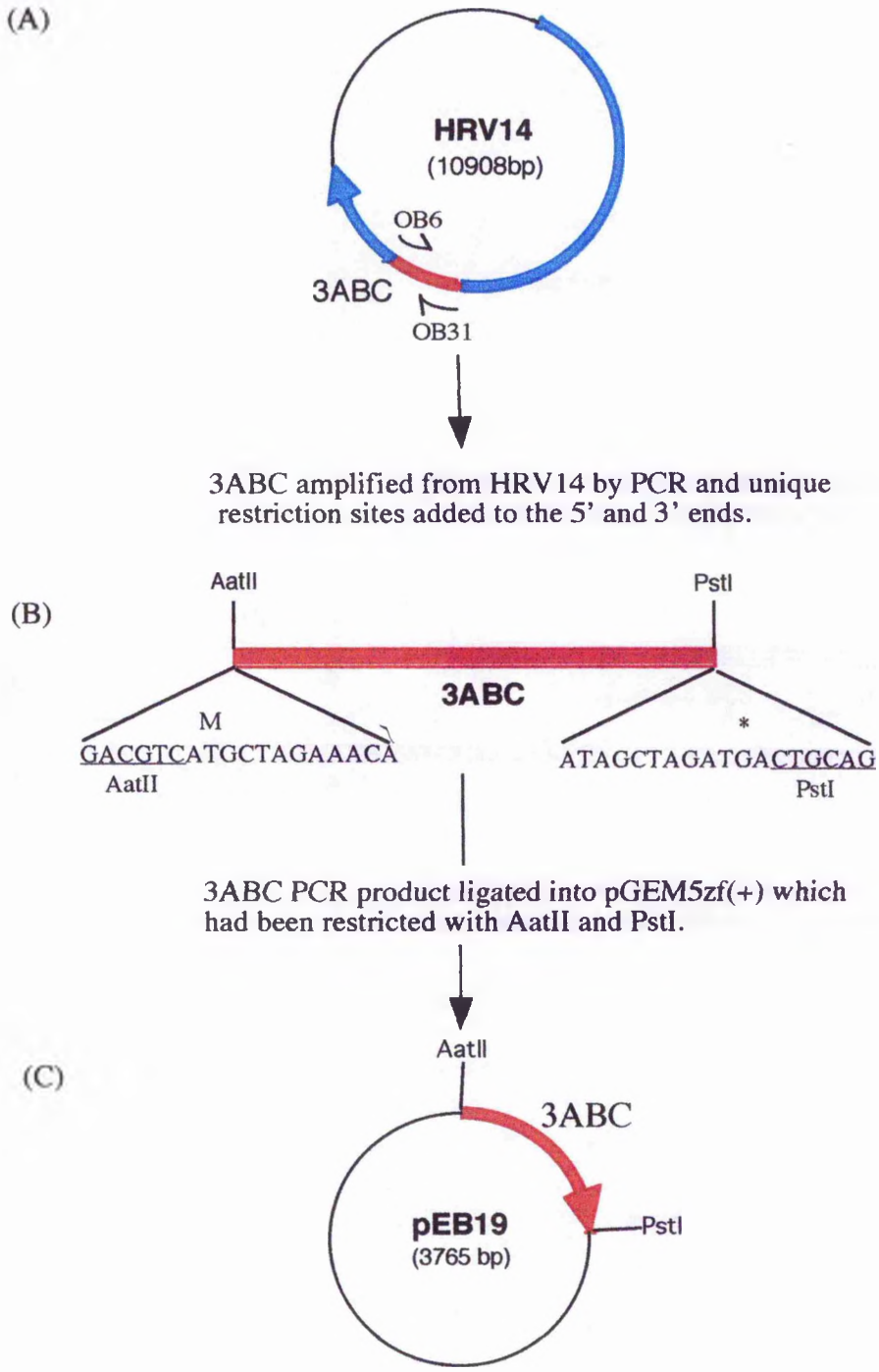
3.7 Construction of pEB19

HRV14 3ABC was amplified by the PCR from the HRV14 full length cDNA clone . Restriction enzyme sites were added to the 5' and 3' ends of the gene by incorporation of the sequences in the oligonucleotide primers designed for the PCR (see section 3.1). An AatII site was added to the 5' end and a PstI site to the 3' end. These would allow easy substitution of the genes into the reporter gene system to replace $3C^{pro}$. As in the other viral components of the reporter system additional amino acids were incorporated in the PCR product to maintain the $3C^{pro}$ cleavage sites. The six amino acids at the 3' end of 2C were incorporated at the 5' end of 3ABC and the six amino acids at the 5' end of $3D^{pol}$ were included at the C-terminus.

The PCR product was restricted with AatII and PstI and then purified. It was ligated into pGEM5zf(+), which had been restricted with AatII and PstI, for use as a control for the reporter polyprotein system (Figure 3.7). An alternative PCR product was also made without the initiation and termination sequences for use in the reporter system.

After transformation into *E.coli* putative clones were screened by restriction digest and of those found to be correct one was selected and a large scale DNA preparation was carried out, providing DNA for subsequent translation analysis.

Figure 3.7 Construction of pEB19. The coding sequence for HRV14 3ABC was amplified by the PCR using oligos OB31 and OB6 (A). The PCR product was digested with AatII and PstI (B) and ligated into pGEM5zf(+) which had ben similarly restricted. The resulting plasmid was pEB19 (C).

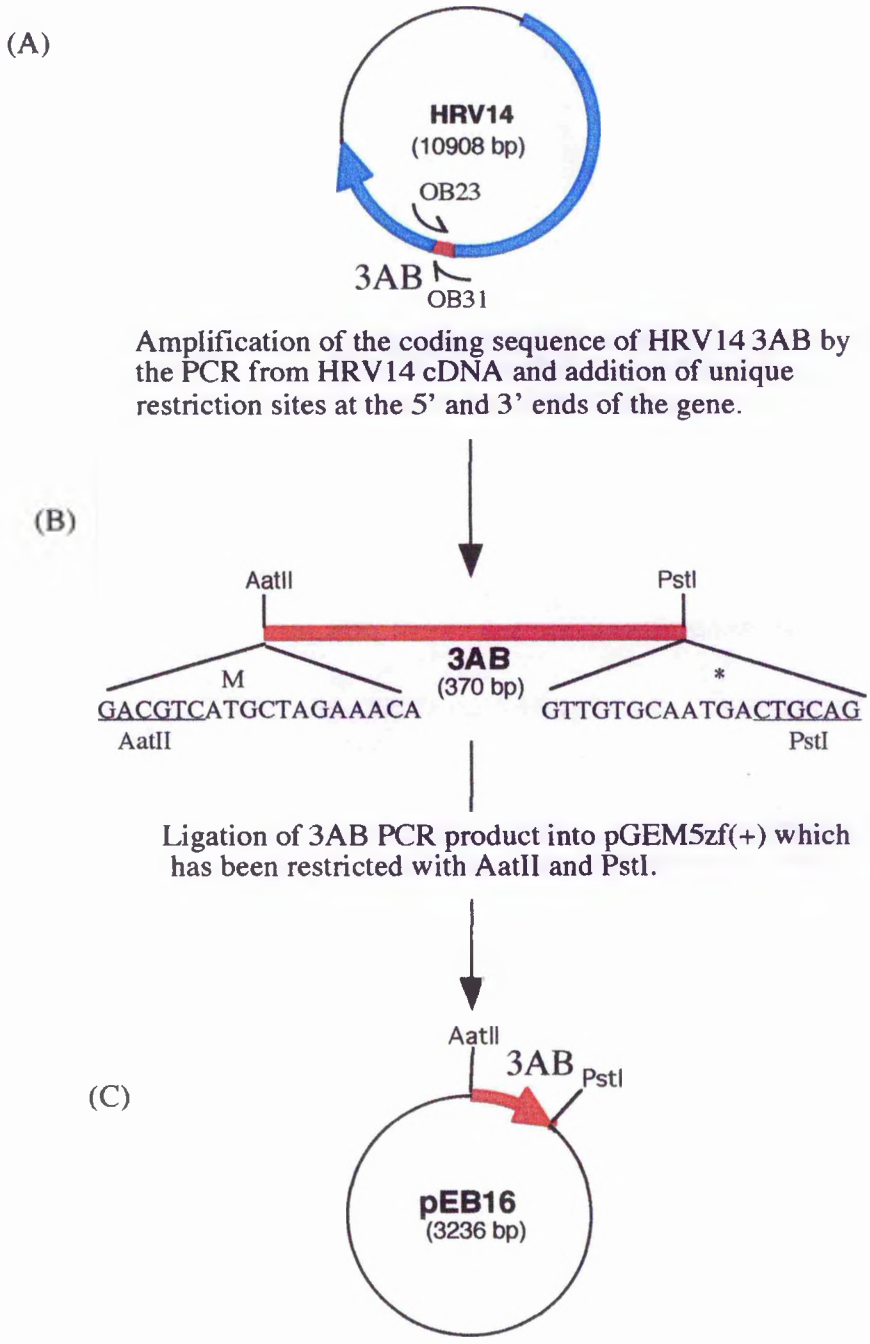


3.8 Construction of pEB16

In order to provide a control for the cleavage of 3ABC and cleavage within the reporter system HRV14 3AB was amplified from HRV14 cDNA cloned into a transcription vector . Oligonucleotide primers were designed to anneal to the 5' and 3' ends of the gene and also to include restriction enzyme sites at either end of the gene for easy ligation into a transcription vector (Figure 3.8). Additional amino acids were not included at the C-terminus of the gene as it was only to be used as a control and was not subject to cleavage by $3C^{pro}$. The same 5' oligonucleotide primer was used to amplify both 3ABC and 3AB so 6 amino acids of the 3' end of 2C were included at the 5 end of 3AB. These were not strictly necessary as the plasmid would not be subject to cleavage by $3C^{pro}$ but since they will not be detrimental the same primer could be used. The additional six amino acids incorporated at the termini of the viral components of the reporter system do not make a significant difference to their molecular weight on translation. Therefore the addition of amino acids at the C-terminus of 3AB was not necessary and this would still provide a reliable control for the cleavage products of 3ABC and the reporter polyprotein system.

A summary diagram of the constructs made is shown in Figure 3.12 on page 108.

Figure 3.8 Construction of pEB16. The coding sequence of HRV14 3AB was amplified by the PCR using oligos OB23 and OB31 (A). The PCR product was restricted with AatII and PstI (B) and ligated into pGEM5zf(+), which had been similarly restricted, to form the plasmid pEB16 (C).

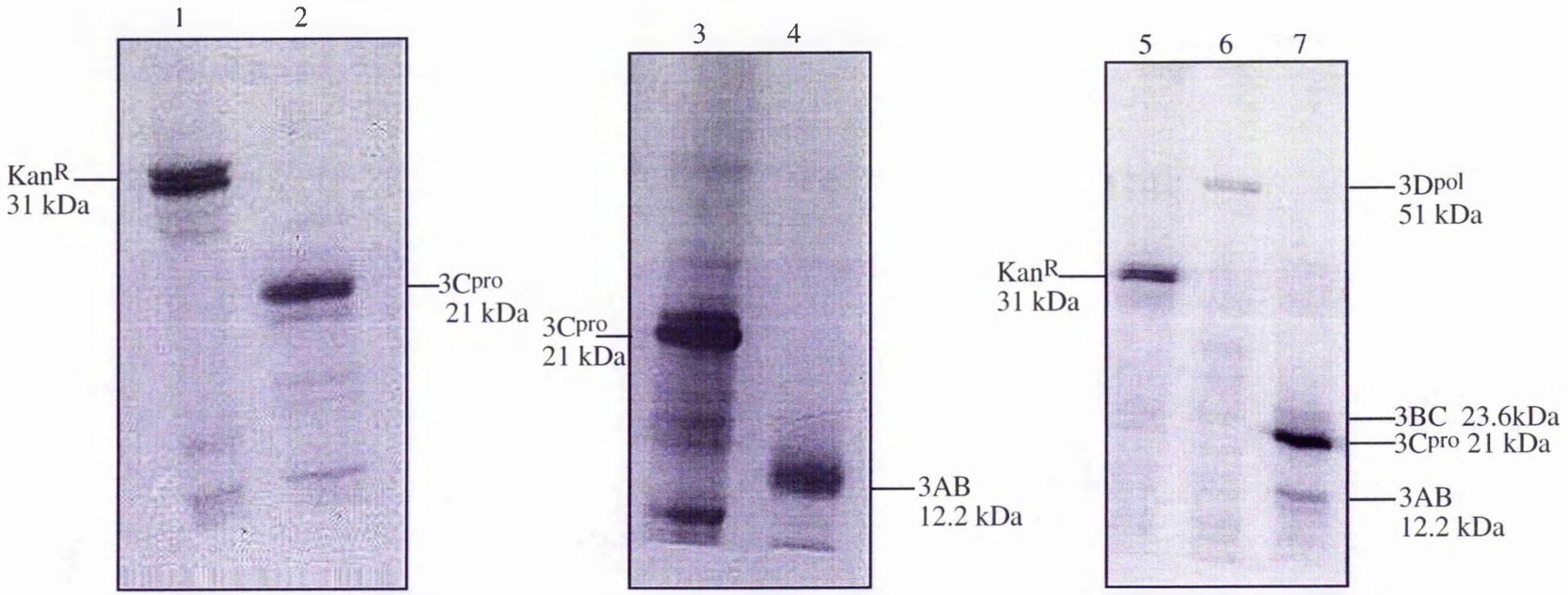


3.9 Translation of control plasmids in a wheatgerm extract transcription translation coupled system.

All control plasmids were used to program WGE TnT reactions to check that the products were of the correct size. All plasmids were translated and the translation products can be seen in figure 3.9.1. pEB20 did not produce any translation products and despite repeated attempts to produce a successful transcript, including re-cloning of the PCR product, no product was obtained. However HRV14 $3CD^{pro}$ was also cloned into the vector pRSETA (see section 5.1.2) and this plasmid, pEB7, was used to proved a control for $3CD^{pro}$.

In the case of plasmids pEB7 [$3CD^{pro}$] and pEB21 [3ABC] these reactions were also used to observe endogenous processing (see sections 3.10 and 3.11).

Figure 3.9 Translation of control plasmids in WGE TnT systems. Control plasmids were used to program TnT reactions to see whether the translation products were of the correct size.



Lane 1 pEB2 KanR in pGEM5zf(+)

Lane 2 pEB3 3Cpro in pGEM5zf(+)

Lane 3 pEB3 3Cpro in pGEM5zf(+)

Lane 4 pEB16 3AB in pGEM5zf(+)

Lane 5 pEB2 KanR in pGEM5zf(+)

Lane 6 pEB27 3Dpol in pGEM5zf(+)

Lane 7 pEB19 3ABC in pGEM5zf(+)

3.10 Endogenous processing of HRV14 3CD^{pro}.

Translation of a plasmid containing the gene for HRV14 3CD^{pro}, pEB20, showed no translation products. Therefore a plasmid encoding 3CD^{pro} in pRSETA was used (Figure 3.10). The use of molecular weight markers allowed the identification of the largest translation product as 3CD^{pro} and the presence of 3C^{pro} (Figure 3.10.1) confirmed this. The other bands are therefore probably internal initiation products and not self-cleavage products. A densitometric analysis of a phosphorimage of the translation products of pEB7 with Mac Bas version 2 software shows that internal initiation products are produced with very high efficiency in pEB7. The percentage initiation has been calculated for different methionine codons in the 3CD^{pro} coding sequence (Table 3.1).

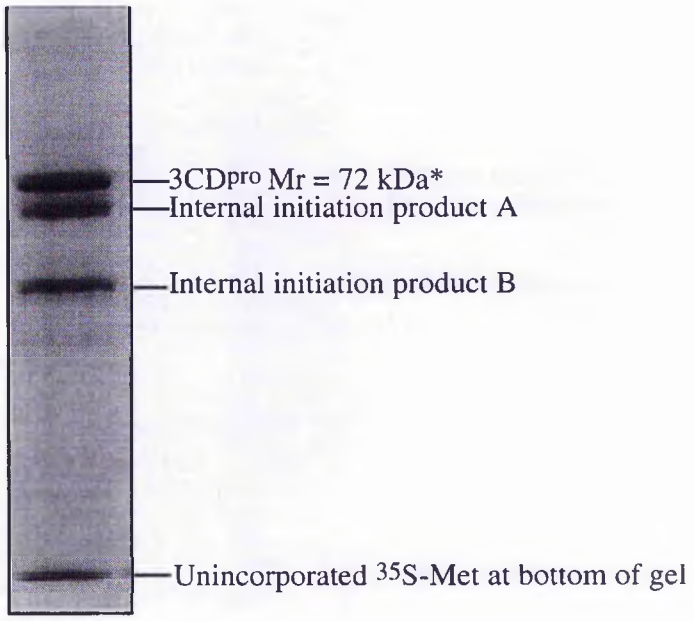
Translation of the plasmid in a rabbit reticulocyte lysate TnT reaction shows the internal initiation products even more clearly (Figure 3.10.1). The kozak consensus sequences for all of the in-frame methionine codons are shown below (Figure 3.10.2) and the molecular weights have been assigned to the bands on the gel using molecular weight markers (not shown). Some internal initiation products have co-migrated in figure 3.10 and percentage initiation has been calculated taking this into account. Mean percentage initiation is as follows:

3CDMet-1 +Δ3CDMet-17	73.1%
Δ3CDMet-121 + Δ3CDMet-135	6.6 %
Δ3CDMet-324	18.2 %

The kozak consensus sequences for the methionines shown to initiate internally are quite strong and this further confirms that they would give visible translation products. Due to time constraints densitometric analysis of 3CD^{pro} translation products in figure 3.10.1 was not possible. This would allow more accurate calculation of percentage initiation as there is better separation of the bands.

The plasmid pEB7 translated well giving discrete bands which were able to be identified as 3CD^{pro} and internal initiation products. No self-cleavage was observed in

Figure 3.10 Translation of pEB20 in a rabbit reticulocyte lysate coupled system.
Translation of a plasmid containing the $3CD^{pro}$ gene from HRV14 in a coupled transcription translation rabbit reticulocyte lysate reaction. Some of the internal initiation products of $3CD^{pro}$ can be seen clearly . Densitometric analysis was carried out to calculate the percentage initiation at each methionine.

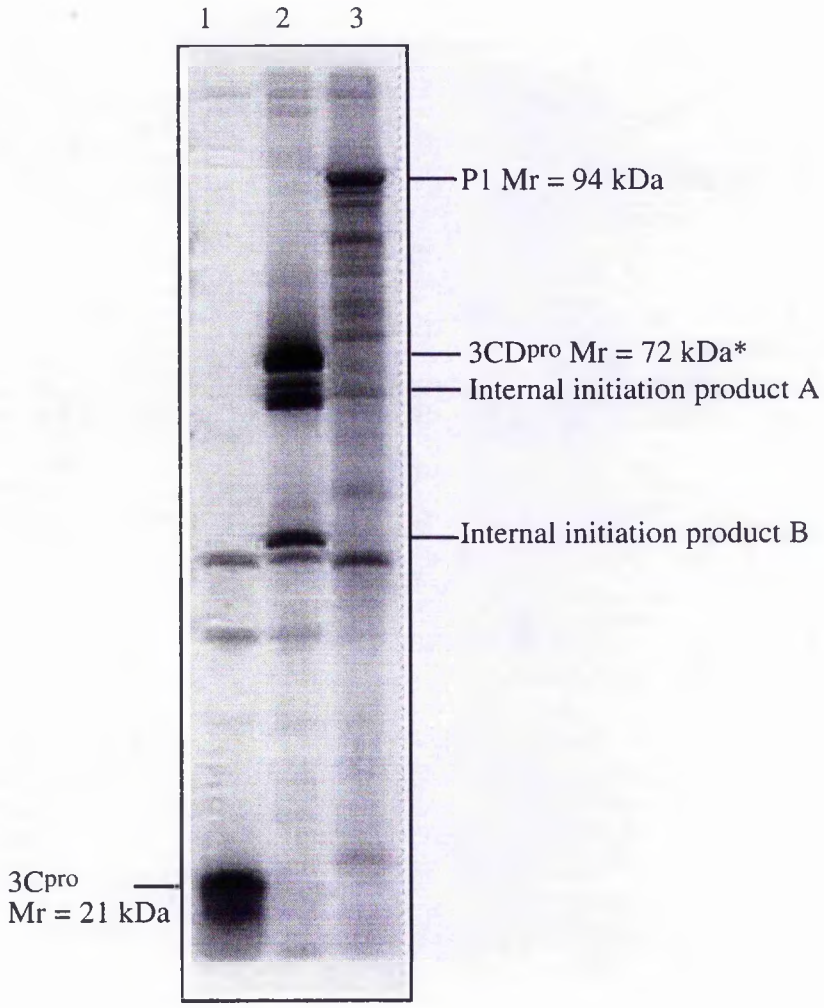


* also includes internal initiation product from Met-17 with a predicted molecular weight of 70.5 kDa.

Internal initiation product A = Δ 3CDMet-121 (Mr = 59 kDa)
+Δ3CDMet-135 (Mr= 57.5 kDa.

Internal initiation product B = Δ 3CDMet-324 (Mr = 36.7 kDa)

Figure 3.10.1 Endogenous processing of 3CD^{pro}. In order to clarify the identification of the internal initiation products of 3CD^{pro} the samples were analysed on a larger SDS-PAGE gel. The other two plasmids were used to provide some size controls.



Lane 1 pEB3 3C^{pro} pGEM5zf(+)

Lane 2 pEB20 3CD^{pro} in pGEM5zf(+)

Lane 3 P1 in pBSKSII+

*also includes Δ3CDMet-17 (Mr = 70.2 kDa)

Internal initiation product A = Δ 3CDMet-121 (Mr = 59 kDa)
+Δ3CDMet-135 (Mr= 57.5 kDa.

Internal initiation product B = Δ 3CDMet-324 (Mr = 36.7 kDa)

the system, *i.e. in cis*.. Self-cleavage *in trans* was also not observed in this system (see section 5.2).

Kozak consensus	GCC	GCC	A/GCC	AUG	G	Predicted Molecular weight
MET 1	TGG	TAC	CAA	AUG	G	M_r = 72.3 kDa
MET 17	AAA	AAC	ATA	AUG	A	Mr = 70.5 kDa
MET 121	CCT	GTA	ACA	AUG	G	Mr = 59.0 kDa
MET 135	ACT	AAC	AGA	AUG	A	Mr = 57.5 kDa
MET 257	ACT	GAA	AAT	AUG	C	Mr = 43.9 kDa
MET 324	ACA	GAA	AAG	AUG	A	Mr = 36.7 kDa
MET 370	TCT	GTT	AAC	AUG	A	Mr = 31.4 kDa
MET 372	AAC	ATG	AGA	AUG	A	Mr = 31.0 kDa
MET 1328	CCT	TGT	TTA	AUG	G	Mr = 27.2 kDa
MET 1233	GGG	CAC	CTG	AUG	G	Mr = 26.6 kDa
MET 1398	GAA	GGT	GGC	AUG	C	Mr = 20.6 kDa
MET 1441	TTC	AAT	TCC	AUG	A	Mr = 19.3 kDa
MET 1648	TTT	ACA	AAA	AUG	A	Mr = 11.8 kDa
MET 1714	CAC	CCA	GTT	AUG	C	Mr = 8.5 kDa
MET 1716	GTT	ATG	CCC	AUG	A	Mr = 8.3 kDa
MET 1801	TCA	TTA	TGC	AUG	T	Mr = 5.3 kDa

Figure 3.10.2 Kozak consensus sequences and predicted molecular weights for internal initiation products of the 3CD^{pro} gene. The residues which are underlined in each sequence are very important for recognition. Residues shown in red in each sequence are those which occur in the kozak consensus sequence.

Band	Lane	1			2			3			4			
	No. of Mets	PSL-BG	PSL-BG No. of mets	% initiation	PSL-BG	PSL-BG No. of mets	% initiation	PSL-BG	PSL-BG No. of mets	% initiation	PSL-BG	PSL-BG No. of mets	% initiation	Mean value%
3CDMet-1 + 3CDMet 17	16 +15 =31	6052.0	195.2	73.5	5731.0	184.9	74.3	4726.0	152.5	75.8	4164	134.3	68.8	73.1
3CD Met 121 +Met 135	14+13 =27	504.9	18.7	7.0	523.0	19.4	7.8	461.8	17.1	8.5	586.5	21.7	11.1	6.6
3CD Met 135	11	570.7	51.8	19.5	489.0	44.5	17.9	349.1	31.7	15.4	432.4	39.3	20.1	18.2
Total	-	-	265.7				248.8		201.3			195.3		

Table 3.1 Calculation of percentage internal initiation from methionine codons in $3CD^{pro}$ in a TnT system.

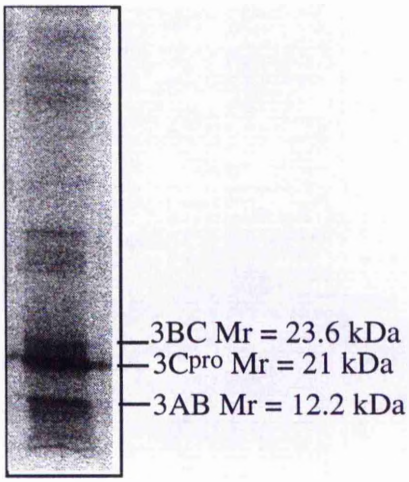
In order to determine the percentage initiation of $3CD^{pro}$ at each of the methionines when incubated in a coupled transcription translation system the SDS-PAGE gel was visualised by phosphorimaging and then quantified using Mac Bas version 2 software. PSL values were calculated and a background measurement subtracted. These values were corrected for the number of methionines in each protein to give a figure which is proportional to the number of moles of each protein product. Percentage initiation was calculated for four translations of the plasmid (results not shown) and an average value calculated.

3.11 Endogenous processing of HRV14 3ABC.

Translation of the plasmid pEB19 in a coupled TnT system shows that HRV14 3ABC processes to give the products 3AB and $3C^{pro}$ (Figure 3.11). Cleavage is occurring at the glutamine-glycine cleavage site between 3AB and $3C^{pro}$ which is seen in the virus when $3C^{pro}$ cleaves N-terminally. This cleavage is fairly rapid in the virus polyprotein. The production of the cleavage products indicates that the protease is able to cleave N-terminally in the WGE TnT system and at the same sites as in the virus. Molecular weight markers indicate the sizes of the cleavage products and allow preliminary identification. Control plasmids pEB3, encoding HRV14 $3C^{pro}$, and pEB16, encoding HRV14 3AB, provide confirmation of the identity of the cleavage products (data not shown). Time constraints prevented densitometric analysis of the cleavage products.

The production of the cleavage product 3BC shows that cleavage is also occuring at a Q-G cleavage site at the N-terminus of 3B. This is not usually utilised in the virus as the cleavage between 3A and 3B occurs during the final stages of virion morphogenesis. The mechanism of this cleavage in the capsid is not known and it is referred to as the maturation cleavage. Previous work has demonstrated that cleavage between 3A and 3B is rarely seen *in vitro* unless all other processing routes are blocked (Parks *et al.*, 1986). $3C^{pro}$ is utilising sites which are not usually cleaved in the virus even though they have the correct amino acid sequence at P_1 and P_1'.

Figure 3.11 Transcription and translation of a plasmid containing HRV14 3ABC. A plasmid containing the coding sequence for HRV14 3ABC was transcribed and translated in a coupled WGE system. The cleavage of 3ABC can be seen, yelding 3AB and $3C^{pro}$ as the cleavage products. There is also a small amount of 3BC present. Cleavage seems to be complete and densitometric analysis was carried out to confirm this.



3.12 Conclusions

The individual components of the artificial reporter polyprotein system have all been amplified from plasmids and ligated into transcription vectors to act as controls in translation systems and show the endogenous processing by $3C^{pro}$ in regions of HRV14 P3. When used to program coupled wheatgerm extract transcription translation systems products of the expected sizes were produced except in the case of pEB20 [$3CD^{pro}$]. This was compensated for by the use of plasmid pEB7 [$3CD^{pro}$ in pRSETA].

- Translation of pEB7 to show endogenous processing in $3CD^{pro}$ yields $3CD^{pro}$ and very strong internal initiation products. There was no observed processing of $3CD^{pro}$ to $3C^{pro}$ and $3D^{pol}$.

- 3ABC processes to give 3AB, $3C^{pro}$ and 3BC. Cleavage is occurring at the N-terminus of $3C^{pro}$ at a glutamine-glycine cleavage site and at a glutamine glycine site at the N-terminus of 3B not utilised in the viral polyprotein processing pathway.

It can therefore be concluded that HRV14 $3C^{pro}$ will process both N and C-terminally in the WGE coupled transcription translation system and this system can be used to monitor cleavage by HRV14 $3C^{pro}$ in artificial reporter polyproteins. $3C^{pro}$ is also utilising cleavage sites which are not part of the viral polyprotein processng pathway. The lack of cleavage in $3CD^{pro}$ is consistent with previous findings (see section 3.13) and does not indicate that $3C^{pro}$ will not cleave C-terminally in this system.

Summary

- Components of a reporter polyprotein system have been cloned into transcription vectors for use as controls and as precursors to the reporter system.
- $3C^{pro}$ will cleave at both its N-terminus in the coupled WGE TnT system with reporter proteins flanking the protease sequence.

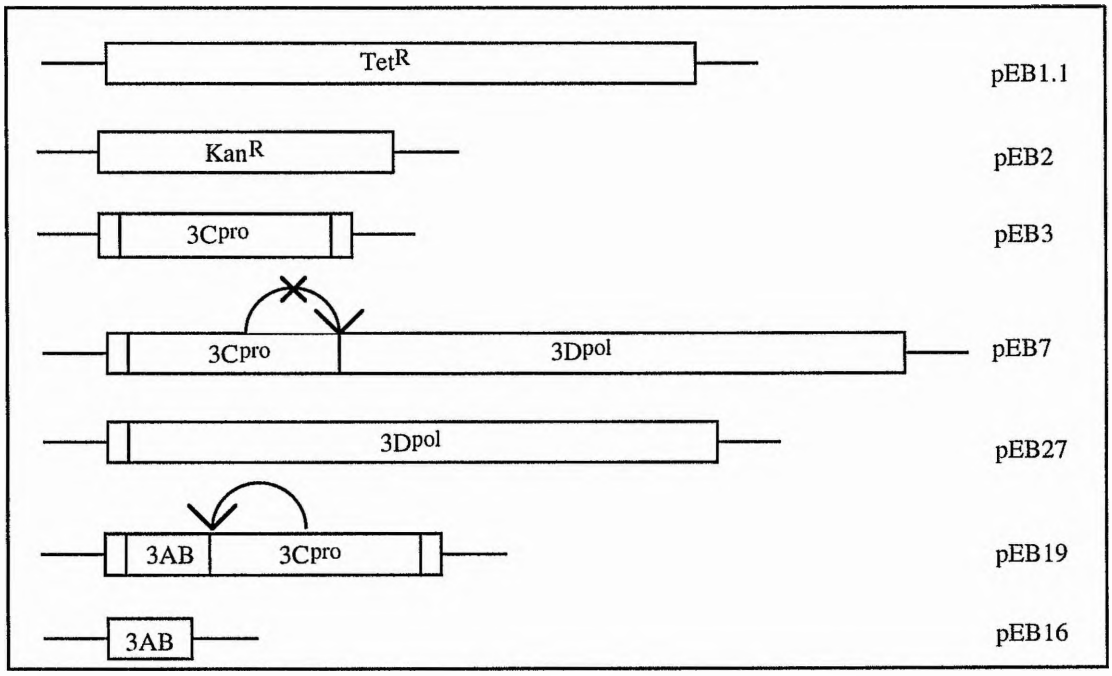


Figure 3.12 Summary diagram to show the plasmids constructed as controls and precursor for the reporter polyprotein system

3.13 Discussion

Plasmids have been constructed which will allow the construction of artificial reporter gene systems. These will be translated to produce artificial polyproteins which will allow cleavage by HRV14 $3C^{pro}$ to be monitored. The potential use of the antibiotic resistance genes as phenotypic "reporters" will allow expression of antibiotic resistance in *E.coli* to be used to indicate the site of cleavage by $3C^{pro}$ in such reporter polyprotein systems.

The construction of control plasmids for the translation of the reporter gene systems has also allowed endogenous cleavage of some viral protein precursors to be monitored. $3C^{pro}$ processes N-terminally in 3ABC.

Endogenous processing of $3CD^{pro}$ shown by translation of the plasmid pEB7 shows that $3CD^{pro}$ does not cleave to $3C^{pro}$ and $3D^{pol}$ in this system. This result is consistent with previous findings *in vitro* (Harris *et al.*, 1992) and *in vivo* (Porter *et al.*, 1993). Cleavage at the 3C/D junction appears to occur more slowly than at other $3C^{pro}$ cleavage sites in the polyprotein (Harris *et al.*, 1992) which indicates that $3CD^{pro}$ has a distinct role in the proteolytic pathway rather than just being a precursor for $3C^{pro}$ and $3D^{pol}$.

Endogenous processing of 3ABC shows cleavage to yield 3AB and $3C^{pro}$. 3ABC has no known distinct function and so cleavage to yield these products is its major function. Cleavage of 3AB is rarely seen *in vitro* unless other processing pathways are blocked (Parks *et al.*, 1989) but in this system $3C^{pro}$ is utilising a glutamine-glycine cleavage site at the N-teminus of 3B which is not used in viral polyprotein processing.

4. Proteolytic activity of HRV14 proteases in an artificial reporter polyprotein system.

In vivo, HRV14 $3C^{pro}$ can cleave within the viral polyprotein either N-terminally or C-terminally. C-terminal cleavage results in the production of 3ABC, $3C^{pro}$ and the polymerase $3D^{pol}$. Cleavage N-terminally results in the production of 3AB and the protease $3CD^{pro}$. Cleavage at a combination of these sites may produce 3AB, $3C^{pro}$, $3D^{pol}$, $3CD^{pro}$ and 3ABC. The 3CD protease processes P1, the capsid precursor, *in trans.* It was shown in chapter 3 that this protease is a stable product of P3 cleavage and had been shown previously that processing to yield $3C^{pro}$ and $3D^{pol}$ occurs very slowly (Harris *et al.*, 1992). This indicates that $3CD^{pro}$ has a role as the uncleaved product. Stability of $3CD^{pro}$ has been demonstrated both *in vitro* (Harris *et al.*, 1992) and *in vivo* (Porter *et al.*, 1993), and that $3CD^{pro}$ lacks polymerase activity (Harris *et al.*, 1992).

The method of control of $3C^{pro}$ cleavage has not yet been determined but the virus must have some means of determining whether cleavage occurs N- or C-terminally or at both sites. Both $3C^{pro}$ and $3CD^{pro}$ are required in the processing pathway of the virus but at different locations and, quite possibly, at different points in time. Both these pathways result in the production of proteins which are essential for the replication of the virus. The 3C protease is also responsible for the primary cleavage of the polyprotein between P2 and P3 and most of the secondary cleavages in the processing pathway.

In order to observe the cleavage activity of $3C^{pro}$ without the influence of other viral sequences an artificial reporter polyprotein system was constructed. HRV14 $3C^{pro}$ was cloned into a transcription vector with two flanking reporter genes. Two antibiotic resistance genes were chosen as the non-viral components of the reporter system to allow possible selection of cleavage products by antibiotic resistance in *E.coli* (see section 4.1.5) as well as the identification of translation products in cell-free translation systems. The antibiotic resistance genes were ligated into a transcription vector flanking the viral sequence of interest. The system was also designed to allow the substitution of

components by digestion with the appropriate restriction enzymes at sites which flank each gene.

In the reporter polyprotein system the action of $3C^{pro}$ can be observed without the presence of other viral or cellular factors. A 'default' pathway may then be established for the autocatalytic action of $3C^{pro}$. The absence of other viral sequences and other factors, such as viral RNA, will allow $3C^{pro}$ to fold and then cleave without any constraints thus establishing a "default" pathway. Additional factors could then be added to the system to try and change the cleavage pattern and establish a possible mechanism(s) used in the virus to control the alternative processing of P3.

Viral proteins could be added to the reporter system *in cis* to observe their effect on the pattern of cleavage by $3C^{pro}$. Other putative switches, such as vRNA (shown to bind to the protease), may be added to the system *in trans* to observe their effect on the cleavage pattern. The virus must have some means of regulating the autoproteolytic cleavage of $3C^{pro}$ as both the N- and C-terminal cleavage products are required during replication, but at different times and in different regions of the polyprotein. This reporter polyprotein system provides a way of monitoring proteolytic cleavage activity in a cell-free system and observing the effects of addition of different factors, such as RNA and other viral proteins, on proteolytic cleavage both *in cis* and *in trans*.

In order to observe the effect of viral sequences on $3C^{pro}$ cleavage *in cis*, other P3 cleavage products were ligated into the reporter polyprotein system between the two reporter genes. pEB15 contains the viral protease sequence $3CD^{pro}$ and pEB25 contains 3ABC. The *in vitro* endogenous processing properties of $3CD^{pro}$ and 3ABC were demonstrated in chapter 3.

4.1 Assembly of artificial reporter polyprotein systems

4.1.1 Construction of pEB4 and pEB13.

The reporter plasmid pEB4 was constructed from the control plasmids shown in chapter 3. pEB13 [Kan^R3C^{pro}] was constructed from pEB2.1 and pEB3.1 to provide both a control for the reporter system and a precursor form (Figure 4.1.1). The plasmid was also designed to reveal the endogenous processing properties of [Kan^R3C^{pro}], $3C^{pro}$ processing, in this case, with a non-viral N-terminal sequence. The second reporter gene ,Tet^R, was then added to the plasmid pEB13 to complete the reporter gene system (Figure 4.1.2).

The $3C^{pro}$ gene has additional sequences added at the 5' and 3' ends to maintain the proteolytic cleavage sites. The 6 amino acids from the 5' end of the 3B sequence have been included in the $3C^{pro}$ gene and the 3' end of $3C^{pro}$ has six additional amino acids from the 5' end of the $3D^{pol}$ gene (see section 3.3).

The Kan^R gene has an initiation codon at its 5' end which will allow the translation of the three genes, (Kan^R, $3C^{pro}$ and Tet^R), as a polyprotein, a termination codon being provided at the 3' end of the Tet^R gene.

Figure 4.1.1 Construction of plasmid pEB13.
Plasmid pEB3.1 was restricted with AatII and PstI (A) to release $3C^{pro}$ (B). pEB2.1 was restricted with AatII and PstI (C) and the $3C^{pro}$ squence was ligated in to form plasmid pEB13 (D).

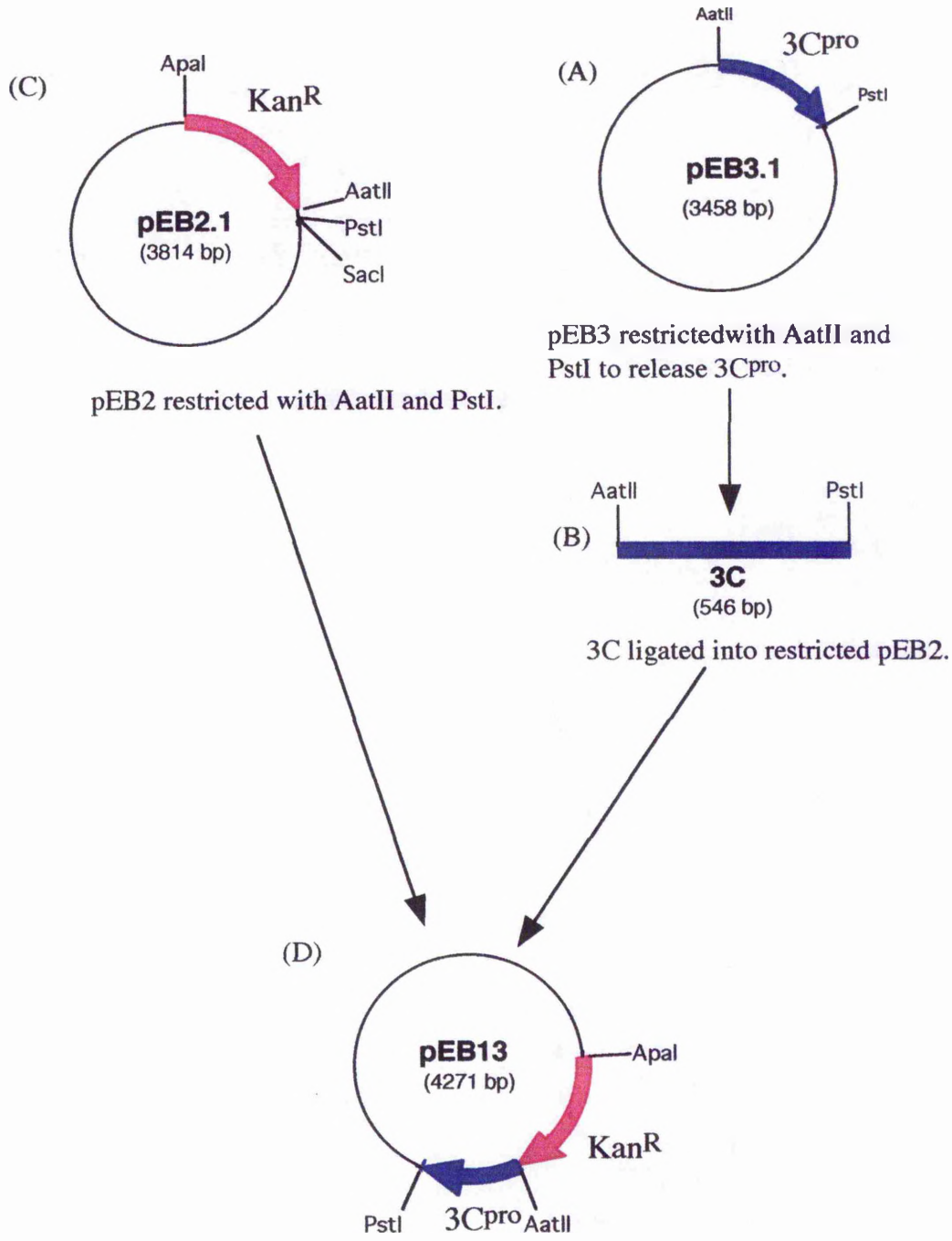
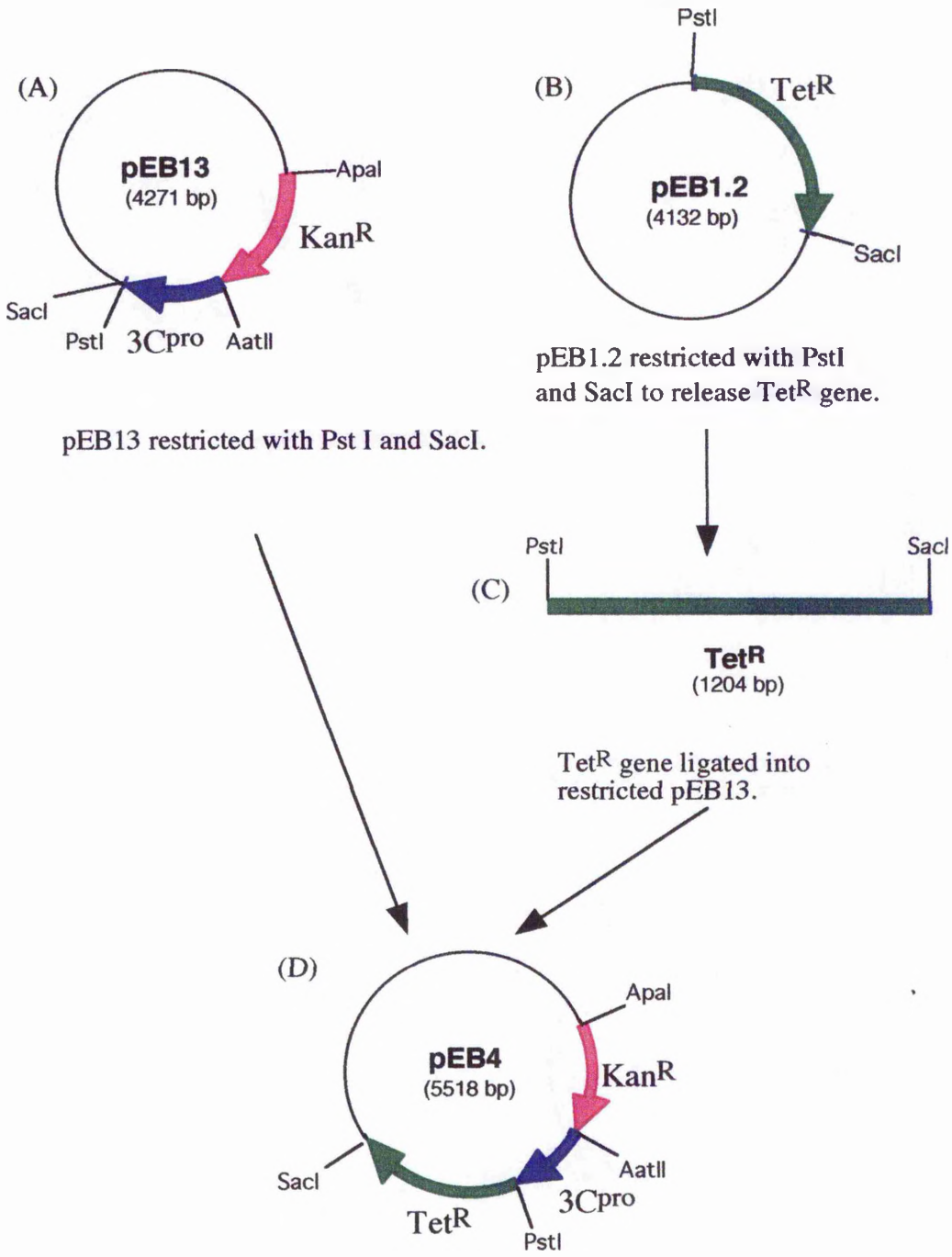


Figure 4.1.2 Construction of pEB4. Plasmid pEB13 was restricted with PstI and SacI (A). Plasmid pEB1.2 was restricted with PstI and SacI (B) to release Tet^R (C). TheTet^R fragment was ligated into restricted pEB13 to form pEB4 (D).



4.1.1.2 Sequencing of pEB4 and control plasmids.

DNA was prepared with a QIagen maxiprep kit and used for DNA sequencing. Control plasmids pEB1, pEB2 and pEB3 were sequenced using the cycle sequencing protocol detailed in the methods section. The plasmid pEB4 was sequenced on an ABI Prism™ automated sequencer using T7 and SP6 polymerase promoter primers and a primer designed to anneal to a sequence within the kanamycin resistance gene (see section 2.2.2).

4.1.2 Construction of pEB4.2

The catalytic nucleophile of poliovirus $3C^{pro}$ has been mutated previously to destroy the catalytic action of the protease. One of the mutations made to destroy this activity was the mutation of the cysteine residue at position 146 to an alanine (Lawson and Semler, 1991), the mutant protease being completely inactive. In order to confirm that the cleavage seen in pEB4 was carried out by $3C^{pro}$ the catalytic nucleophile was mutated by overlap PCR. Oligonucleotide primers were designed to introduce this mutation into the $3C^{pro}$ PCR product. Figure 4.1.3 shows the mutation which will be made in the $3C^{pro}$ sequence.

HRV14 $3C^{pro}$ sequences

Wild type	140										150
Amino acid		A	T	K	T	G	Q	C	G	G	V	L
Nucleotide	GCA	ACA	AAA	ACT	GGG	CAG	TGT	GGA	GGT	GTG	CTG	
Mutant												
Nucleotide	GCA	ACA	AAA	ACT	GGG	CAG	GCT	GGA	GGT	GTG	CTG	
Amino acid		A	T	K	T	G	Q	A	G	G	V	L

Figure 4.1.3 To show the mutation of the catalytic nucleophile in HRV14 $3C^{pro}$.

Oligonucleotide primers were designed such that the same restriction enzyme sites were incorporated at the 5' and 3' ends as in the original $3C^{pro}$ PCR product, *i.e.* AatII and PstI. The PCR product was digested with AatII and PstI (the enzymes whose sites had been engineered at the ends of the mutant PCR product) and then ligated into pEB4 which had been cut with AatII and PstI to release the wild type $3C^{pro}$ sequence (Figure 4.1.4).

As there was only one amino acid difference between the sequences of the mutant and the wild type $3C^{pro}$, the putative clones were checked with restriction enzyme digests and then by nucleotide sequencing. One of the putative clones was found to have the Cys->Ala mutation in the $3C^{pro}$ sequence at the correct position and this was used to program the TnT system (Figure 4.1.4.1).

Figure 4.1.4 Construction of pEB4.2. pEB4 was restricted with AatII and PstI (A) to release $3C^{pro}$ (B). Mutant $3C^{pro(C146A)}$ was amplified from a HRV14 cDNA clone by overlap PCR. The PCR product was restricted with AatII and PstI (C) and ligated into pEB4, similarly restricted, to form pEB4.2 (D).

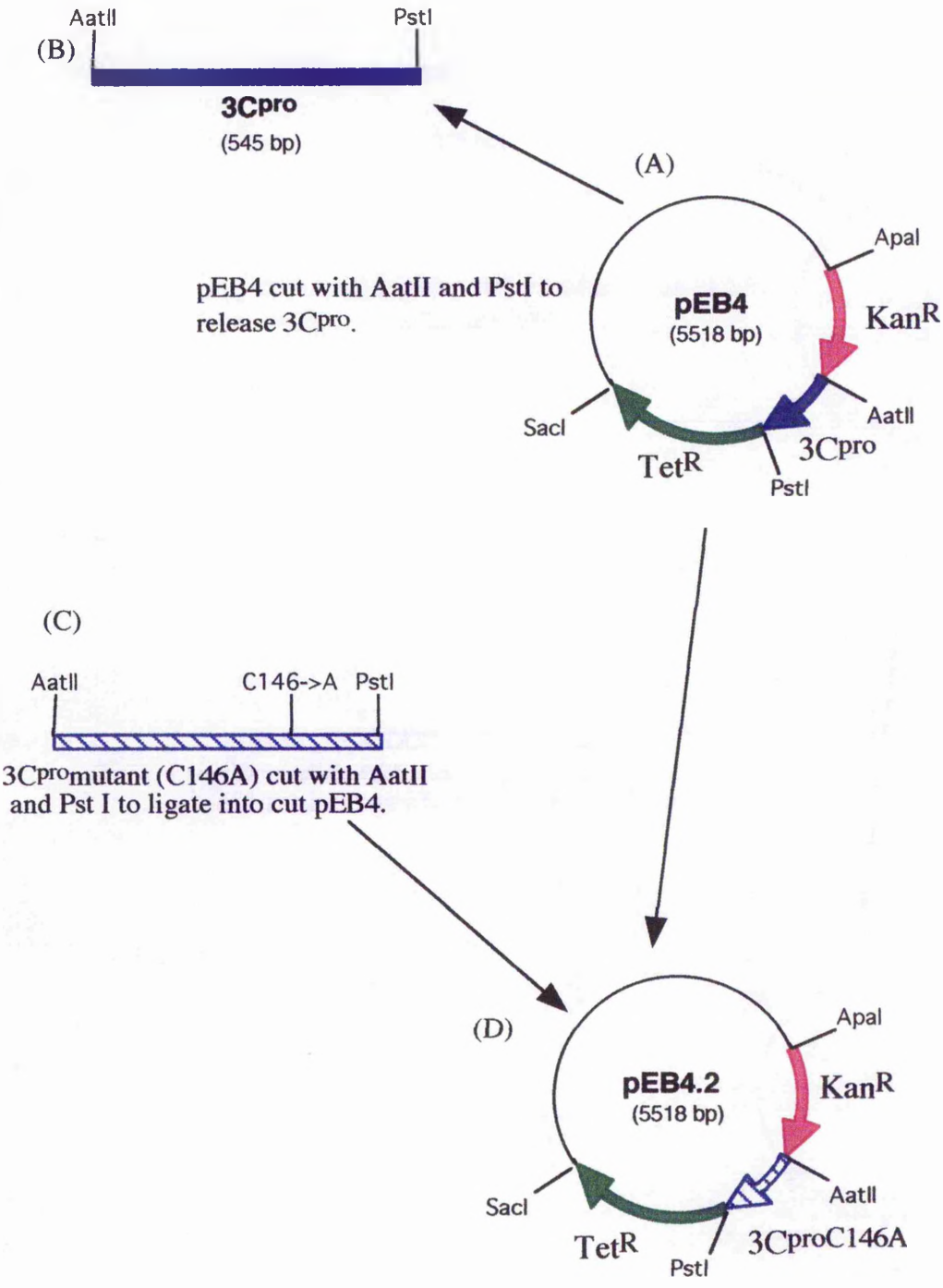
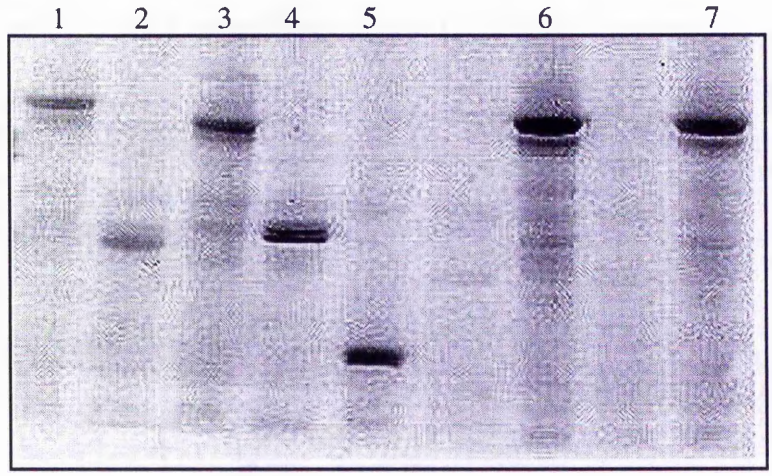


Figure 4.1.4.1 Translation of pEB4.2. Plasmids were used to program WGE coupled TnT reactions and the results analysed by SDS-PAGE.



Lane	Plasmid	Insert	Linearised by restriction with
1	pEB4	[Kan^{R}3C^{pro}Tet^{R}]	
2	pEB4	[Kan^{R}3C^{pro}Tet^{R}]	x AatII [1]
3	pEB4	[Kan^{R}3C^{pro}Tet^{R}]	x PstI [2]
4	pEB2	[Kan^{R}]	
5	pEB3	HRV14[3C^{pro}]	
6	pEB5	[3C^{pro}Tet^{R}]	
7	pEB4.2	[Kan^{R}3C^{pro}C146A Tet^{R}]	

1. AatII linearises pEB4 between the 3' end of the Kan^{R} gene and the 5' end of the 3C^{pro} gene.
2. PstI linearises pEB4 between the 3' end of the 3C^{pro} gene and the 5' end of the Tet^{R} gene.

4.1.3 Construction of plasmids pEB15 and pEB15.2.

$3CD^{pro}$ was amplified from HRV14 cDNA by the PCR using the oligonucleotide primers shown in Table 2.1. Restriction sites were added to the 5' and 3' ends of the genes in order to allow ligation into the plasmids encoding the reporter polyprotein system. In order to maintain the $3C^{pro}$ cleavage sites, as in pEB4, six amino acids from the viral sequence flanking the sequence of interest was included in the PCR products. Products were then digested with restriction enzymes and cloned into pEB4 which had been digested to release $3C^{pro}$ (Figure 4.1.5). All the HRV14 components have the same restriction sites at the 5' and 3' ends (*i.e.* AatII and PstI) for ease of ligation. Plasmid pEB15.2 was constructed to show the action of $3C^{pro}$ between two entities and can be compared more easily with pEB4 (Figure 4.1.6) Plasmid pEB15 shows the action of $3C^{pro}$ in the presence of other viral components as well as the two reporter genes.

As there is no $3C^{pro}$ cleavage site at the C-terminus of $3D^{pol}$ in the viral polyprotein no additional sequences were included in the artificial system.

Figure 4.1.5 Construction of pEB15. pEB20.2 was restricted with AatII and PstI (A) to release 3CD^{pro} (B). pEB4 was restricted with AatII and PstI (C) to release 3C^{pro} (D). The restricted 3CD^{pro} fragment was ligated into pEB4, similarly restricted, to form the plasmid pEB15 (E).

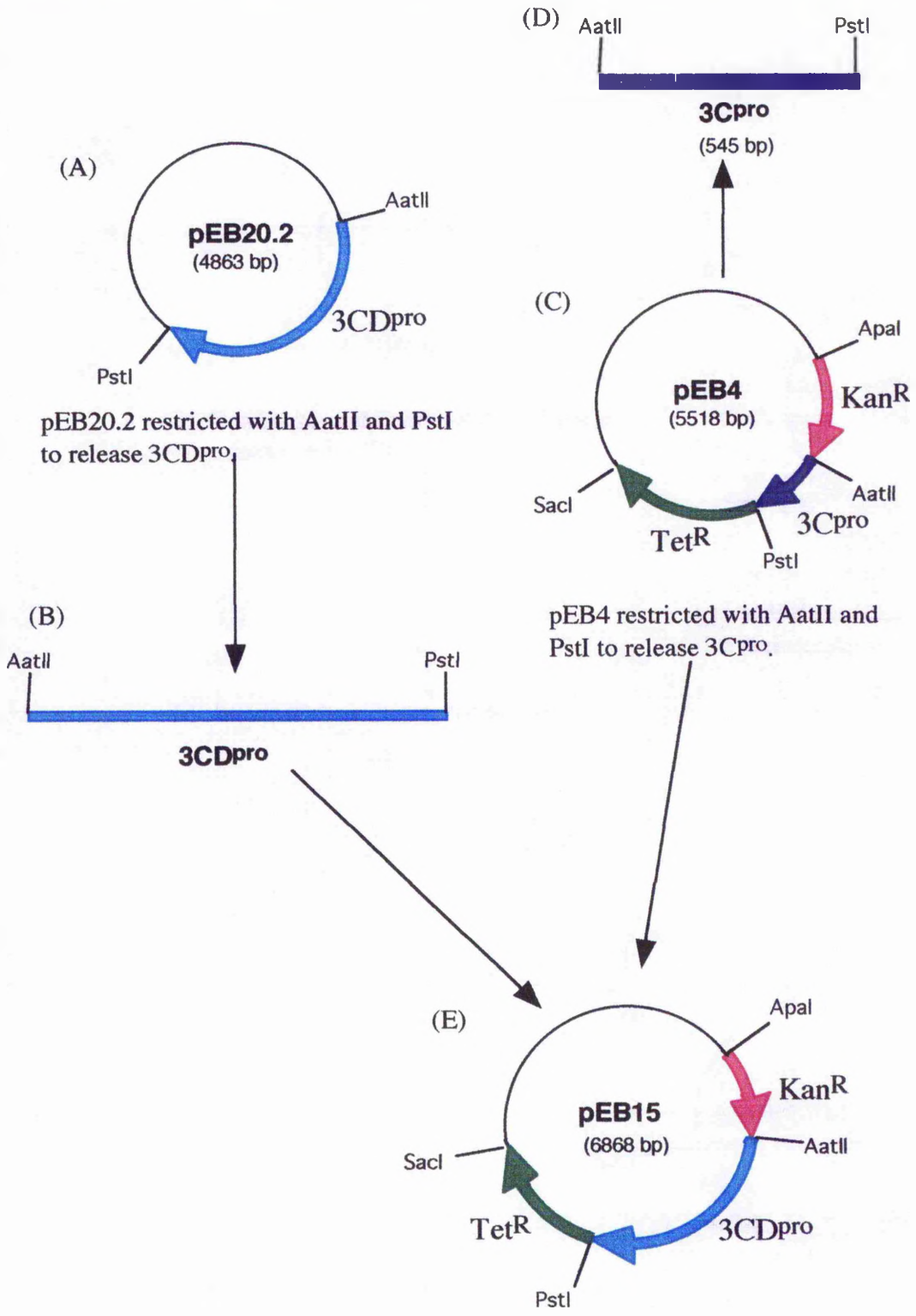
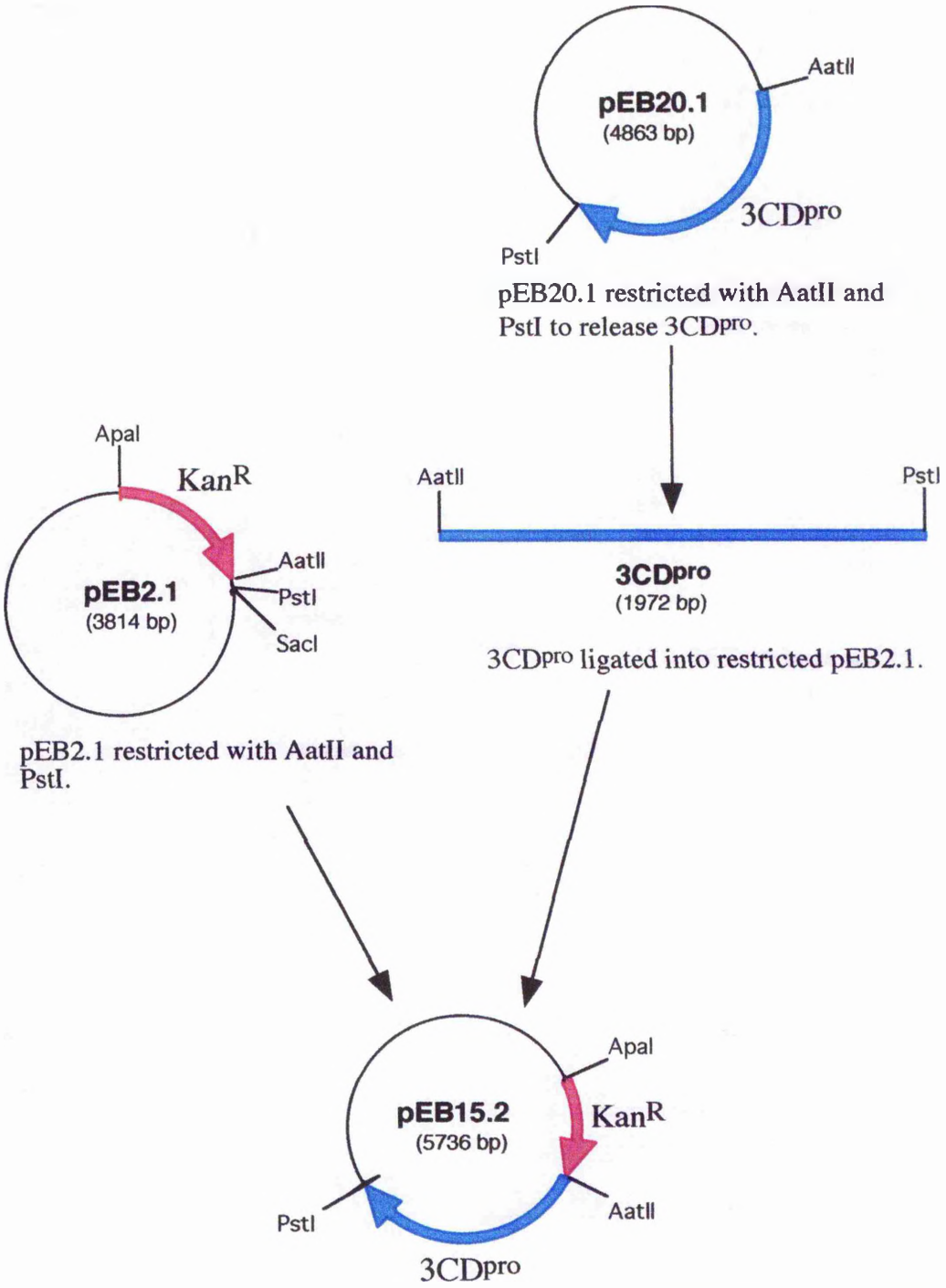


Figure 4.1.6 Construction of pEB15.2. The plasmid pEB20.1 was restricted with AatII and PstI (A) releasing 3CD^{pro} (B). pEB2.1 was restricted with AatII and PstI (C). The 3CD^{pro} fragment was ligated into pEB2.1, similarly restricted, to form the plasmid pEB15.2 (D)

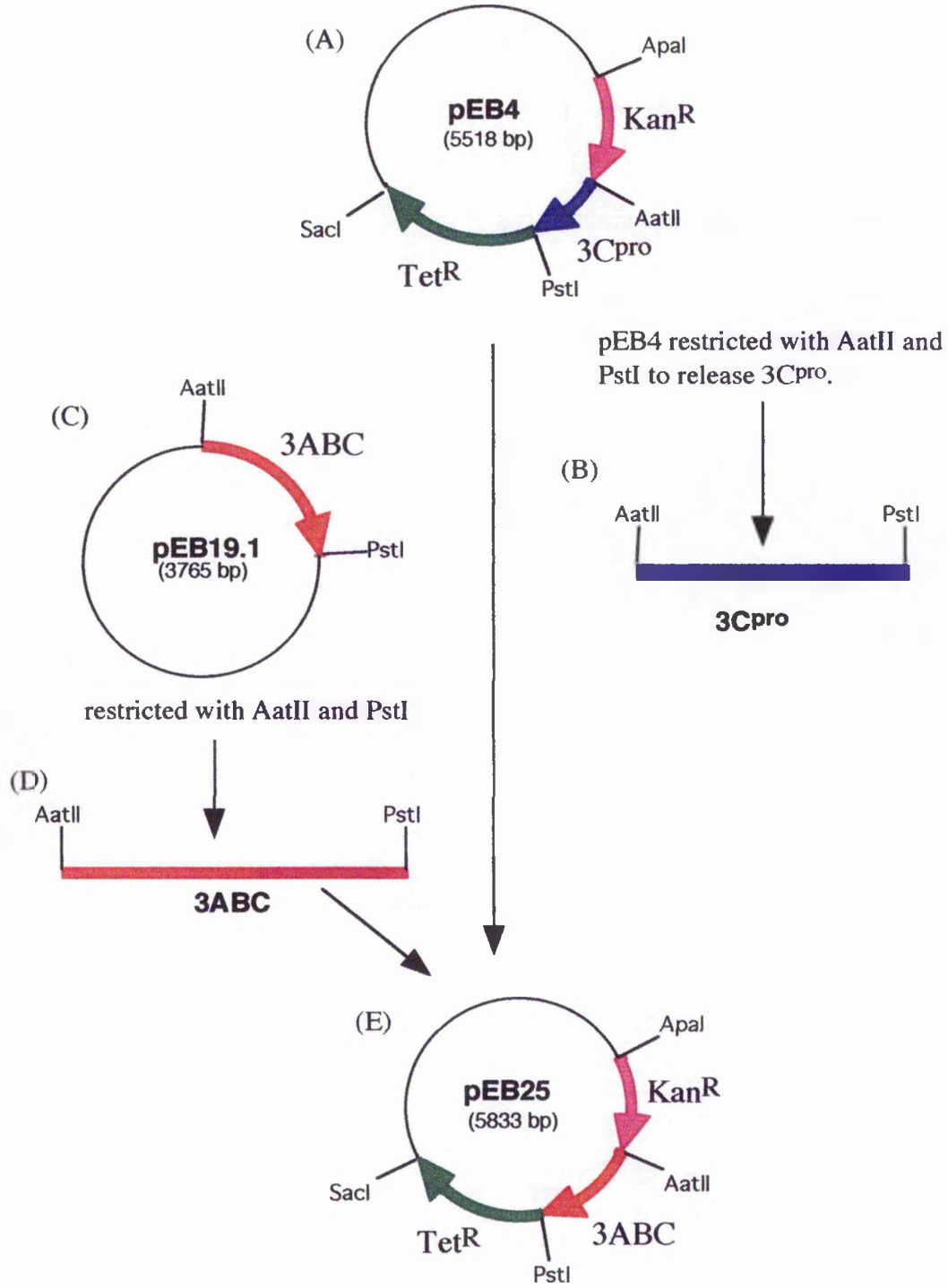


4.1.4 Construction of pEB25.

HRV14 3ABC was released from plasmid pEB19.1 by restriction with AatII and PstI. The purified restriction fragment was then ligated into pEB4 which had also been digested with AatII and PstI releasing $3C^{pro}$ (Figure 4.1.7).

The final plasmid pEB25 [Kan^R3ABCTet^R] will be used to show the influence of upstream viral sequences on cleavage by $3C^{pro}$.

Figure 4.1.7 Construction of pEB25. pEB4 was restricted with AatII and PstI (A) to release $3C^{pro}$ (B). pEB19.1 was restricted with AatII and PstI (C) to release 3ABC (D). The 3ABC fragment was ligated into pEB4, similarly restricted, to form the plasmid pEB25 (E).



4.1.5 Expression of self-processing antibiotic resistance polyproteins in *E. coli*

The reporter system was designed to allow investigation into the analysis of cleavage products by their antibiotic resistance. The two reporter genes, Kan^R and Tet^R, confer resistance to kanamycin and tetracycline respectively. It was intended that cleaved products would express resistance to different combinations of the two antibiotics.

It was known from previous observations that the kanamycin resistance gene product used would not confer resistance with substantial C-terminal extensions.. Therefore cleavage at the N-terminus of $3C^{pro}$ could give kanamycin resistant clones whereas cleavage at the C-terminus of $3C^{pro}$ could give kanamycin sensitive clones (Figure 4.1.8). It was not known whether the Tet^R gene product would show sensitivity to additional sequences at its N- or C-termini.

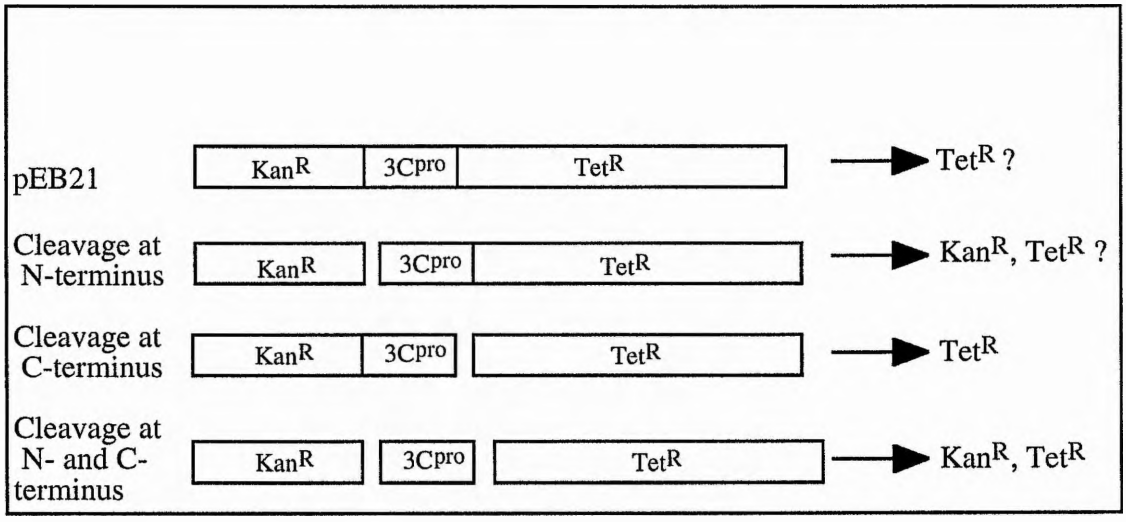


Figure 4.1.8 Predicted antibiotic resistance of pEB4 and its cleavage products.

4.1.5.1 Construction of expression vectors

In order to express antibiotic resistance the genes coding for the artificial reporter polyproteins had to be removed from the transcription vector by restriction digest and ligated into a suitable expression vector. To utilise the unique restriction sites added to the polyproteins the vector pBSIIKS+ was used. Cut fragments were gel purified and ligated into cut pBSIIKS+ overnight at 16 °C with T4 DNA ligase. Putative clones were

checked by restriction digest and a large scale DNA preparation was made of correct clones.

4.1.5.2 Expression of antibiotic resistance in *E.coli*

To monitor expression of antibiotic resistance, reporter gene plasmids were transformed into *E.coli* using the calcium chloride method or the modified Hanahan method. The *E.coli* strain XL1-BLUE was used initially but was found to have an endogenous tetracycline resistance phenotype so the strain JM109 was used. Transformed *E.coli* were plated onto LB agar plates containing appropriate antibiotics. All plates contained ampicillin (at a final concentration of 100 μg/ml) as the expression vector used encoded ampicillin resistance.

When pEB21 and the appropriate control plasmids were grown on LB agar containing antibiotics the results shown in Table 4.1 were obtained. All plasmids were resistant to ampicillin as expected due to the expression vector containing the ampicillin resistance sequences.

		Antibiotic used			
Plasmid	Insert	Amp	Amp + Kan	Amp + Tet	Amp + Kan + Tet
pEB21	[Kan^R3$C^{pro}$$Tet^R$]	√	X	√	X
pEB22	Tet^R	√	X	X	X
pEB23	$3C^{pro}$	√	X	X	X
pEB24	Kan^R	√	√	X	X

Table 4.1 Results of antibiotic expression in *E.coli*. A tick indicates growth of bacterial colonies after 16 hours incubation at 37 °C and a cross indicates no colonies after the same growth period. Ampicillin was used at 100μg/ml, kanamycin at 10 μg/ml and tetracycline at 25 μg/ml.

The control plasmid for kanamycin resistance, pEB24, conferred ampicillin resistance and also conferred kanamycin resistance. It did not express tetracycline resistance. The $3C^{pro}$ control plasmid, pEB23, conferred ampicillin resistance but neither tetracycline resistance nor a kanamycin resistant phenotype. The Tet^R control plasmid, pEB22, conferred ampicillin resistance but failed to confer neither kanamycin resistance

(as expected) nor tetracycline resistance, which was unexpected. The plasmid was sequenced and the nucleotide sequence of the insert was found to be correct. Several attempts were made to rectify the lack of tetracycline resistance in the control plasmid. The tetracycline concentration of the agar plates was varied between 1 μg/ml and 20 μg/ml but this did not affect the lack of tetracycline resistance conferred. The gene was amplified again by the PCR and cloned but this did not result in a plasmid which conferred a tetracycline resistant phenotype. The template used for the PCR amplification of the Tet^R gene, pBR328, was tested for tetracycline resistance and found to be resistant at 10 μg/ml. It was concluded therefore that the gene template was functional and the fault may have been due to the additional sequences added by the PCR or the extraneous sequences in the vector. In pBSIIKS+ the Tet^R gene was cloned as a fusion with β-Gal. This required the addition of IPTG to induce the promoter. The experiment was carried out with IPTG on the agar plates and the induction should have taken place. It therefore seems likely that the lack of function was due to the extraneous β-Gal sequences at the N-terminus. Due to time constraints it was not possible to continue this investigation into the non-functional gene.

pEB21 conferred ampicillin resistance and tetracycline resistance. There are three possible explanations for this. Since it was known that the kanamycin resistance gene will not confer resistance with an C-terminal extension in the coding sequence it was possible that $3C^{pro}$ had cleaved C-terminally in the artificial polyprotein. It may also be that there was no cleavage and the tetracycline resistance gene will function with an N-terminal addition. The third possibility is that additional sequences had inhibited the action of the kanamycin gene product.

Further work will allow this system to be analysed more fully but the expression of tetracycline resistance by pEB21 looks promising. The system may be able to be used for indicating the site (N- or C-terminal) of cleavage.

The coding sequences of the other reporter plasmids (pEB15 and pEB25) were not transferred into the expression vector as the inconclusive results from the expression of pEB21 did not seem to warrant the additional work. Once the problems with the

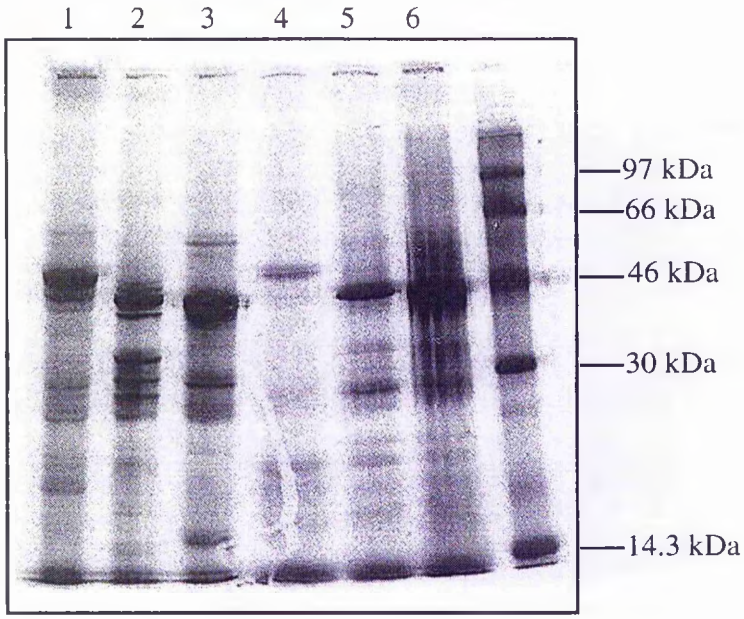
tetracycline expression have been overcome it may be useful to put [Kan^R3CD$^{pro}Tet^R$] and [Kan^R3ABCTet^R] into expression vectors to test the system further.

4.2 Comparison of translation in rabbit reticulocyte lysate and wheatgerm extract transcription translation coupled systems.

In order to determine which system would be best for the transcription of the reporter system, plasmid DNAs were used to program reactions from each TnT system. Equal amounts of the plasmid DNAs (500 ng) were used in each reaction and they were incubated for equal amounts of time (90 minutes). Reactions were stopped by the addition of SDS-PAGE loading buffer and equal volumes were loaded onto 10% SDS-PAGE gels (Figure 4.2.1). As can be seen from the PAGE analysis the distribution of radiolabel is the same in both rabbit reticulocyte lysate (RRL) and wheatgerm extract (WGE) reactions.

Further reactions were carried out using the WGE system as this is more efficient for translating linearised constructs. The linearised reporter constructs provided controls for the system and so the translation of these was important in all subsequent experiments to aid the identification of translation products.

Figure 4.2.1 Comparison of translation in rabbit reticulocyte lysate (RRL) and wheatgerm extract (WGE) TnT systems. Equal amounts of plasmid DNAs were used to program RRL or WGE reactions to compare processing in each system. Results were analysed by SDS-PAGE.



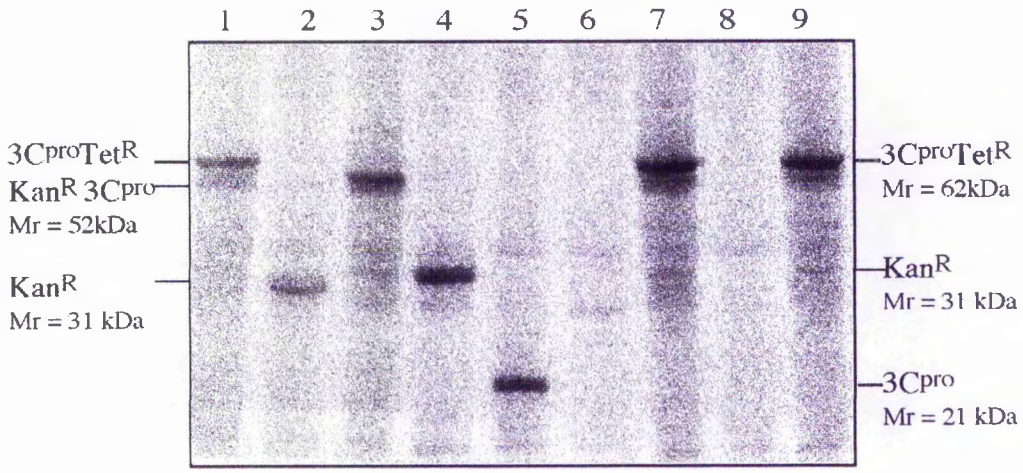
Lane	Plasmid	TnT system used
1	pEB4	WGE
2	pEB15	WGE
3	pEB25	WGE
4	pEB4	RRL
5	pEB15	RRL
6	pEB25	RRL

4.3 Translation *in vitro* of pEB4.

Translation of the plasmid pEB4, encoding [$Kan^R 3C^{pro} Tet^R$], in a wheatgerm extract coupled transcription translation system showed the major translation products to be [$3C^{pro}Tet^R$] and Kan^R (Figure 4.3 a). The identity of the band representing [$3C^{pro}Tet^R$] was identified by comparison with the product of plasmid pEB5 and the use of size standards. This indicated that $3C^{pro}$ cleaved preferentially at its N-terminus in this system. A band representing the translation product of the kanamycin gene in pEB4 was also observed which represents the other portion of the processed polyprotein. No uncleaved polyprotein is detected amongst the translation products. The uncleaved polyprotein has a predicted molecular weight of 93 kDa and the largest protein observed by SDS-PAGE analysis had a molecular weight of approximately 60 kDa (Figure 4.3.1). This implied that it was the cleavage product [$3C^{pro}Tet^R$] which migrated to the same position as the translation product of the plasmid pEB5, [$3C^{pro}Tet^R$], which has a predicted molecular weight of 62 kDa. The absence of a band migrating at ~93 kDa on the gel indicated that $3C^{pro}$ cleaves the polyprotein entirely (at one site or another).

Some other cleavage products are seen apart from [$3C^{pro}Tet^R$] and [Kan^R] which implies that cleavage was also taking place at the C-terminus of $3C^{pro}$ but at lower levels. These products were produced in much smaller quantities than the N-terminal cleavage products so $3C^{pro}$ appeared to cleave preferentially at its own N-terminus in the artificial polyprotein system used here.

Figure 4.3 a Transcription and translation of reporter plasmid pEB4. WGE TnT reactions were programmed with the plasmid DNAs indicated . Samples of reactions were analysed by SDS-PAGE. Gels were exposed to X-ray film overnight.



Lane	Plasmid	Insert	Linearised by restriction with
1	pEB4	[KanR3CproTetR]	
9	pEB4	[KanR3CproTetR]	
2	pEB4	[KanR3CproTetR]	x AatII [1]
3	pEB4	[KanR3CproTetR]	x PstI [2]
4	pEB2	[KanR]	
5	pEB3	HRV14[3Cpro]	
6	pEB1	[TetR]	
7	pEB5	[3CproTetR]	
8	pEB14	[KanR3Cpro]	

1. AatII linearises pEB4 between the 3' end of the KanR gene and the 5' end of the 3Cpro gene.
2. PstI linearises pEB4 between the 3' end of the 3Cpro gene and the 5' end of the TetR gene.

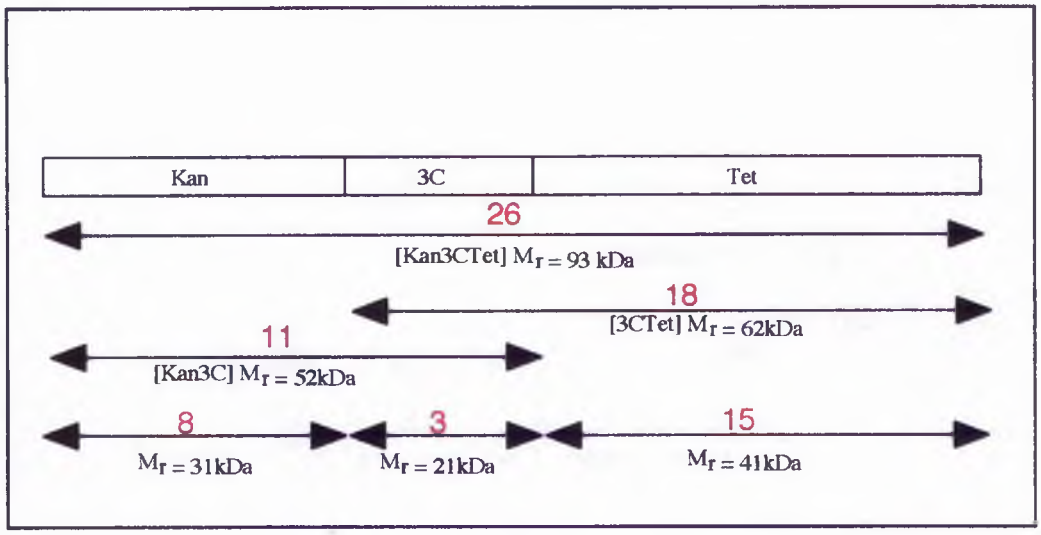


Figure 4.3.1 Molecular weights of predicted translation products of pEB4. Numbers in red show the number of methionine codons in each protein.

There are slight size discrepancies between the control plasmids and the sizes of the translation products from pEB4, due to the strategy used to maintain the $3C^{pro}$ cleavage sites in the reporter polyprotein system (Figure 4.3.3).

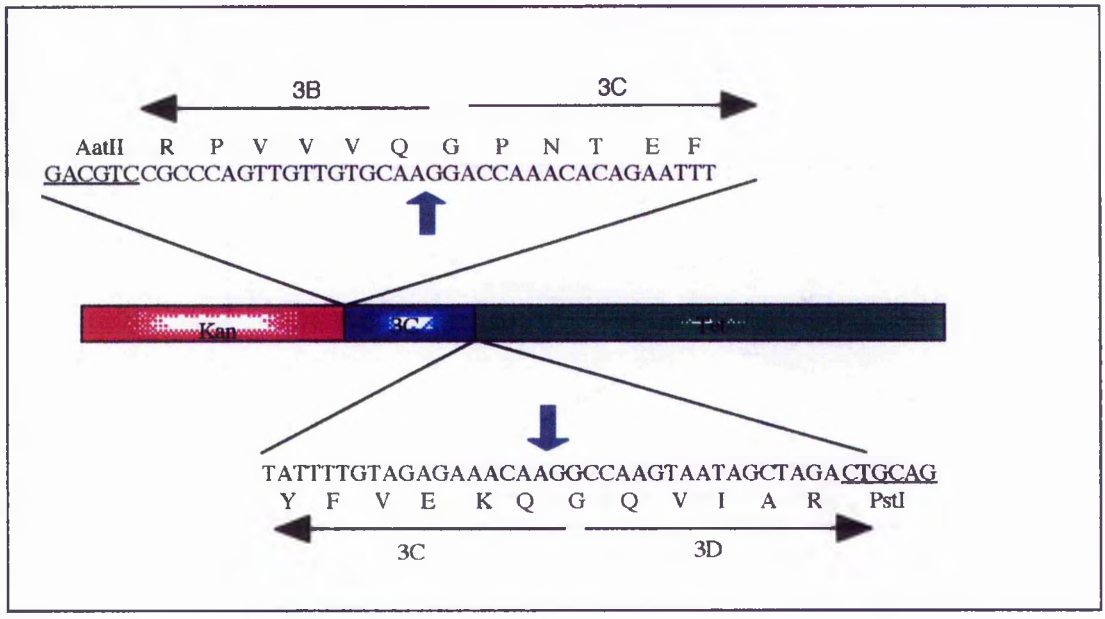


Figure 4.3.2 Diagram to show the discrepancy between the enzyme cleavage sites and $3C^{pro}$ proteolytic cleavage sites.

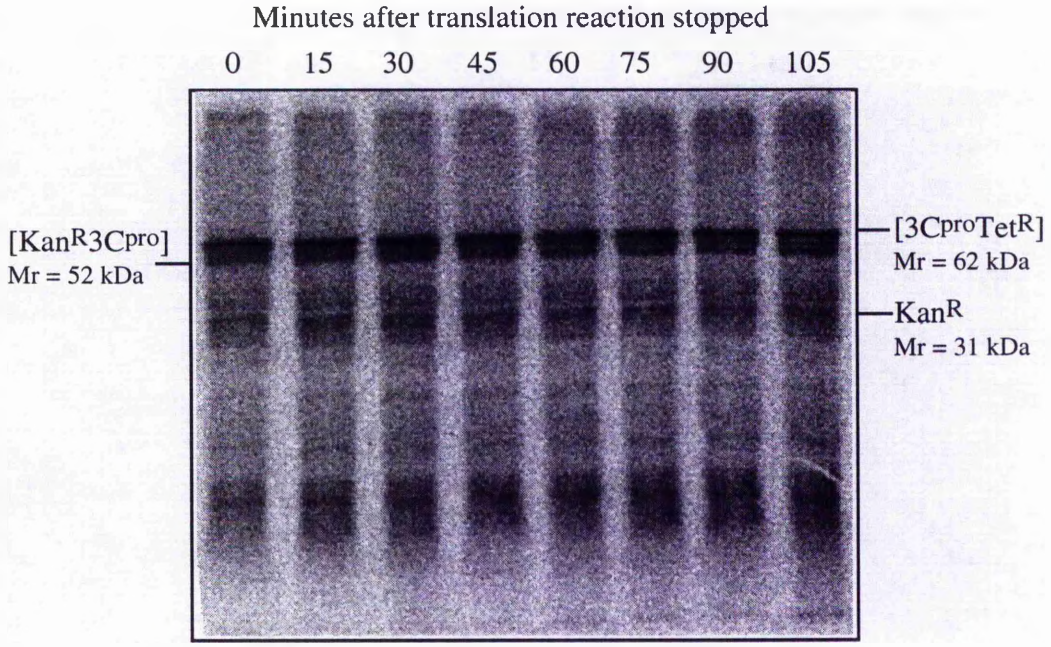
4.3.1 Densitometric analysis of the degradation of the cleavage products of pEB4.

Inspection of the gel revealed the band representing Kan^R to be of low intensity compared to that of [$3C^{pro}Tet^R$]. Initially this was thought to be due to the degradation of the translation product by proteases within the translation system. In order to clarify this point, the stability of the translation products was monitored over a period of time. The reaction was incubated at 30°C for 45 minutes and protein synthesis stopped by the addition of cycloheximide to a final concentration of 0.8 μg/ml and mRNA degraded by the addition of RNaseA (added to a final concentration of 0.6 μg/ml). Incubation was continued at 30 °C and samples were taken at regular intervals. Samples were added to an equal volume of SDS-PAGE loading buffer to stop any further degradation. The results are shown in Figure 4.3.3. It can be seen that the intensity of the kanamycin band appeared to be constant throughout the reaction period and no significant degradation occurred. In confirmation densitometric analysis was carried out on the SDS-PAGE gel (Table 4.2).

The number of methionines in the kanamycin resistance gene (8) is much lower than that of [$3C^{pro}Tet^R$] (18) and therefore the level of incorporation of ^{35}S-Met is much lower. Densitometric analysis was carried out using Mac-Bas version 2 software on an image generated by phosphorimaging. Each lane was analysed by searching for peaks within the band and the intensity measured in PSL units. A value was also obtained for the intensity of the bands with the background count subtracted. These values were then corrected for the number of methionines in the protein and a ratio calculated of [$3C^{pro}Tet^R$] : [Kan^R]. The results of the analysis can be seen in Table 4.2.

The corrected values for the Kan^R bands in each lane did not show a decrease indicating that there was little or no degradation of [Kan^R]. The ratio of [$3C^{pro}Tet^R$] to Kan^R varied between 0.58 and 1.05 but there was no trend seen in the data. This indicates that any degradation of the Kan^R band is not occurring until 90 minutes after protein production has been stopped and the lower intensity of the Kan^R band is due to the lower number of methionines not degradation. It should also be noted that the KanR product migrates in a very diffuse manner in this system which may account for the low intensity observed.

Figure 4.3.3 Degradation of pEB4. A TnT reaction was programmed with pEB4 and incubated for 45 minutes at 30°C. The reaction was stopped by the addition of cycloheximide and RNaseA and subsequently incubated at 30°C for varying periods of timeas shown. Translation reactions were stopped by the addition of loading buffer and aliquots were then loaded onto an SDS-PAGE gel.



The amount of the kanamycin resistance gene produced from the control plasmid pEB1 was, however, a great deal higher as the translation product is a lot shorter than the polyprotein in pEB4. Both of the Kan^R genes were cloned in the same context, *i.e.* they have the same additional sequences and initiation codon so the difference in observed intensity is surprising. The Kan^R product derived by processing of pEB4 migrates in a much more diffuse manner than that of the discrete gene product from pEB1. The incorporation level of methionine would be the same as for the [Kan^R] band in pEB4 but the number of protein molecules produced may be greater due to the smaller protein product being translated more rapidly due to its size.

Lane	Band	PSL	PSL-BG	No. of mets	PSL-BG/ no. of mets	Ratio of 3CTet : Kan
1	$3C^{pro}Tet^R$	1216.0	789.2	18	43.8	0.66
	Kan^R	1060.0	525.6	8	65.7	
2	$3C^{pro}Tet^R$	1153.0	1153	18	64.05	0.91
	Kan^R	561.4	561.4	8	65.7	
3	$3C^{pro}Tet^R$	1061.0	1061	18	58.9	0.88
	Kan^R	534.9	534.9	8	66.9	
4	$3C^{pro}Tet^R$	1059.0	1059	18	58.8	0.96
	Kan^R	492.0	492	8	61.5	
5	$3C^{pro}Tet^R$	1123.0	1123	18	62.4	1.05
	Kan^R	476.8	476.8	8	59.6	
6	$3C^{pro}Tet^R$	1061.0	1061	18	58.9	0.96
	Kan^R	488.6	488.6	8	61.05	
7	$3C^{pro}Tet^R$	1129.0	1129	18	62.7	0.65
	Kan^R	762.2	762.2	8	95.2	
8	$3C^{pro}Tet^R$	851.2	851.2	18	47.3	0.58
	Kan^R	654.1	654.1	8	81.7	

Table 4.2 Results of densitometric analysis of degradation of pEB4. The PSL- BG (background) figures were divided by the number of methionines in each protein product to give a figure proportional to the number of moles of protein within each protein band. These figures were then used to calculate a ratio for the amount of [$3C^{pro}Tet^R$] compared to that of Kan^R.

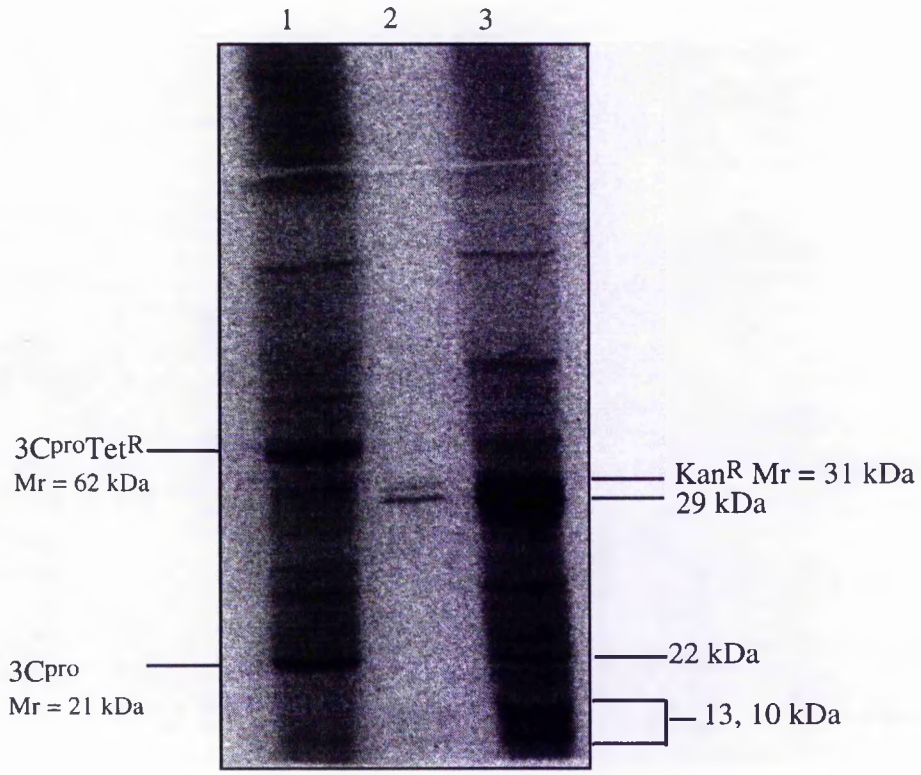
4.3.2 Identification of translation products of the Kan^R gene in pEB4 by immunoprecipitation.

Despite the results of the densitometric analysis on the translation products confirming that there was no degradation of Kan^R further proof of the identity of these band was felt to be necessary. An immunoprecipitation reaction was carried out on the translation products of pEB4 (Figure 4.3.4). No antibodies were available for the identification of the Tet^R products or HRV14 3C^{pro}. The antibody used was supplied by 5Prime ->3Prime and is an antibody to the neomycin phosphotransferase gene NPTII, recognising the kanamycin resistance gene product. In order to clarify the identification of the Kan^R translation product of pEB4 it was immunoprecipitated with this antibody. The translation reaction was carried out with pEB4, digested with AatII to give the kanamycin resistance gene translation product. The antibody identified two translation products from pEB4. These were thought to be the product of the kanamycin resistance gene and a product from internal initiation within the gene.

Protein synthesis begins when the ribosome scanning along the mRNA finds the first AUG start codon. Sequences flanking the AUG start codon influence its recognition by eukaryotic ribosomes. Optimal initiation is found when the AUG is surrounded by a consensus sequence known as the Kozak consensus sequence (Kozak, 1986). The consensus sequence, G C C G C C **A/G** C C **A U G** **G**, also has within it nucleotides which are very important for recognition (shown in bold). Some ribosomes may therefore scan along the mRNA and through the first AUG to initiate at another AUG downstream producing internal initiation products.

The nucleotide sequence of the kanamycin resistance gene was scanned for internal initiation sites and their Kozak consensus sequences checked. The full length translation product of the kanamycin gene has a relative molecular weight of 31 kDa and there are seven other methionine codons which are in frame with the initiation codon for the full length product. The immunoprecipitation reaction seen in Figure 4.3.4 precipitated two proteins, which were identified as the kanamycin protein from the initial methionine and an internal initiation product from Met-19 with a relative molecular weight of 29 kDa.

Figure 4.3.4 Identification of translation products of the Kan^R gene in pEB4 by immunoprecipitation. The plasmid pEB4 was digested with AatII and then used to program a WGE TnT reaction. The products of this were immunoprecipitated with NPTII and the resulting products analysed by SDS-PAGE.
The two bands in Lane 2 are the full length translation product of the kanamycin resistance gene (MW = 31 kDa) and an internal initaition product with a molecular weight of 29 kDa.



Lane 1 pEB4

Lane 2 Immunoprecipitation of [pEB4 x AatII]

Lane 3 pEB4 x AatII

Kozak consensus sequences and predicted molecular weights for all of the methionine codons which are in frame can be seen below (Figure 4.3.5).

Kozak consensus	GCC	GCC	A/GCC	AUG	G	Predicted Molecular weight
MET 1		GGG	CCC	AUG	G	M_r = 31 kDa
MET 19	AAT	TCC	AAC	AUG	G	Mr = 29 kDa
MET 69	ACA	GAT	GAG	AUG	G	Mr = 23.5 kDa
MET 79	ACG	GAA	TTT	AUG	C	Mr = 22 kDa
MET 156	CAA	TCA	CGA	AUG	A	Mr = 13 kDa
MET 182	TGG	AAA	GAA	AUG	C	Mr = 10 kDa
MET 258	AAT	CCT	GAT	AUG	A	Mr = 2 kDa
MET 266	TTT	CAT	TTG	AUG	C	Mr = 0.5 kDa

Figure 4.3.5 Kozak consensus sequences and predicted molecular weights for internal initiation products of the kanamycin resistance gene.

The consensus sequence for Met-19 is stronger than that for the wild type initiation codon which may explain the higher observed intensity of the band for this product. Densitometric analysis of this indicates that initiation is greater from Met-19 than from Met-1. The insertion of a restriction site at the 5' end of the gene has changed the kozak sequence and this has probably led to the lesser degree of initiation. The optimisation of the consensus initiation sequence during the construct design process would have eliminated the problem of low levels of initiation at the initial methionine and aided the interpretation of results by giving better translation products from this codon.

As all TnT reactions were carried out in the wheatgerm extract TnT system the context sequence for translation initiation has been found to be slightly different from that of vertebrates. The context for monocot mRNAs is c(a/c)(A/G)(A/C)c AUG GCG (Joshi et al., 1997) which exhibits overall similarity to the vertebrate consensus GCC GCC A/GCC AUG G.

4.3.3 Effect of C-terminal sequence on cleavage activity of $3C^{pro}$.

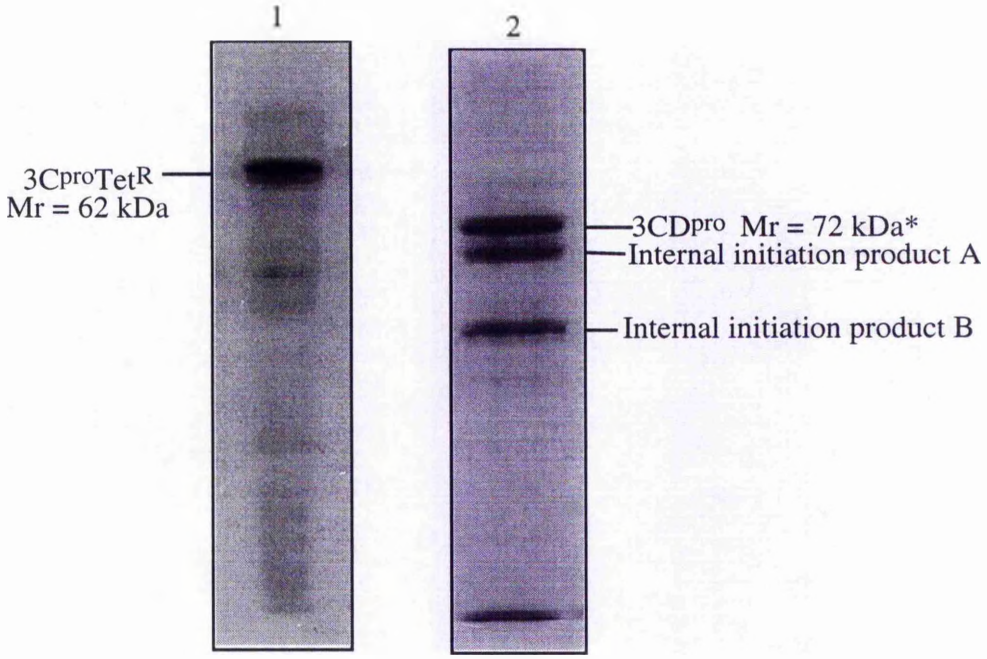
Translation of the plasmid pEB4 [$Kan^R3C^{pro}Tet^R$] shows that $3C^{pro}$ cleaves both N- and C-terminally in the reporter polyprotein system. When *only* an upstream sequence was present (as seen with [Kan^R3C^{pro}]; Figure 4.3 a Lane 3) N-terminal cleavage does not seem to be very efficient, with [Kan^R3C^{pro}] being the major translation product. Low levels of Kan^R and $3C^{pro}$ were, however, observed. The effect of the C-terminal sequence is apparent when there are both N- and C-terminal sequences present in the reporter system. The presence of Tet^R seems to greatly enhance cleavage at the N-terminus of $3C^{pro}$ giving rise to cleavage products of Kan^R and [$3C^{pro}Tet^R$]. If $3C^{pro}$ has C-terminal extension, only, (as in [$3C^{pro}Tet^R$]; Figure 4.3.a Lane 7), then cleavage seemed to be at very low levels.

Comparison between these two constructs is purely qualitative as the difference between translation products derived from enzyme restriction of plasmids and proteolytically cleaved polyproteins does not allow quantitative analysis. The same applies to comparison of cleavage in $3CD^{pro}$ with cleavage in [$3C^{pro}Tet^R$]. The cleavage in $3CD^{pro}$ occurs in the virus but not in this system. Cleavage of [$3C^{pro}Tet^R$] is not an "artificial" cleavage in the sense that the $3C^{pro}$ Q-G cleavage site has been maintained.

The effect of a C-terminal sequence on cleavage by $3C^{pro}$ seemed to be the enhancement of cleavage at the N-terminus and not the C-terminus of $3C^{pro}$. The nature of the sequence beyond the cleavage site did not seem to be important. Both $3D^{pol}$ and Tet^R produced the same enhancing effect.

Figure 4.3.6 Effect of C-terminal sequence on cleavage activity by $3C^{pro}$. These tracks (taken from different gels) show the different cleavage patterns of $3C^{pro}$ with C-terminal sequences. It can be seen from Lane 1 that [$3C^{pro}Tet^{R}$] is a stable product when derived from the plasmid pEB5. (NOTE: The same occured in translation of pEB4 [$Kan^{R}3C^{pro}Tet^{R}$]).

Lane 2 shows translation of $3CD^{pro}$ which similarly did not cleave to give $3C^{pro}$ and $3D^{pol}$.



Lane 1 pEB5 [$3C^{pro}Tet^{R}$] in pGEM5zf(+)
Lane 2 pEB7 [$3CD^{pro}$] in pRSETA

*also includes the internal initiation product Δ3CDMet-17 (Mr = 70.2 kDa
Internal initiation product A = Δ3CD Met-121 (Mr = 59 kDa)
+Δ3CD Met-135 (Mr = 57.5 kDa)
Internal initiation product B = Δ3CDMet-324 (Mr = 36.7 kDa)

4.3.4 Conclusions

The summary diagram below (Figure 4.3.7) shows the cleavage patterns of $3C^{pro}$ in a polyprotein system flanked by two reporter proteins. $3C^{pro}$ processes efficiently with both N- and C-terminal extensions. In order to cleave efficiently at its N-terminus $3C^{pro}$ needs a C-terminal extension whether this is viral ($3D^{pol}$) or non-viral (Tet^R) in origin does not seem to be important. This can be seen by the different processing patterns in pEB4 and pEB4 when digested with PstI prior to translation. N-terminal cleavage is very efficient when $3C^{pro}$ has a C-terminal extension of Tet^R but when this is removed by restriction digest prior to translation cleavage is much less efficient.

The sequence C-terminal to $3C^{pro}$ did not have an effect on the efficiency of cleavage at the C-terminus of $3C^{pro}$. pEB5 [$3C^{pro}Tet^R$] did not cleave at all in this form or as a product of pEB4 translation. $3CD^{pro}$ did not cleave to give $3C^{pro}$ and $3D^{pol}$ which was consistent with previous work (see section 3.10). The effect of the identity of the C-terminal sequence on C-terminal cleavage did not, therefore, seem to be important.

Antibiotic resistance studies (section 4.1.5), although inconclusive in their own right, appeared to contradict results from the TnT experiments. The major cleavage seen in pEB4 is at the N-terminus of $3C^{pro}$ yielding Kan^R and [$3C^{pro}Tet^R$]. Antibiotic resistance was observed to tetracycline but not to kanamycin. As it is known that the kanamycin resistance gene will not function with an N-terminal extension, $3C^{pro}$ must either be cleaving C-terminally or not cleaving at all. However, the greater amount of *trans* cleavage found to occur in virally infected cells may explain the discrepancy between the results. In the *in vitro* system used small quantities of protein are being produced and even with long incubation periods *trans* processing is rarely seen. The larger amounts of protein produced in cells allow *trans* processing to occur more frequently and this may be the reason why the C-terminal cleavage seems to be occuring preferentially in the bacterial system rather than the N-terminal cleavage which is seen in the *in vitro* system.

The TnT results show cleavage at both N and C-termini of $3C^{pro}$ although it does not appear to be equal. The preferential cleavage is at the N-terminus to give Kan^R and

[$3C^{pro}Tet^{R}$] which would be expected to give resistance to kanamycin and possibly to tetracycline as well in the antibiotic expression system. The antibiotic resistance experiments could infer that cleavage is occurring at the C-terminus but there are other possible explanations for this (see section 4.1.5).

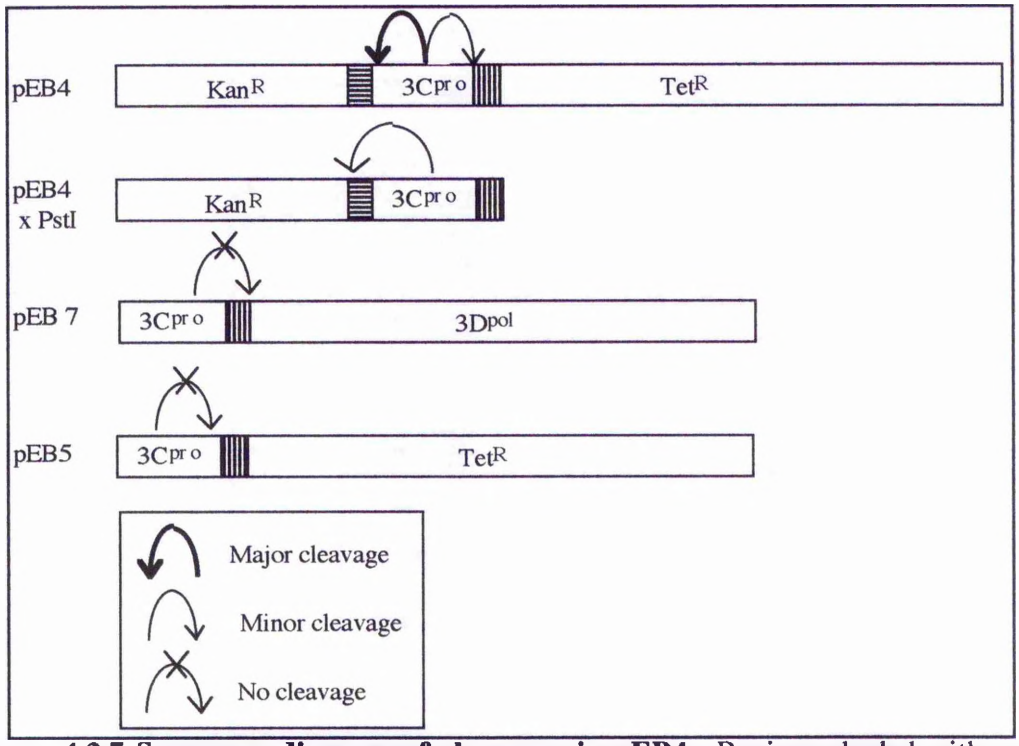


Figure 4.3.7 Summary diagram of cleavages in pEB4. Regions shaded with vertical lines are the C-terminal sequences from 3B and regions shaded with horizontal lines are the N-terminal sequences from 3D.

Translation of the plasmid pEB4.2 [$Kan^{R}3C^{proC146A}Tet^{R}$] resulted in a single band (Figure 4.1.4.1), however, this appears to be of approximately the same size as the translation product of pEB5 [$3C^{pro}Tet^{R}$], 62 kDa. The mutation in the protease sequence should destroy the proteolytic activity of the enzyme resulting in translation of the whole polyprotein [$Kan^{R}3C^{pro}Tet^{R}$], 93 kDa. Due to time constraints this was not investigated further. The discrepancy may be due to inadequate migration on the SDS-PAGE gel. If migration was allowed for a longer period of time a size difference may be seen. The sequence of the mutant protease was correct when sequenced.

Summary

- A reporter polyprotein system has been produced which allows cleavage by HRV14 $3C^{pro}$ *in vitro* without the influence of other viral gene sequences.
- $3C^{pro}$ will cleave both N-and C-terminally in the presence of reporter proteins.
- In this system cleavage occurs preferentially at the N-terminus of $3C^{pro}$ but there is some C-terminal cleavage.
- The presence of a C-terminal sequence is required for efficient N-terminal cleavage.
- The identity of the C-terminal sequence is not important. Cleavage is not more efficient with $3D^{pol}$ than Tet^{R}.

4.4 *In vitro* translation of pEB15 and pEB15.2.

Incubation of the constructs pEB15 [Kan^{R}3CD^{pro}Tet^R] and pEB15.2 [Kan^{R}3CD^{pro}] in a coupled transcription /translation system gave an indication of the cleavage activity of 3C^{pro} when other viral sequences (3AB, 3D^{pol}) were present. pEB15 [Kan^{R}3CD^{pro}Tet^R] shows the proteolytic activity of 3C^{pro} in the Kan/Tet reporter system when 3D^{pol} is also present (Fig 4.4.a). pEB15 was used to program a WGE TnT reaction and translation products were analysed by SDS-PAGE. Products observed were Kan^R, 3C^{pro}, 3D^{pol}, [Kan^{R}3C^{pro}] and surprisingly Tet^R. Cleavage occurs at both the N-and C-termini of 3C^{pro} to give the products Kan^R, 3C^{pro} and 3D^{pol} and Tet^R. Cleavage at the C-terminus of 3C^{pro} yields [Kan^{R}3C^{pro}] and the remaining portion of the polyprotein is cleaved at the C-terminus of 3D^{pol} to yield 3D^{pol} and Tet^R. No [3D^{pol}Tet^R] is observed in the PAGE analysis indicating that cleavage at the C-terminus of 3D^{pol} is very efficient. The predicted molecular weights for these cleavage products can be seen below (Figure 4.4.1).

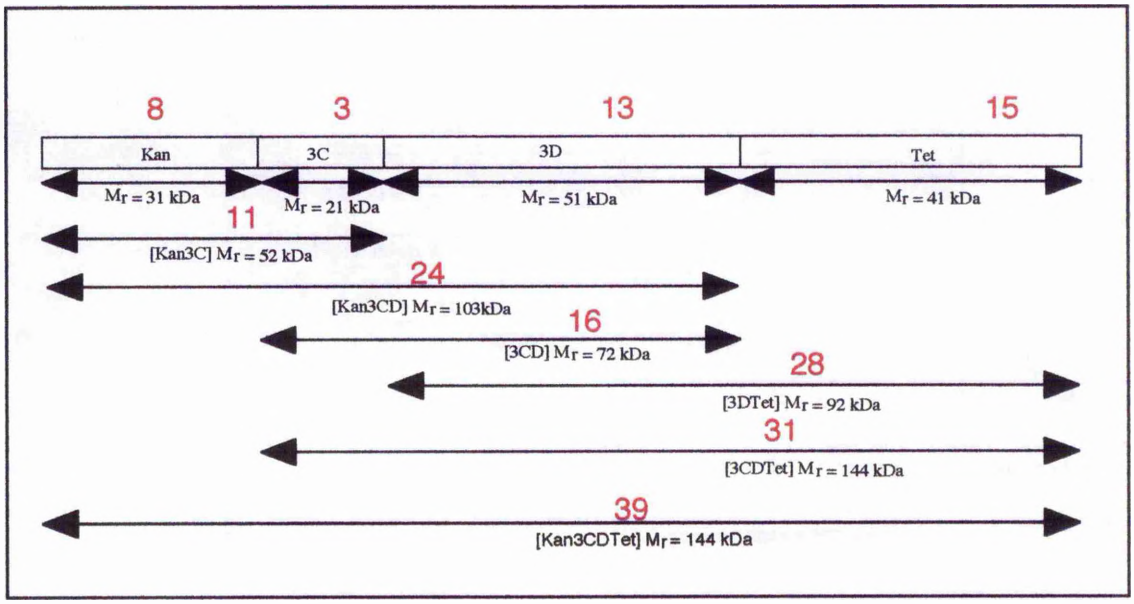
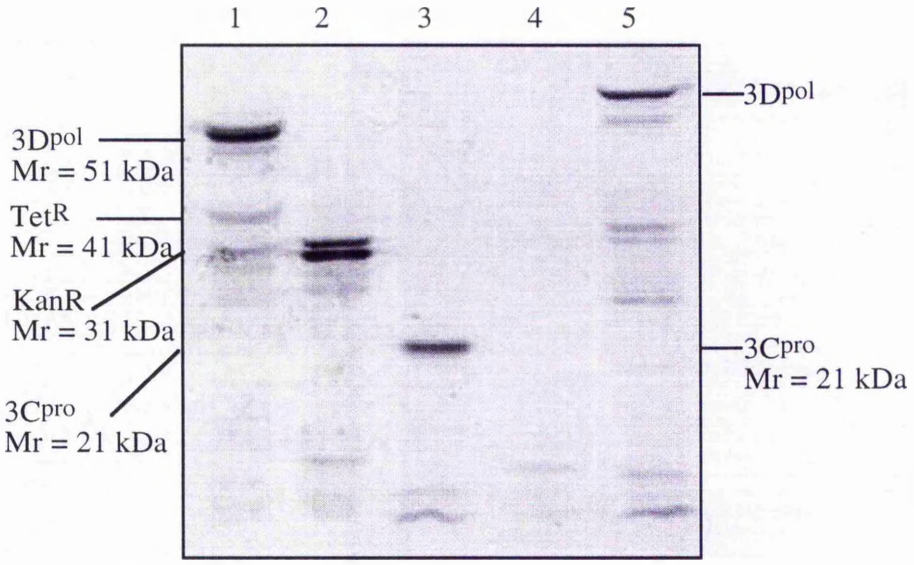


Figure 4.4.1 Molecular weights of predicted cleavage products of pEB15. The numbers in red are the number of methionine codons in each protein.

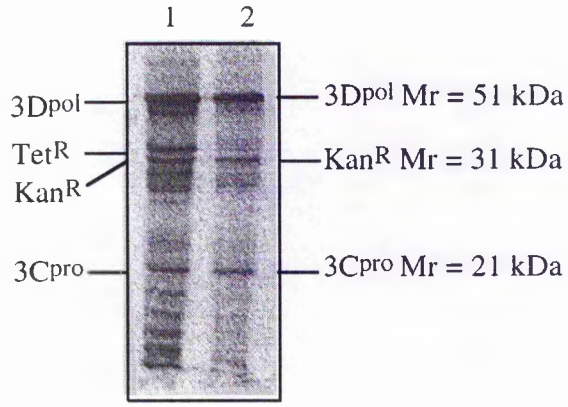
Figure 4.4 a Translation of pEB15. Plasmid DNA was incubated in a WGE TnT reaction for 90 minutes at 30 °C. Samples of the reaction were then analysed by SDS-PAGE.



Lane	Plasmid	Insert
1	pEB15	[Kan^{R}3CD^{pro} Tet^R]
2	pEB2	[Kan^R]
3	pEB3	[3C^{pro}]
4	pEB7	[3CD^{pro}]
5	pEB27	[3D^{pol}]

NOTE: There is a size discrepancy between the 3D^{pol} gene product in Lanes 1 and 5. The identity of the 3D^{pol} band in the translation of pEB15 (Lane 1) is confirmed by comparison with pEB15 x PstI (Figure 4.4b). The plasmid pEB27 needs to have the identity of the insert conformed by DNA sequencing.

Figure 4.4 b Longer exposure of translation of pEB15. Samples of translation products of pEB15 and pEB15 x PstI (the plasmid was digested prior to incubation in the TnT sytem) were analysed by SDS-PAGE. The digested plasmid was used to provide a size control for the translation products of the reporter plasmid. The gel was exposed to X-ray film for 48 hours.



Lane 1 pEB15 [Kan^R3CD^{pro}Tet^R]

Lane 2 pEB15x PstI

Cleavage is seen at the C-terminus of $3D^{pol}$ which releases Tet^R, this is not a cleavage seen in the virus. The plasmid was digested at the restriction site at the 3' end of $3CD^{pro}$ prior to translation to provide a control for the size of [Kan^R3CD^{pro}] and its cleavage products (Fig 4.4 a). This produced bands corresponding to Kan^R, $3C^{pro}$ and $3D^{pol}$. There was no uncleaved [Kan^R3CD^{pro}] or [$3CD^{pro}$] seen. $3C^{pro}$ cleaves the polyprotein totally in this situation.

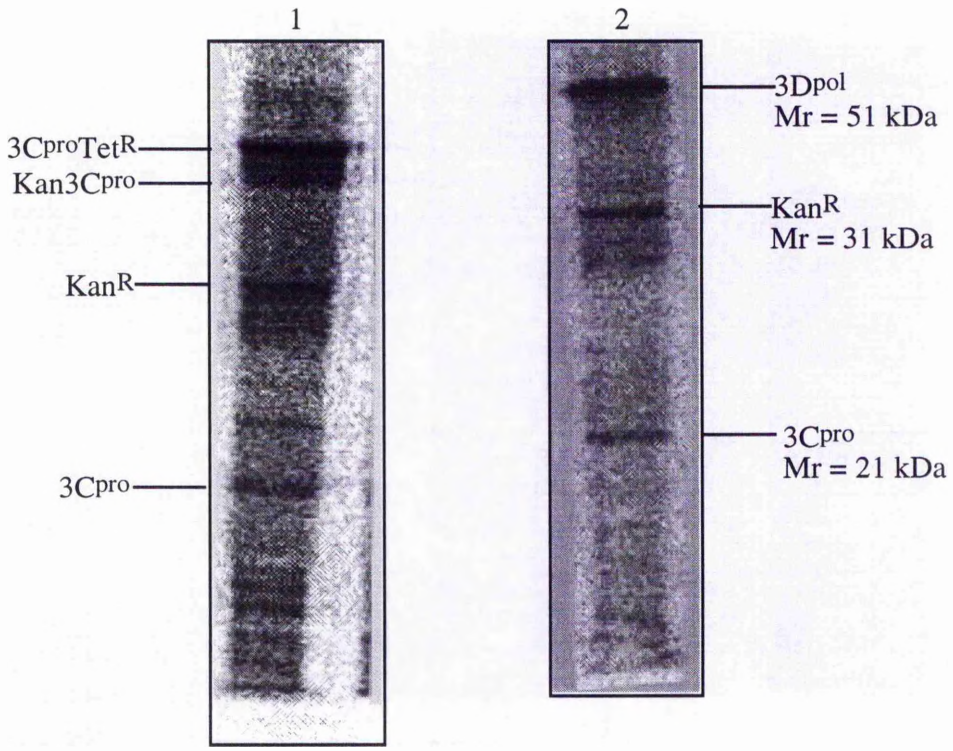
With $3CD^{pro}$ present in the reporter system $3C^{pro}$ processes at all available sites within the polyprotein. The analysis showed there to be equal processing N- and C-terminally to give all products in equimolar amounts. In PstI digested pEB15 [Kan^R3CD^{pro}] (Figure 4.4 b) there is also equal processing at both N- and C-termini. This confirms the conclusion made earlier that $3D^{pol}$ has an enhancing effect on the cleavages made by $3C^{pro}$.

This can be compared with processing in pEB4 (Figure 4.4.2). The plasmid pEB15.2 [Kan^R3CD^{pro}] was translated *in vitro* in a wheatgerm extract coupled transcription translation system to observe the action of $3C^{pro}$ between two genes as a comparison with pEB4. Comparison is more reliable between these two constructs as both artificial polyproteins have been cloned into transcription vectors. The constructs have not been digested with enzymes at sites within the polyprotein to produce these products.

pEB4 yields Kan^R and [$3C^{pro}Tet^R$] as its major cleavage products. This is due to cleavage at the N-terminus of $3C^{pro}$. Cleavage at the C-terminus of $3C^{pro}$ is not as strong and the products of this are only seen in small amounts. Processing in pEB15.2 yields Kan^R, $3C^{pro}$ and $3D^{pol}$ all in approximately equal amounts. Cleavage is occurring at both ends of $3C^{pro}$ with equal efficiency. This ties in with the observation made in the previous section that $3C^{pro}$ cleaves more efficiently at its C-terminus if it has $3D^{pol}$ downstream.

Some internal initiation products from the Kan^R gene can be seen in the translation of pEB15. The sizes of these are shown in Figure 4.3.4.

Figure 4.4.2 Comparison of cleavage in pEB4[Kan^{R}3CTet^R] and pEB15.2 [Kan^{R}3CD^{pro}]. To compare cleavage when 3C^{pro} has both N- and C-terminal extensions these two lanes can be compared. They are taken from different gels but can still be compared qualitatively.



Lane 1 pEB4 [Kan^{R}3C^{pro}Tet^R] in pGEM5zf(+)

Lane 2 pEB15.2 [Kan^{R}3CD^{pro}] in pGEM5zf(+)

4.4.1 Densitometric analysis of the cleavage products of pEB15

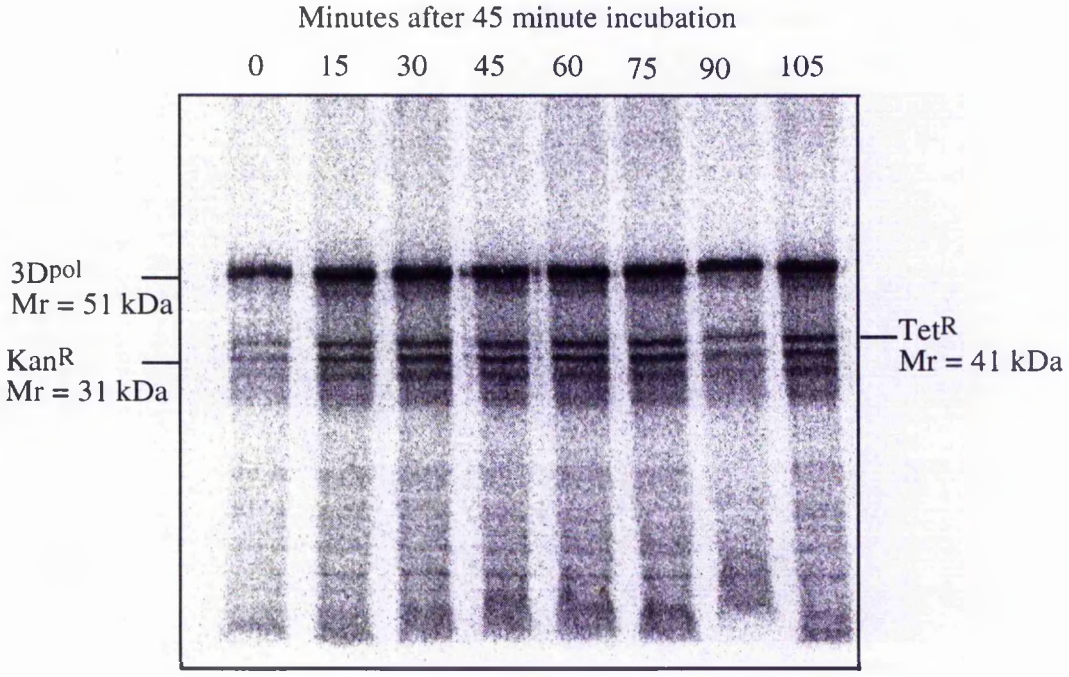
The cleavage of pEB15 and pEB15.2 yields the products [Kan^R3C^{pro}] and $3D^{pol}$, $3C^{pro}$, Kan^R and Tet^R. In pEB4 it was observed that the Kan^R band migrated in a diffuse manner and as of low intensity. This was initially thought to be due to degradation. In order to determine whether this was the case in pEB15 the course of the translation reaction was monitored. The reaction was incubated at 30 °C for 45 minutes and then protein synthesis was stopped by the addition of cycloheximide (to a final concentration of 0.8μg/ml) and RNaseA was added (to a final concentration of 0.6μg/ml) to degrade any mRNA present. The reaction was then incubated at 30 °C and samples taken at regular intervals. The results of this can be seen in Figure 4.4.3.

Densitometric analysis was carried out on the products of the degradation reaction. The SDS-PAGE gel was visualised by phosphorimaging and the bands were quantified using Mac Bas version 2 software. PSL values were obtained and gave a value for the intensity of the bands. A background level was subtracted from the PSL values (PSL-BG) and the values were corrected for the number of methionines in each protein product. As labelling in the reaction is by ^{35}S-Met the number of methionines available in the protein for labelling has an effect on the intensity of the band. PSL values must therefore be corrected for this before they can be compared with each other. Division of PSL-BG by the number of methionines in the protein gives a value which is proportional to the number of moles of each protein product present on the gel. A ratio of the products to each other was also calculated (Table 4.3).

The corrected PSL value for the Kan^R band shows that it is not degraded over the time course monitored in this reaction. The ratio of the protein products to each other indicates that there is significantly more $3D^{pol}$ produced than the other products measured.

The Kan^R band does not migrate in such a diffuse manner when produced as a cleavage product of the polyprotein [$Kan^R3CD^{pro}Tet^R$] encoded by pEB15.. The reason for the contrast with the Kan^R product of pEB4 is not known. The upstream sequences of the two genes are the same but the downstream sequences are different. The downstream

Figure 4.4.3 Degradation of pEB15. pEB15 was usedto program a wheatgerm extract TnT reaction and incubated at 30 °C for 45 minutes. The reaction was then stopped by the addition of cycloheximide and RNase A. In order to study degradation of the products the reaction was incubated at 30° C and samples taken at 15 minute intervals. Degradation was stopped by the addition of loading buffer and samples were analysed on an SDS-PAGE gel.



sequences may have an effect on the protein produced and cause the disparity in the migration of the protein.

Lane	Band	PSL	PSL-BG	No. of mets	PSL/No. of mets	Ratio
1	3D	771.3	674.3	13	51.8	1 : 3.35 : 0.53
1	Tet	192	128.6	15	8.24	
1	Kan	167.2	123.6	8	15.45	
2	3D	1852.0	1654.0	13	127.2	1 : 3.47 : 0.53
2	Tet	399.3	277.3	15	19.56	
2	Kan	387.8	293.4	8	36.68	
3	3D	2151	1857.0	13	142.8	1 : 2.95 : 0.53
3	Tet	452.2	322.5	15	25.82	
3	Kan	499.0	387.3	8	48.41	
4	3D	1590	1468.0	13	112.9	1 : 2.96 : 0.53
4	Tet	397.4	288.4	15	20.34	
4	Kan	389.4	305.1	8	38.14	
5	3D	1801	1674.0	13	128.8	1 : 3.12 : 0.53
5	Tet	443	325.3	15	22.0	
5	Kan	433.0	330.1	8	41.3	
6	3D	1770	1646.0	13	126.6	1 : 3.12 : 0.53
6	Tet	438.4	315.0	15	21.6	
6	Kan	421.2	324.5	8	40.56	
7	3D	1307	1160.0	13	89.2	1 : 4.38 : 0.53
7	Tet	249.3	169.8	15	10.85	
7	Kan	222.9	162.8	8	20.35	
8	3D	1962	1447.0	13	111.3	1 : 3.05 : 0.53
8	Tet	465.2	277.2	15	19.5	
8	Kan	376.8	292.4	8	36.55	

Table 4.3 Results of densitometric analysis of degradation of pEB15. The PSL- BG (background) figures were divided by the number of methionines in each protein product to give a figure proportional to the number of moles of protein within each protein band. These figures were then used to calculate a ratio for the amount of Tet^R compared to that of Kan^R and $3D^{pol}$.

As the predicted molecular weight of the $3D^{pol}$ product, 51 kDa, is of approximately the same as that of the $Kan3C^{pro}$ product, 52 kDa, this band may represent *both* these cleavage products and the value given for $3D^{pol}$ is, in all probability, a combined figure for both $3D^{pol}$ and $[Kan^R3C^{pro}]$. If the values are corrected using the number of

methionines in both the $3C^{pol}$ and [$Kan^{R}3C^{pro}$] protein products there is a reduction in the ratio of 3D+Kan3C : Tet : Kan but the ratio is still a long way from 1:1 (Table 4.4).

Lane	Band	PSL-BG	No. of mets	PSL-BG/no. of mets	Ratio
1	3D+Kan3C	674.3	24	28.09	1 : 1.82 : 0.53
2	3D+Kan3C	1654	24	56.40	1 : 1.54 : 0.53
3	3D+Kan3C	1857	24	77.38	1 : 1.60 : 0.53
4	3D+Kan3C	1468	24	61.16	1 : 1.60 : 0.53
5	3D+Kan3C	1674	24	69.75	1 : 1.69 : 0.53
6	3D+Kan3C	1646	24	68.58	1 : 1.69 : 0.53
7	3D+Kan3C	1160	24	48.33	1 : 2.37 : 0.53
8	3D+Kan3C	1447	24	60.29	1 : 1.65 : 0.53

Table 4.4 Densitometric analysis of degradation of pEB15 corrected for presence of $Kan^{R}3C^{pro}$ in $3D^{pol}$ band. The PSL-BG value for the $3D^{pol}$ band was divided by the number of methionines in [$Kan^{R}3C^{pro}$] as this is approximately the same size as $3D^{pol}$.

4.4.2 Conclusions

All processing in pEB15 is highly probably *in cis* as processing *in trans* is slower than *cis* processing and the reactions were not incubated for a sufficient length of time. In order to try and enhance *trans* processing pEB15 was incubated overnight in a TnT reaction but there was still no further processing seen (see chapter 5). The processing event at the 3D/Tet junction does not mimic a cleavage performed in the virus as there is no sequence in the viral genome C-terminal to $3D^{pol}$. However, there are possible cleavage sites in the $3D^{pol}$ sequence which may allow for this cleavage in the context of the reporter polyprotein system. For possible sites see Figure 6.3. It may be possible that $3C^{pro}$ could perform cleavages downstream of $3D^{pol}$ if cleavage sites were present. In the case of [Kan^R3CD^{pro}Tet^R] $3C^{pro}$ could scan along the nascent protein to find any possible cleavage sites and then cleave at any possible sites. It may also be that $3C^{pro}$ will cleave to rid itself of any sequences whether they are viral or not but this doesn't seem to fit in with the non-cleavage of [3C^{pro}Tet^R].

Speculation can be made as to the order of cleavages performed by $3C^{pro}$ in pEB15. The cleavage between $3C^{pro}$ and $3D^{pol}$ must be the last to be performed if it is to occur *in cis*. The first cleavage to be performed would be either between Kan^R and $3C^{pro}$ or between $3D^{pol}$ and Tet^R. The second cleavage would be one of the two above which has not been performed first. A summary of the order of cleavages can be seen below (Figure 4.4.5).

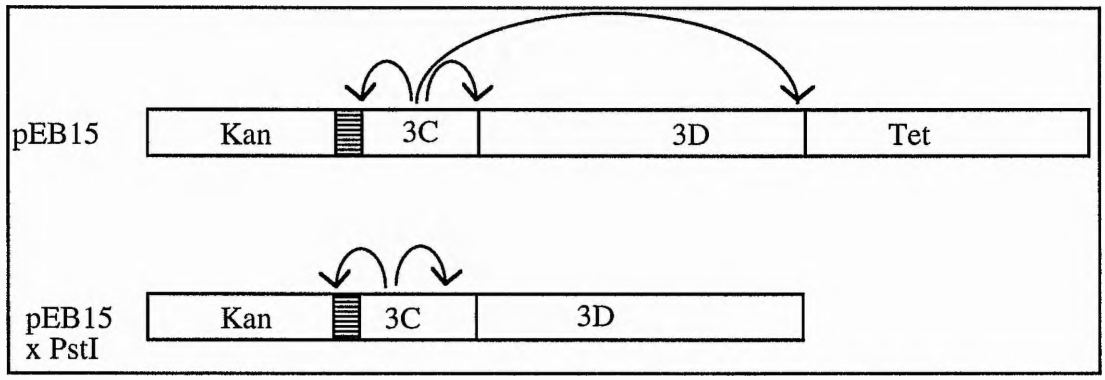


Figure 4.4.4 Summary diagram of cleavages in pEB15. The regions shaded with horizontal lines indicate the C-terminal region of 3B which is included to maintain the $3C^{pro}$ 3B/C cleavage site..

Summary

- With $3CD^{pro}$ in the reporter polyprotein system $3C^{pro}$ processes at all available sites to yield Kan^R, $3C^{pro}$, $3D^{pol}$ and Tet^R.
- There is equal processing both N- and C-terminally.
- The presence of the reporter proteins does not inhibit cleavage by $3C^{pro}$ either N- or C--terminally.
- From densitometric analysis of the cleavage products it appears that there may be some [Kan^R3C^{pro}] present.
- In comparison with pEB4 [$Kan^R3C^{pro}Tet^R$] pEB15.2 [Kan^R3CD^{pro}] shows stronger processing at the C-terminus of $3C^{pro}$ indicating that $3D^{pol}$ has a positive effect in regulating cleavage at the C-terminus of $3C^{pro}$.
- Processing occurs in pEB15 at the C-terminus of $3D^{pol}$ indicating that $3C^{pro}$ is capable of cleaving at sites which are not present in the viral sequence.

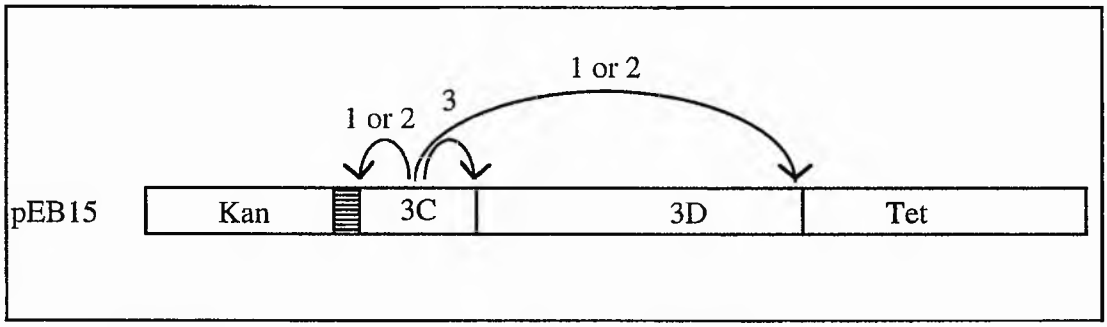


Figure 4.4.5 Proposed order of cleavages in pEB15. Shaded region indicates 3B sequences included to maintain 3B/C cleavage site.

4.5 Translation *in vitro* of pEB25.

The plasmid pEB25 [Kan^{R}3ABCTet^R] was designed to determine whether the sequence upstream of 3C^{pro} in the virus had any effect on cleavage in the reporter polyprotein system. Translation products correspond to [3C^{pro}Tet^R], Kan^R and 3AB (Figure 4.5 a&b), indicating that cleavage occurs at the N-terminus of 3C^{pro} and also at the N-terminus of 3AB. There is a band at approximately 20 kDa which is the same size as 3C^{pro} but as there is no corresponding Tet^R band it seems more likely that this is one of the internal initiation products of Kan^R. The molecular weights of the predicted cleavage products of pEB25 can be seen in Figure 4.5.1.

The cleavage products were identified by digesting the polyprotein construct before translation to give defined translation products and by the use of control plasmids (Figure 4.5 a). Digestion by PstI prior to translation would linearise the plasmid at the C-terminus of 3C^{pro} and the products of translation of the resulting linearised plasmid [Kan^{R}3ABC] are [Kan^{R}3AB] and 3C^{pro} (Figure 4.5 b). This indicates that the translation products of pEB25 are Kan^R, (3C^{pro}) and 3AB. There appears to be no band representing Tet^R. Sequencing of the plasmid confirmed that the sequence was correct.

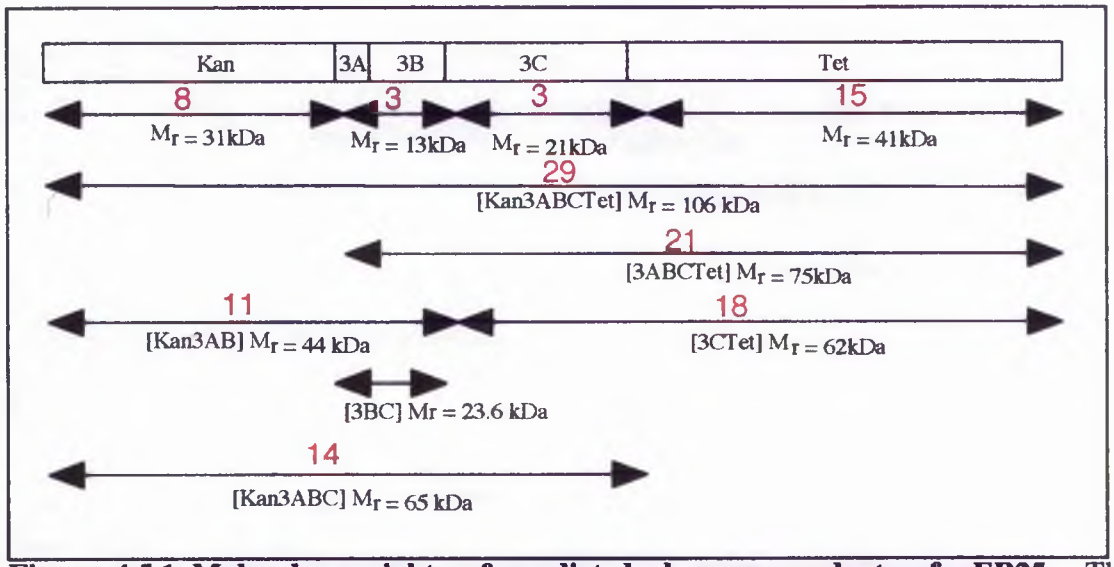
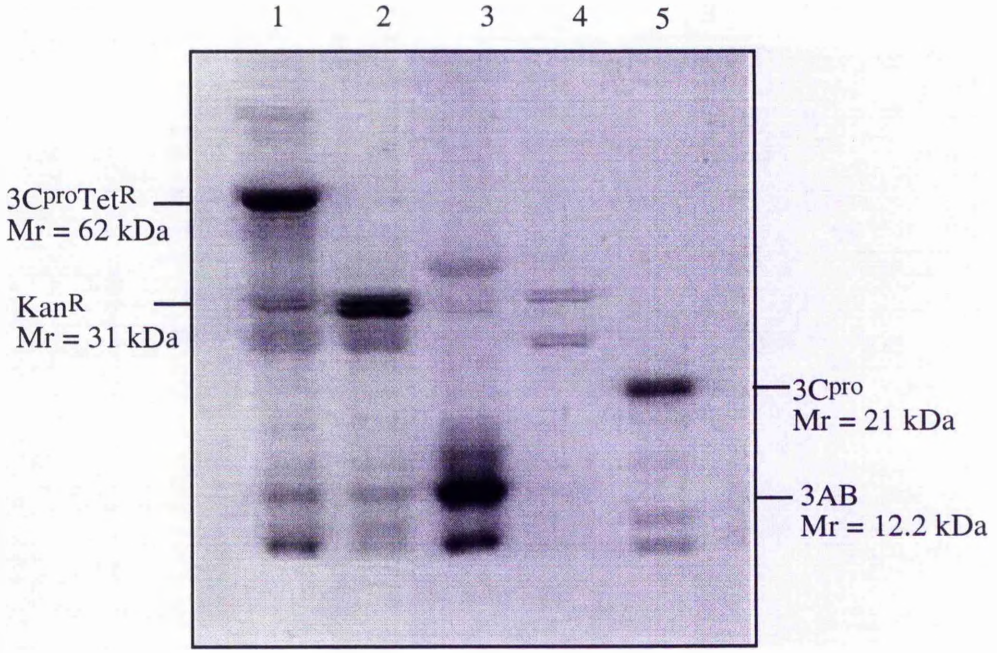


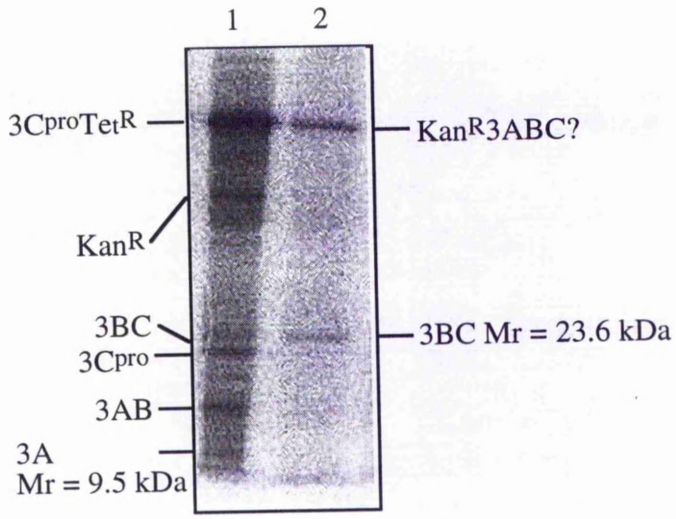
Figure 4.5.1 Molecular weights of predicted cleavage products of pEB25. The numbers in red are the number of methionine codons in each protein.

Figure 4.5 a Translation of pEB25. Plasmid DNAs were used to program a WGE TnT reaction and incubated for 90 minutes at 30 °C. Samples of the reaction products were analysed by SDS-PAGE. pEB19 (Lane 4) did not produce any translation products. The bands seen are artifacts of the TnT system.



Lane	Plasmid	Insert
1	pEB 25	[KanR3ABCTetR]
2	pEB 2	[KanR]
3	pEB 16	[HRV14 3AB]
4	pEB 19	[HRV14 3ABC]
5	pEB 3	[HRV14 3Cpro]

Figure 4.5 b Translation products of pEB25 and pEB25 x PstI (the plasmid was digested prior to incubation in the TnT reaction). The gel was exposed to film for 48 hours to obtain a clearer image.

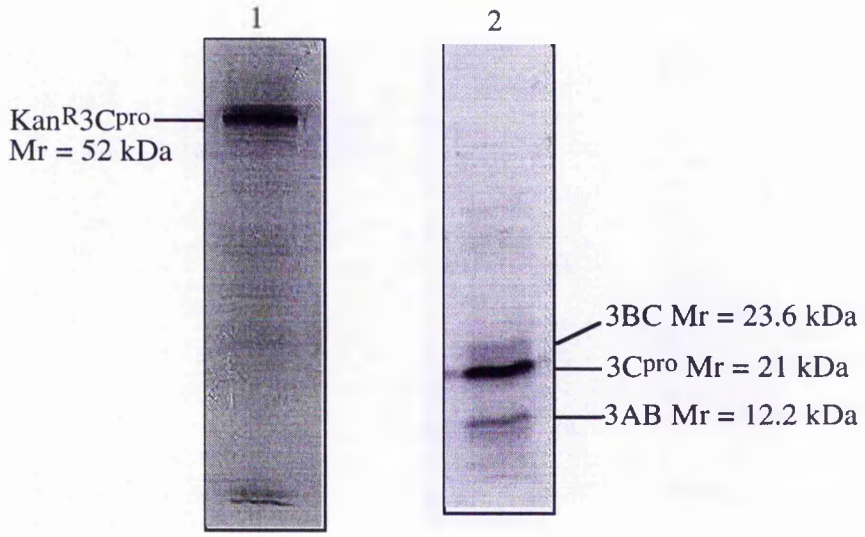


Lane 1 pEB25

Lane 2 pEB25 x PstI

The cleavage at the N-terminus of 3AB is at a glutamine-glycine pair, the primary cleavage which releases P3 from P2. If the order of cleavage in the reporter polyprotein system is the same as *in vivo*, [Kan^R3ABC] would be cleaved to Kan^R and 3ABC. $3C^{pro}$ will then cleave at its N-terminus to give 3AB and $3C^{pro}$. As there is no evidence of a band for $3C^{pro}$ this does not seem to be the case. The presence of 3AB seems to positively regulate the cleavage of 3ABC at the N-terminus of $3C^{pro}$. Cleavage of [Kan^R3C^{pro}] is not as efficient as cleavage of 3ABC (Figure 4.5.2). It appears that a downstream sequence is required for cleavage between Kan^R and 3ABC. This can be seen by comparing cleavage in pEB25 and pEB25 restricted with PstI. Tet^R positively regulates cleavage in all constructs, pEB4, pEB15 and pEB25.

Figure 4.5.2 Comparison of cleavage in pEB13 and pEB19. In order to compare cleavage by $3C^{pro}$ with different N-terminal extensions the translation products of these two plasmids were compared. The lanes are from different gels but can be compared qualitatively.



Lane 1 pEB14 [Kan^R3C^{pro}] in pGEM5zf(+)

Lane 2 pEB19 [HRV14 3ABC] in pGEM5zf(+)

4.5.1 Conclusions

The replacement of $3C^{pro}$ in the reporter system with 3ABC allows the influence of upstream viral sequences to be studied. The cleavage products of pEB25 [Kan^R 3ABCTet^R] are [$3C^{pro}Tet^R$], Kan^R , 3BC and 3AB (Figure 4.5). Cleavage occurs at the N-terminus of 3AB which is equivalent to the primary cleavage by $3C^{pro}$ between P2 and P3in the viral polyprotein. The production of [$3C^{pro}Tet^R$] indicates that cleavage also occurs at the N-terminus of $3C^{pro}$. There is no cleavage observed at the C-terminus of $3C^{pro}$ in this system which is indicated by the lack of a band representing Tet^R. This compares with the lack of cleavage seen in pEB5 [$3C^{pro}Tet^R$].

Comparing cleavage of 3ABC with cleavage of [Kan^R3C^{pro}] shows that cleavage of 3ABC is more efficient. This may be due to the action of 3AB in enhancing cleavage at the N-terminus of $3C^{pro}$ or it may be due to the fact that 3AB is a viral sequence whereas Kan^R is a foreign sequence. 3AB does not seem to enhance cleavage at the C-terminus of $3C^{pro}$ as there is not any cleavage seen at this site in this system.

When pEB25 is digested with PstI before incubation in the TnT system the cleavage products are [Kan^R3ABC] and 3BC. There is no evidence of [Kan^R3A] however as the remaining portion of the cleaved polyprotein. There is a $3C^{pro}$ Q-G cleavage site between 3A and 3B but in the viral polyprotein this site is rarely utilised. Cleavage still occurs at the N-terminus of $3C^{pro}$ but there is no cleavage at the N-terminus of 3A. This implies that a C-terminal extension may be required for $3C^{pro}$ to process at its N-terminus. The analysis of the construct [Kan^R3ABCD] would be useful to establish whether the inclusion of a viral sequence downstream of $3C^{pro}$ has the same effect as the presence of Tet^R.

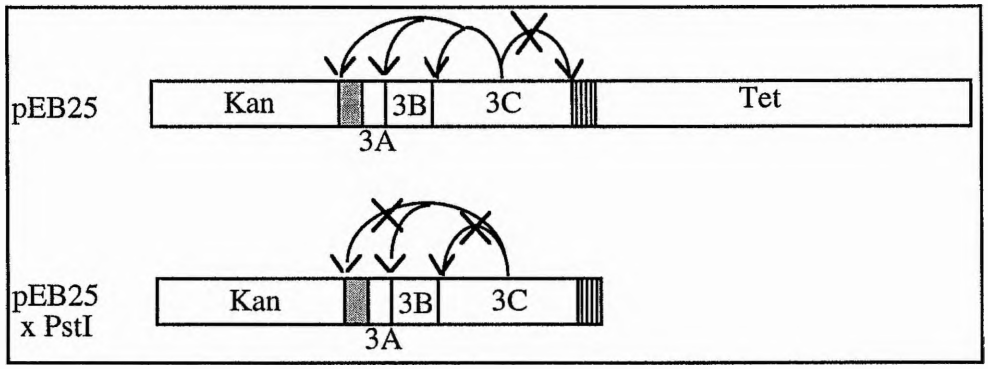


Figure 4.5.3 Summary of cleavages in pEB25. The regions shaded with dots are the C-terminal region of 2C and the regions shaded with vertical lines are the N-terminal region of 3D.

Summary

- When the HRV14 P3 cleavage product 3ABC is used as the viral sequence in the reporter polyprotein system cleavage occurs at the N-terminus of $3C^{pro}$ to give [$3C^{pro}Tet^{R}$].
- Cleavage also occurs at the N-terminus of 3AB giving Kan^{R} and 3AB as cleavage products.
- Cleavage occurs at the N-terminus of 3B to yield 3BC as a cleavage product
- 3AB has a possible function enhancing the N-terminal cleavage of $3C^{pro}$.
- 3AB does not have any enhancing effect on C-terminal cleavage by $3C^{pro}$.
- $3C^{pro}$ seems to require a C-terminal sequence in order to process at the N-terminus of 3AB.

4.6 Discussion

- **The 3C protease favours cleavage at its N-terminus in the artificial reporter polyprotein system.** In this polyprotein reporter system $3C^{pro}$ will cleave with both N- and C-terminal extensions. The degree of processing appears to be dependent on which extensions are present and, in some circumstances, the nature of the extension (*i.e.* viral or non-viral). Between the two non-viral reporter proteins $3C^{pro}$ cleaves both N- and C-terminally but favours cleavage at its N-terminus. If the C-terminal extension is removed from $3C^{pro}$ N-terminal processing still occurs but at a lower level than with the C-terminal extension present. By comparing [$Kan^R3C^{pro}Tet^R$] and [Kan^R3CD^{pro}] the effect of viral *versus* non-viral sequences can be observed. [Kan^R3CD^{pro}] cleaves to give Kan^R, $3C^{pro}$ and $3D^{pol}$ whereas [$Kan^R3C^{pro}Tet^R$] cleaves to give [Kan^R3C^{pro}] and [$3C^{pro}Tet^R$] which are not cleaved to give further products. Cleavage in [$Kan^R3C^{pro}Tet^R$] is less complete than in [Kan^R3CD^{pro}]. The cleavage products of [Kan^R3CD^{pro}] are produced in equal amounts and there is a large amount of uncleaved polyprotein present.

- **Effect of N-terminal extension on cleavage by $3C^{pro}$.** If the N-terminal extension is not present then differences in cleavage may also be seen depending on the nature of the C-terminal sequence. [$3C^{pro}Tet^R$] does not cleave to give $3C^{pro}$ and Tet^R either as a polyprotein or when it is a product of processing from [$Kan^R3C^{pro}Tet^R$]. $3CD^{pro}$ does not cleave to give $3C^{pro}$ and $3D^{pol}$ which ties in with earlier observations made *in vivo* and *in vitro*.

  The nature of the N-terminal extension of $3C^{pro}$ has an effect on cleavage. Cleavage occurs with any N-terminal extension whether viral or non-viral. If 3AB is present upstream of $3C^{pro}$ then cleavage occurs more efficiently at the N-terminus of $3C^{pro}$. This can be seen by comparing cleavage in [$Kan^R3C^{pro}Tet^R$] with cleavage in [$Kan^R3ABCTet^R$].

  Cleavage of [$Kan^R3ABCTet^R$] is however also dependent on the presence of a downstream sequence. [Kan^R3ABC] only cleaves once at the N-terminus of $3C^{pro}$ to

release Kan^R3AB and $3C^{pro}$. When Tet^R is present at the C-terminus of [Kan^R3ABC] cleavage also occurs between Kan^R and 3AB. This agrees with the results seen in pEB4 and pEB15 that presence of a C-terminal extension regulates N-terminal cleavage.

- **Cleavage in pEB15 shows unexpected cleavages by $3C^{pro}$ downstream of $3D^{pol}$.** Cleavage in pEB15 yields Kan^R, $3C^{pro}$, $3D^{pol}$, Tet^R and possibly [Kan^R3C^{pro}]. The cleavage between $3D^{pol}$ and Tet^R is unexpected as $3C^{pro}$ does not perform a cleavage downstream of $3D^{pol}$ in the virus. The viral polyprotein does not have any coding sequences C-terminal to $3D^{pol}$ and it was not thought that $3C^{pro}$ would cleave any sequence downstream of $3D^{pol}$. Cleavage by $3C^{pro}$ downstream of $3D^{pol}$ is discussed further in chapter 6.

5. Human Rhinovirus 3C proteolytic processing *in trans*.

Cleavages performed in the processing pathway of picornaviruses occur *in cis* (an intramolecular reaction) or *in trans* (an intermolecular reaction). There is one exception to this in the maturation cleavage whose mechanism has yet to be determined (see section 1.5.5). In *cis* cleavages the protease is acting at a cleavage site whilst still part of the same polyprotein. These cleavages are often rapid and occur co-translationally. They are frequently referred to as autocatalytic and follow zero order kinetics, and, therefore are insensitive to dilution. Cleavages performed *in trans* are exemplified by a protein containing proteolytic activity cleaving another peptide or protein at a substrate cleavage site. This type of reaction follows second order kinetics and is sensitive to dilution. Cleavages which are performed by the protease *in trans* occur more slowly than those *in cis*.

In poliovirus the primary cleavage product P1 is cleaved *in trans* to yield the capsid proteins (Hanecak *et al.*, 1982). The protease responsible was identified as $3C^{pro}$ however, additional sequences from either 3B and/or $3D^{pol}$ were also thought to be required (Hanecak *et al.*, 1982;Ypma-Wong and Semler, 1987). Surprisingly P1 was found to be processed *in trans* by $3CD^{pro}$, rather than $3C^{pro}$, to give the structural proteins 1AB, 1C and 1D (Jore *et al.*, 1988;Ypma-Wong *et al.*, 1988). The final cleavage of 1AB to 1A and 1B, known as the maturation cleavage, is mediated by an unknown entity. In some other picornaviruses, such as EMCV, P1 is processed by $3C^{pro}$ alone (Parks *et al.*, 1986).

In an alternative cleavage pathway in poliovirus $3CD^{pro}$ is processed to yield 3C' and 3D' (Hanecak *et al.*, 1982;Pallansch*et al.*, 1984;Rueckert*et al.*, 1979). This cleavage is catalysed by $2A^{pro}$ (Toyoda *et al.*, 1986) acting *in trans*. Cleavage occurs at a conserved tyr-gly pair in the $3D^{pol}$ sequence (Figure 5). It gives a product, 3C', which is larger than $3C^{pro}$ but not as large as $3CD^{pro}$ and another (3D')which is smaller than $3D^{pol}$. This cleavage had also been observed previously in the P3 region of human rhinovirus 1a (HRV-1a; McLean *et al.*, 1976) . No function has been found for the alternative cleavage

products and they have been shown not be essential for poliovirus replication in tissue culture (Lee and Wimmer, 1988). The Y-G site is completely conserved in the entero- and rhinovirus group with the exception of HRV16 where a single base change results in an aspartic acid residue instead of the tyrosine (see figure 6.4)

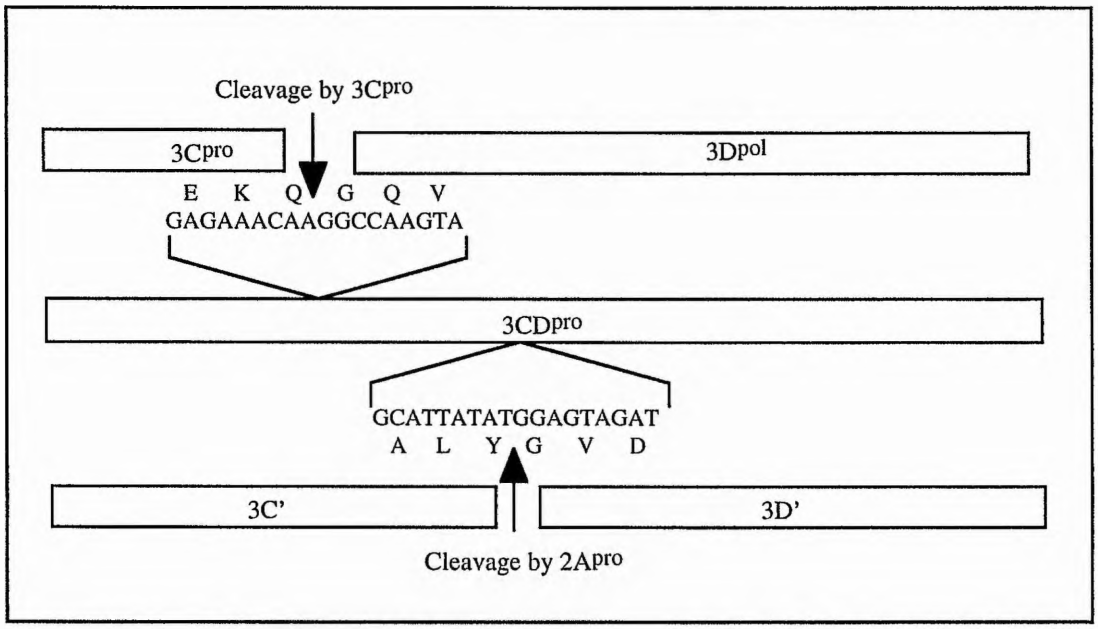


Figure 5 Diagram to show alternative cleavage products of 3CD^{pro} due to cleavage *in cis* by 3C^{pro} or cleavage *in trans* by 2A^{pro}.

It was hypothesised that the alternative cleavage product 3C' may be able to act in the same way as 3CD^{pro} and process P1 *in trans.* Previous experiments, however, have shown that the whole 3CD^{pro} sequence is required for P1 processing in poliovirus (Jore *et al.*, 1988;Ypma-Wong *et al.*, 1988). Deletions of 3D^{pol} resulted in a lack of processing (Jore *et al.*, 1988) and some transcripts containing up to 75% of 3D^{pol} will not process P1 in poliovirus (Ypma-Wong *et al.*, 1988).

In HRV14 P1 is processed *in trans* by 3CD^{pro} to produce the capsid proteins. As no function has yet been ascribed to 3C' it was decided to see whether the alternative cleavage product had the same ability to cleave P1 *in trans* in HRV14.

5.1 Construction of plasmids.

5.1.1 pGEM7zf(+)GFP-2A.

A plasmid containing the coding sequences for green fluorescent protein (GFP) and HRV14 $2A^{pro}$ was constructed by Vanessa Cowton in the laboratory. GFP was amplified by the PCR from a plasmid containing the wild type GFP sequence donated by Simon Santa Cruz at the Scottish Crop Research Institute. The 2A protease was amplified by PCR from a plasmid containing the complete HRV14 sequence using oligonucleotide primers designed to the 5' and 3' ends of the gene. The genes were then ligated into the multiple cloning site of pGEM7zf(+) (Figure 5.1.1) to form an artificial polyprotein and provide a control for the N-terminal cleavage by $2A^{pro}$. If cleavage occurs at the N-terminus of $2A^{pro}$ then two cleavage products will be seen, GFP and $2A^{pro}$, in the TnT reaction. If cleavage has not occurred then one product will be seen representing the polyprotein precursor GFP-2A.

5.1.2 pGEM-T 3C'.

The product of the alternative cleavage in $3CD^{pro}$, 3C', was amplified by the PCR from a plasmid containing the full length sequence of HRV14. The $3C^{pro}$ 5' primer employed previously was used and a new primer was designed to complement reverse sequence of the 3' end of the nucleotide sequence. The restriction sites added were AatII at the 5' end and PstI at the 3' end. The PCR product was cloned into pGEM-T, a transcription vector which utilises the additional adenine left by PCR at the ends of the products. The resulting plasmid (pGEM-T 3C') was analysed by restriction digest and then sequenced with T7 and SP6 polymerase promoter oligonucleotide primers to confirm that it was correct. It is important to confirm orientation of the insert so that the correct polymerase may be used in the TnT reaction. The insert was found to be under the control of the T7 polymerase promoter and this polymerase was used in the TnT reactions (Figure 5.1.2).

Figure 5.1.1 pGEM7zf(+) GFP-2A. Constructed by Vanessa Cowton. GFP and HRV14 2A^{pro} were amplified by the PCR and then ligated into pGEM7zf(+) to form an artificial polyprotein.

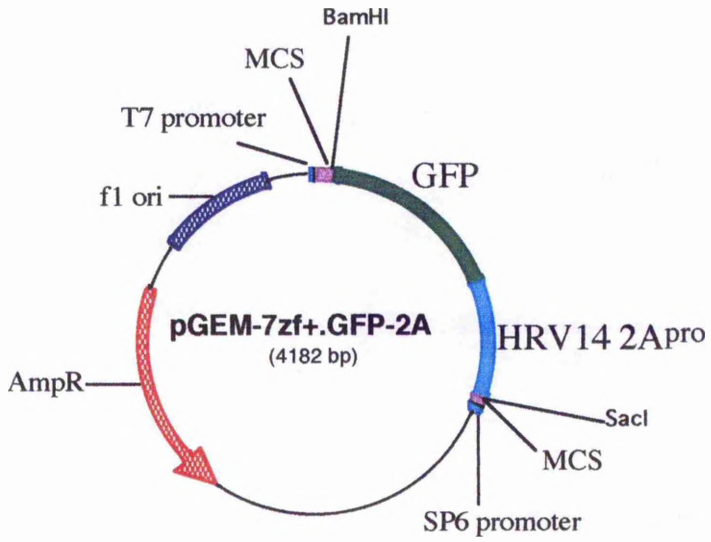
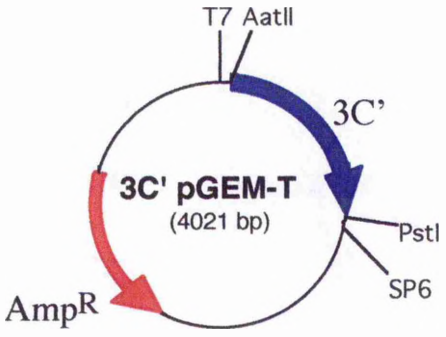


Figure 5.1.2 3C' pGEM-T. Constructed by Vanessa Cowton. 3C' was amplified by the PCR from a plasmid containing the full length HRV14 coding sequence. It was ligated into pGEM-T and sequenced to confirm the orientation of the insert.



5.1.3 pEB7.

The 3CD protease gene was amplified by the PCR from a plasmid containing the full length coding sequence for HRV14 using primers designed to the 5' and 3' ends of the gene (see table 2.1). Restriction sites were incorporated into the primer sequence to add sites to the ends of the PCR product for easy ligation into the chosen vector. A HindIII site was added to the 5' end and a KpnI site to the 3' end. The PCR product was restricted with these enzymes and ligated into pRSETA which had been similarly restricted. The resulting plasmid (pEB7; Figure 5.1.3) was used to program TnT systems and to express the protein in *E.coli.*

5.1.4 pEB8.

HRV14 P1 was amplified from a plasmid containing the complete coding sequence of HRV14 by the PCR. Restriction sites were added to the 5' and 3' ends to facilitate ligation into a transcription vector. An SpeI site was added to the 5' end and a KpnI site to the 3' end. In order to make it easier to see whether the vector had been digested prior to ligation a plasmid containing an insert was used to provide the vector. The chosen vector, pBluescript IIKS+, had previously been used by Mike Flint in the laboratory to make a plasmid containing FMDV $3CD^{pro}$ (pMF16c). The insert was released from the plasmid using KpnI and SpeI leaving the vector with the correct sticky ends for ligation with the cut P1 PCR product (Figure 5.1.4) to produce the plasmid pEB8.

All these constructs were used to program reticulocyte lysate systems as controls.

Figure 5.1.3 Construction of pEB7. The 3CD^{pro} coding sequence was amplified by the PCR from a plasmid containing the full length HRV14 coding sequence, using primers OB14 and OB16 (A). The PCR product was restricted with KpnI and HindIII and ligated into pRSETA which had been similarly restricted. The plasmid created, pEB7, was designed to program RRL and WGE TnT reactions and to express 3CD^{pro} in *E.coli* .

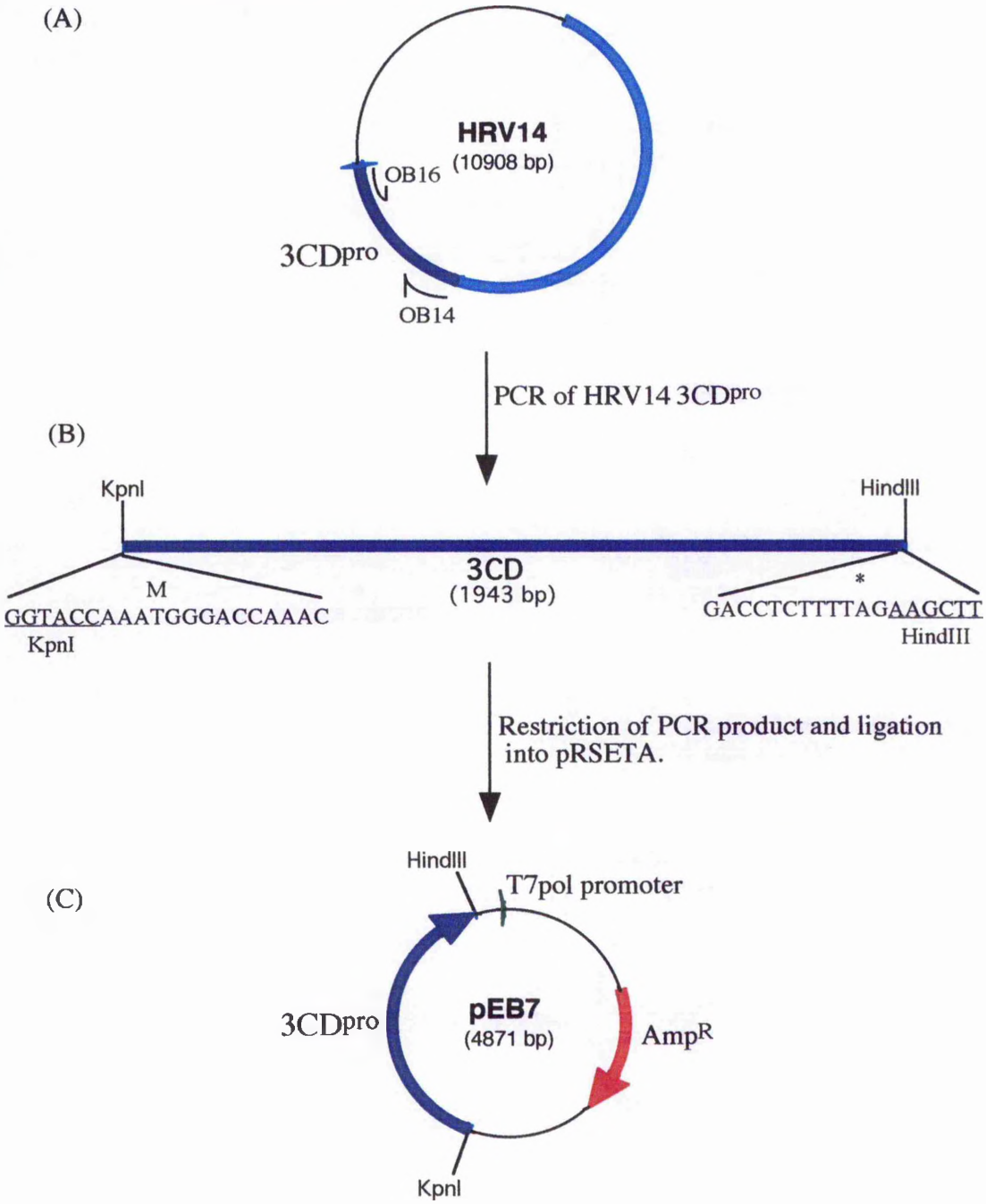
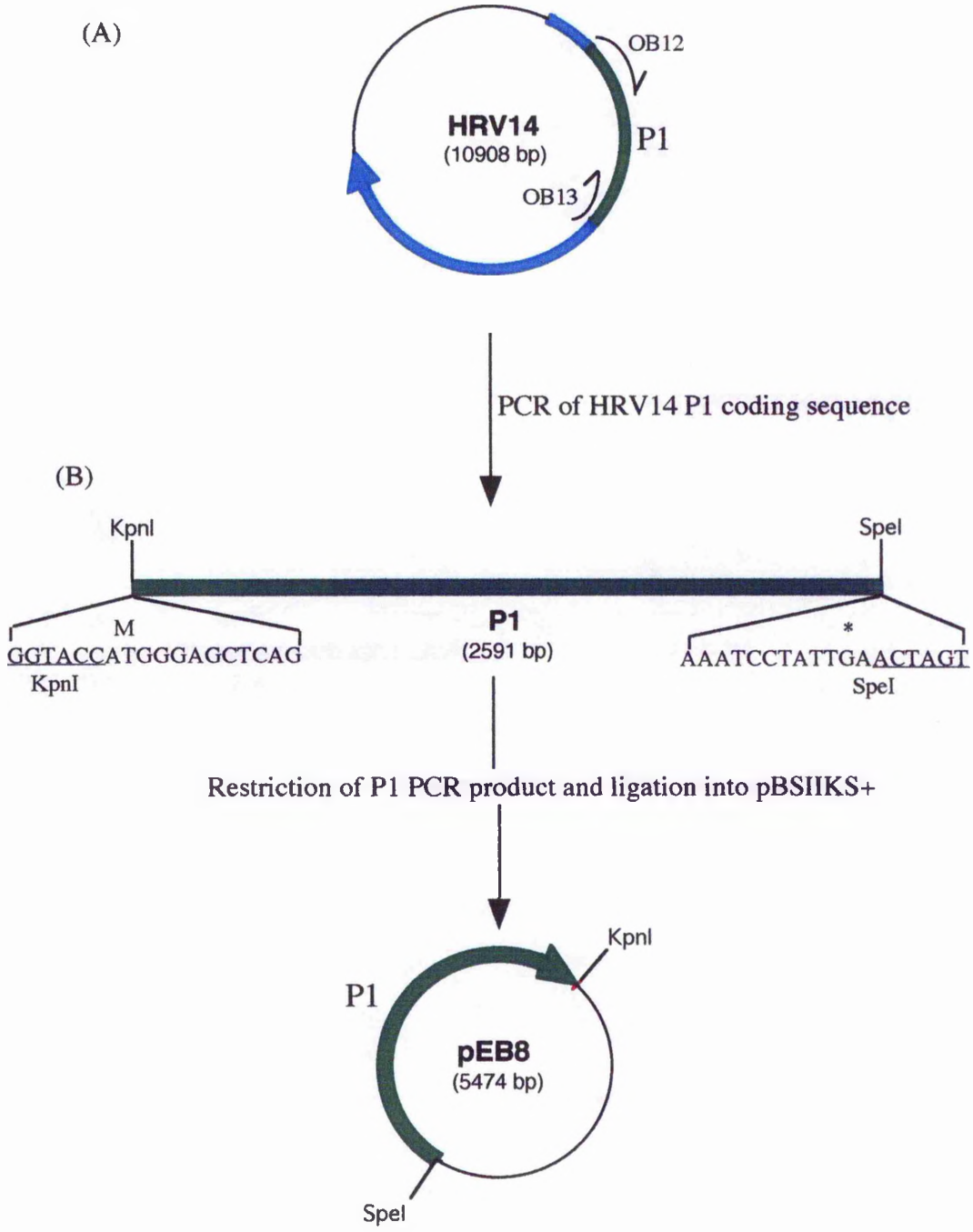


Figure 5.1.4 Construction of pEB8. The coding sequence for HRV14 P1 was amplified by the PCR from a plasmid encoding the full length HRV14 sequence, using oligos OB12 and OB13 (A). The PCR product was then restricted with KpnI and SpeI (B) and ligated into pBSIIKS+ which had been similarly restricted. The resulting plasmid pEB8 was used to program RRL TnT reactions.



5.2 Processing of HRV14 3CD^{pro} by 2A^{pro} *in trans*.

The control plasmids for this experiment were translated to check whether they gave translation and cleavage products of the expected sizes (Figure 5.2.2). The plasmid DNA was used to program a rabbit reticulocyte lysate (RRL) TnT reaction, incorporating ^{35}S-methionine as the radiolabel. After incubation for 90 minutes at 30 °C translation was stopped by the addition of SDS-PAGE loading buffer and samples were loaded onto polyacrylamide gels (Figure 5.2.1). 10% and 15% running gels were used to assess which would be best for visualising the small products of cleavage. The rabbit reticulocyte lysate system was used in these experiments as it had already been determined, in previous experiments carried out by Vanessa Cowton in the laboratory, that the cleavage products of GFP-2A are more easily visualised in this system than in the wheatgerm extract system.

The tracks on both the 10% and 15% gels were clear and it was decided to use a 10% gel for all further experiments. GFP-2A cleaves to give GFP and 2A^{pro} (Figure 5.2.1 a &b Lane 2). The 2A protease cleaves *in cis* at its own N-terminus. Cleavage is total in this instance as no band representing GFP-2A is seen on the gel. The predicted molecular weight of GFP-2A is 44 kDa. GFP has a predicted molecular weight of 27 kDa and 2A^{pro} 17 kDa.

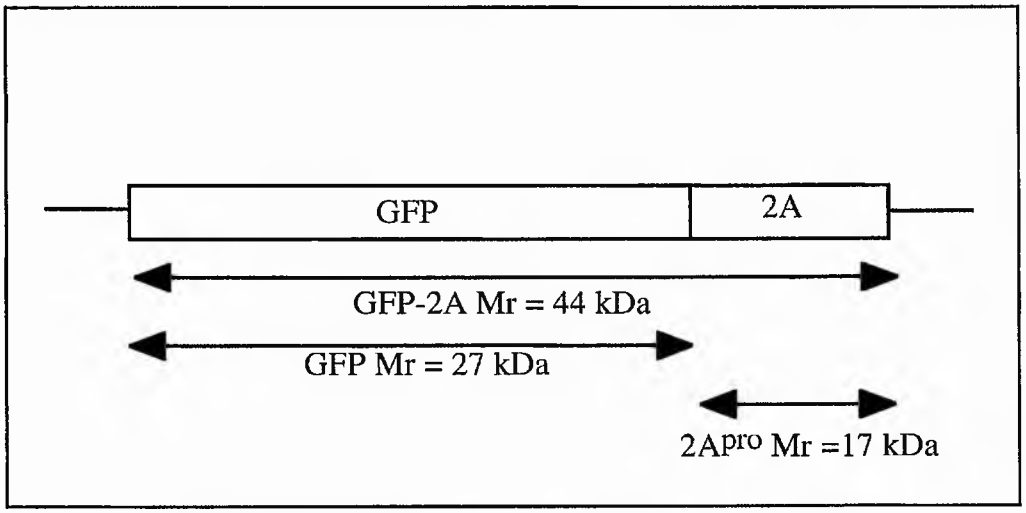
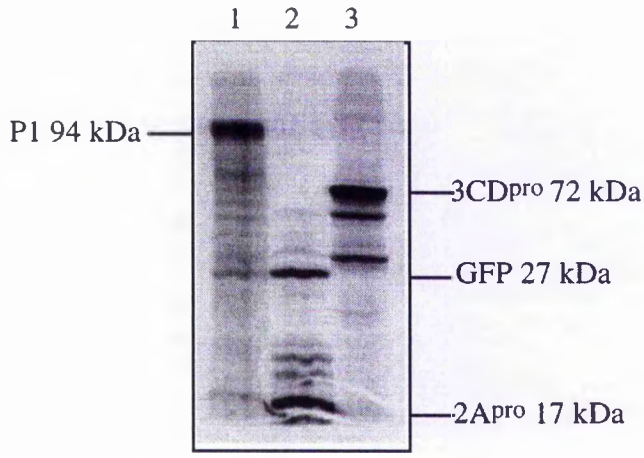
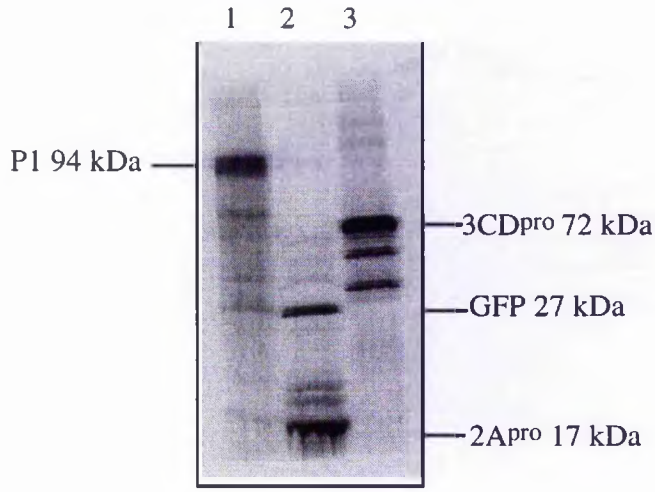


Figure 5.2.2 Diagram to show the predicted molecular weights of the cleavage products of GFP-2A

Figure 5.2.1 Translation of control plasmids. Control plasmids were used to program RRL TnT reactions labelled with ^{35}S-methionine. Translation products were analysed by SDS-PAGE.



A. 10 % polyacrylamide gel



B. 15 % polyacrylamide gel

Lane 1 pEB8 HRV14 P1 in pBBSKS+

Lane 2 pGEM7zf(+)GFP-2A

Lane 3 pEB7 HRV14 3CDpro in pRSETA

Translation of pEB7 gave three prominent bands. These were identified as the full length $3CD^{pro}$ translation product (72kDa) and two internal initiation products (discussed in chapter 3). Figure 5.2.3 shows the predicted molecular weights for the expected translation products. These were identified and the percentage initiation from each calculated (see section 3.10). This indicates that $3CD^{pro}$ is stable in this system which is consistent with results from *in vivo* work.

Translation of pEB8 gave the expected P1 translation product with a predicted molecular weight of 94 kDa.

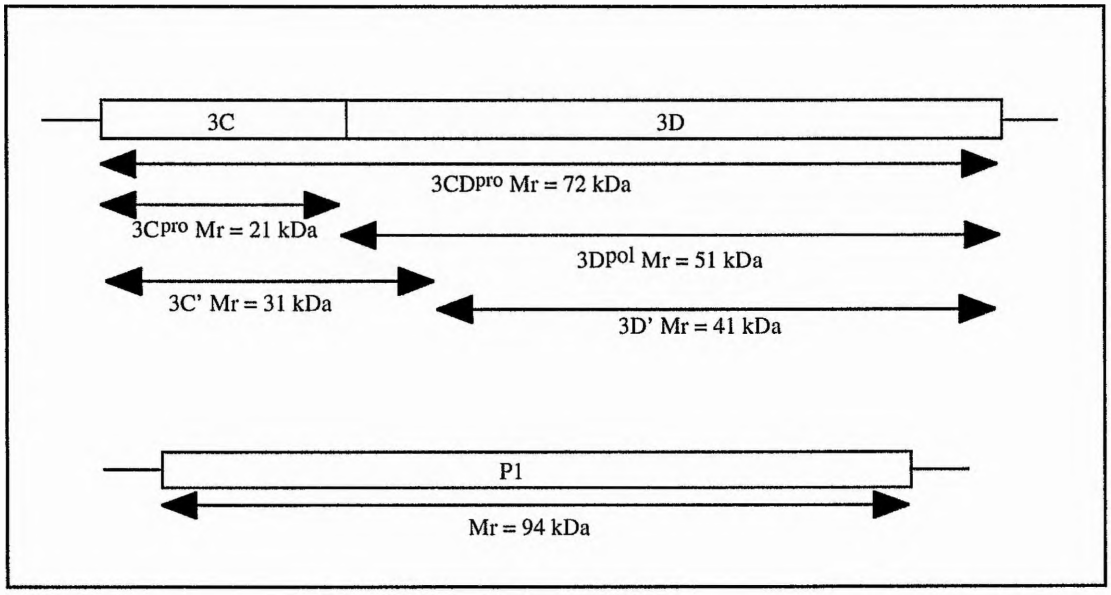


Figure 5.2.3 Predicted molecular weights of the translation products of pEB7 and pEB8.

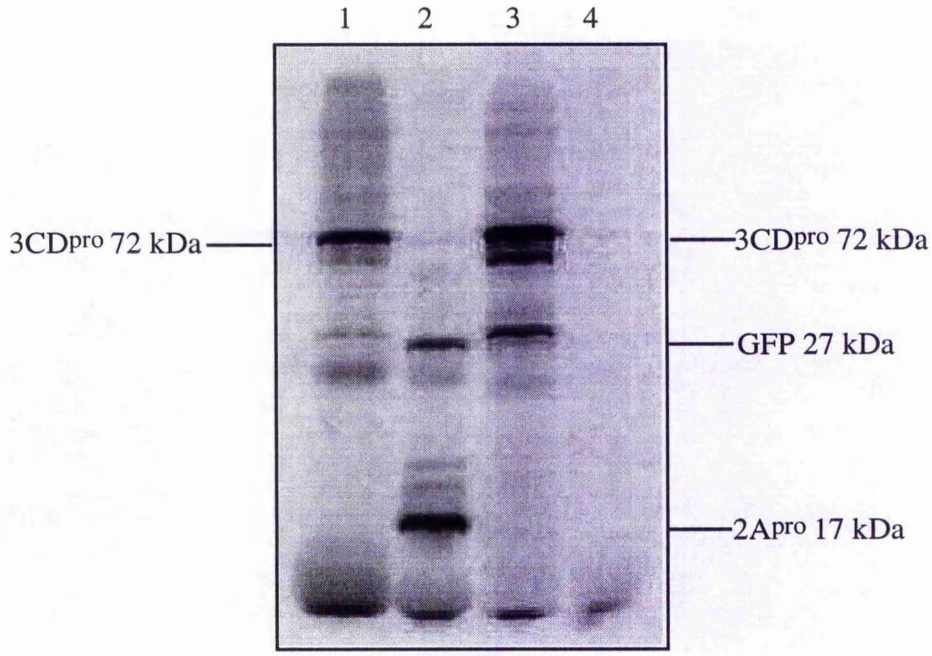
5.3 Cleavage *in trans* of 3CD^{pro} by 2A^{pro}.

In order to observe the cleavage of HRV14 3CD^{pro} by HRV14 2A^{pro} *in trans*, unlabelled 2A^{pro} was synthesised in a rabbit reticulocyte lysate (RRL) system. An amino acids minus leucine mix was substituted for the radiolabelled ^{35}S-methionine to provide methionine for the reaction. The reaction was incubated at 30 °C for 90 minutes and then stopped. The reaction was stopped by the addition of cycloheximide (to a final concentration of 0.8 μg/ml), to stop protein synthesis, and RNaseA (to a final concentration of 0.6 μg/ml), to degrade any mRNA. Labelled substrate, *i.e.* 3CD^{pro} was synthesised in a reaction with ^{35}S-methionine as the radiolabel in the same way. An aliquot (7.5 μl) of the unlabelled GFP-2A reaction was then mixed with 5.0 μl of the labelled 3CD^{pro} reaction, overlaid with mineral oil and incubated overnight at 30 °C.

The 3C' coding sequence was amplified by the PCR and cloned into a transcription vector to use as a control. It was known that 2A^{pro} was active *in cis* (GFP-2A cleaved to yield GFP and 2A^{pro}) in the TnT reaction. A reaction was therefore set up to determine whether 2A^{pro} would be active *in trans* and cleave 3CD^{pro} to yield the alternative cleavage products 3C' and 3D'.

A sample of the overnight reaction was analysed on an SDS-PAGE gel with control reactions to provide size standards (Figure 5.3.1). It can be seen from the gel that there may be a low level of cleavage in 3CD^{pro} to give the alternative cleavage product 3C'. The predicted molecular weight of the 3C' cleavage product is 31 kDa. The plasmid containing the 3C' sequence did not produce any translation products. The DNA had been cleaned by phenol chloroform extraction, ethanol precipitation and spun column chromatography with G-50 sepharose. The plasmid was sequenced and the nucleotide sequence of the insert was found to be correct. The lack of translation product may be due to an error in the promoter sequence of the transcription vector used so the insert needs to be ligated into another vector for further experiments.

Figure 5.3.1 Cleavage of HRV14 3CD^{pro} *in trans*. Plasmid DNA was used to program RRL TnT reactions. Control reactions were carried out to provide size markers for the*trans* processing reaction between 2A^{pro} and 3CD^{pro} (Lane 4).



Lane 1 3CD+ GFP-2A
Lane 2 GFP-2A
Lane 3 3CD
Lane 4 3C'

The reaction in Lane 1 was carried to monitor processing of 3CD^{pro} by 2A^{pro} *in trans*. 3CD^{pro} was synthesised in an RRL reaction with ^{35}S-methionine as a radiolabel. 2A^{pro} was synthesised in an unlabelled reaction. 7.5 µl of the unlabelled protease was added to 5.0 µl of the labelled substrate and incubated overnight at 30 °C.

5.4 Processing of P1 *in trans*.

5.4.1 Processing of HRV14 P1 *in trans*.

In order to see whether the alternative cleavage product 3C' is capable of performing the P1 processing events, an experiment was devised to monitor these *trans* reactions. The substrate, P1, was synthesised in a ^{35}S-labelled RRL TnT reaction and the protease source, either 3CD^{pro} or 3C', was synthesised in an unlabelled RRL reaction. Both reactions were stopped after 90 minutes at 30 °C by the addition of cycloheximide and RNaseA (see section 5.3). Unlabelled protease (7.5µl) was added to 5µl of the labelled substrate. The mixture was overlaid with mineral oil and incubated at 30 °C overnight. Samples from the reactions were then analysed on SDS-PAGE gels with samples of the reaction components as controls (Figure 5.4.1).

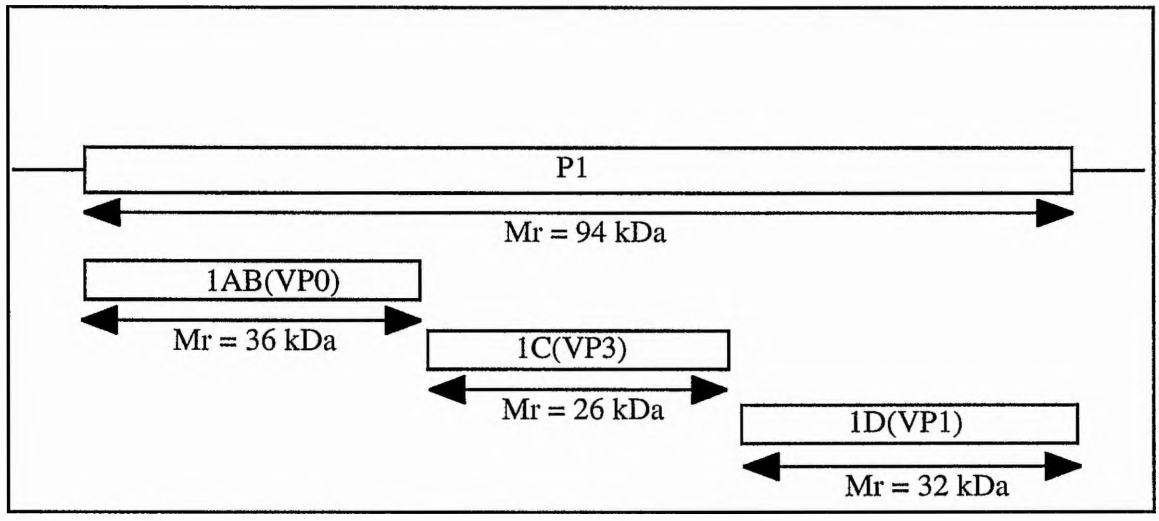
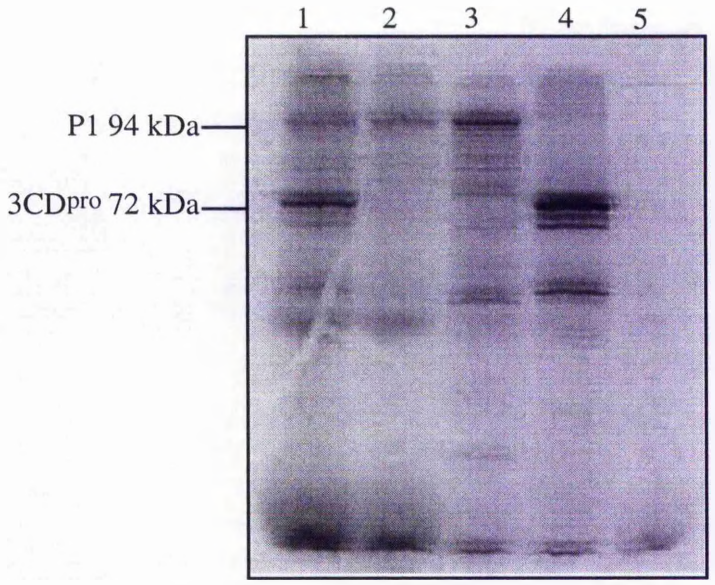


Figure 5.4.2 Predicted molecular weights of cleavage products of P1.

If cleavage occurred *in trans* then translation products representing the capsid proteins 1AB, 1C and 1D would be seen on the gel. Lane 1 has bands representing the uncleaved P1 polyprotein and 3CD^{pro} which can be identified by comparison with the

Figure 5.4.1 Processing *in trans* of HRV14 P1. The reactions in Lanes 1 and 2 were incubated overnight at 30 °C. The reaction consisted of 7.5µl of an unlabelled protease source and 5.0 µl of a labelled substrate. These were synthesised separately in RRL reactions incubated at 30 °C for 90 minutes.



Lane 1 3CD + P1
Lane 2 3C' + P1
Lane 3 P1
Lane 4 3CD
Lane 5 3C'

control lanes. The protease source was labelled in error so the translation products of this can also be seen. There are extra bands in lane 1 which may be processing products of P1 but it is unlikely that processing has occurred *in trans* due to the failure of $2A^{pro}$ to process $3CD^{pro}$. If the extra bands are processing products then they are not very distinct and processing is at a very low level. Previous experiments have indicated that P1 is processed to VP1 (1D) but no VP0 (1AB) or VP3 (1C) is seen (Jore *et al.*, 1988). The predicted molecular weight for HRV14 1C is 26 kDa as shown in Figure 5.4.2 above. These extra bands are not present in the 3C' + P1 lane (Lane 2) indicating that there was no *trans* processing with 3C' as the protease source in the reaction.

The discrepancy between the intensity of the P1 bands in lanes 1 and 3 is due to unequal loading between these lanes. The amount of P1 loaded in the control lane (Lane 3) is approximately five times that loaded in Lane 1.

To create conditions that may favour *trans* processing DTT was added to the reactions and co-translation reactions were also tried.

5.4.2 Addition of DTT and adjustment of reaction components.

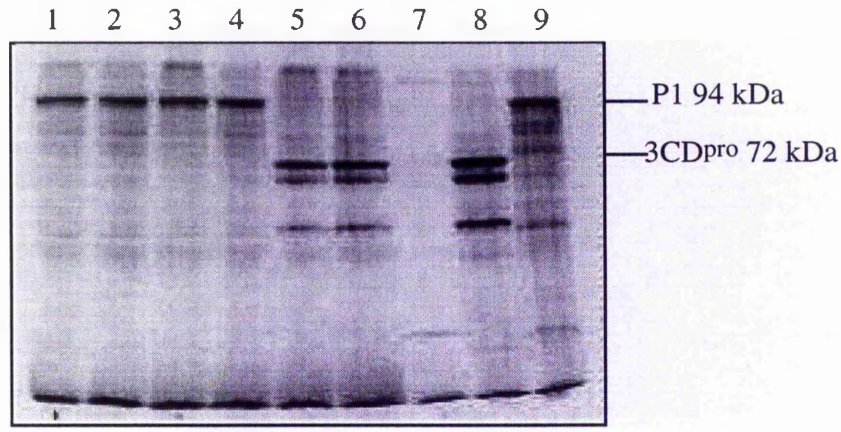
The reactions were repeated in much the same way as before with labelled and unlabelled components being synthesised separately and then incubated together overnight. The volumes of reactants were changed so that equal volumes (5 μl) of the substrate and protease were added to the overnight incubation. A second set of reactions was also carried out with the same volumes of reactants (5 μl) and the addition of DTT (2mM final concentration) to the overnight reaction. DTT is a reducing agent which stops the oxidation of the protease by protection of the thiol groups. This allows the protease to be active over a longer period of time which would be favourable for *trans* processing. Samples of the reactions were analysed by SDS-PAGE (Figure 5.4.3).

It can be seen from figure 5.4.3 that no *trans* processing occured in the overnight reactions between $3CD^{pro}$ and P1 or 3C' and P1. The only translation product seen is that of uncleaved P1 which is the same as that in the control lane. If *trans* processing had taken place P1 would have been expected to be processed into the capsid proteins 1C and

1D and the precursor 1AB. The predicted molecular weights of these proteins can be seen in figure 5.4.2. The addition of 2mM DTT (lanes 1 and 3) did not have any effect on the lack of *trans* processing.

The 3CD protease does not seem to have undergone any *trans* processing on incubation with GFP-2A. The bands present in both the lane with DTT and the one without show the same bands as the control $3CD^{pro}$ lane. GFP-2A was expected to cleave $3CD^{pro}$ at the alternative tyr-gly cleavage site in $3D^{pol}$ to produce 3C'. The $2A^{pro}$ in this artificial polyprotein is known to be active *in cis* as it cleaves itself away from GFP but it does not seem to be active *in trans*..

Figure 5.4.3 Addition of DTT to *trans* processing reactions. Reactions in Lanes 1 - 6 were synthesised separately (the protease source unlabelled and the substrate labelled with ^{35}S-methionine) and then combined and incubated overnight at 30 °C.



Lane 1 P1 + 3C' with 2 mM DTT
Lane 2 P1 + 3C' no DTT
Lane 3 P1+ 3CD with 2mM DTT
Lane 4 P1 + 3CD no DTT
Lane 5 3CD + GFP-2A with 2mM DTT
Lane 6 3CD + GFP-2A no DTT
Lane 7 3C'
Lane 8 3CD
Lane 9 P1

5.4.3 Co-translations.

To enhance *trans* processing between $2A^{pro}$ and $3CD^{pro}$ and 3C' and P1 co-translation reactions were carried out. Equal amounts (500ng) of the two constituent DNAs were added to a 20μl RRL TnT reaction and incubated at 30 °C for 90 minutes. Reactions were then stopped by the addition of cycloheximide and RNaseA (see section 5.3). The reactions were overlaid with mineral oil and incubated overnight at 30 °C.

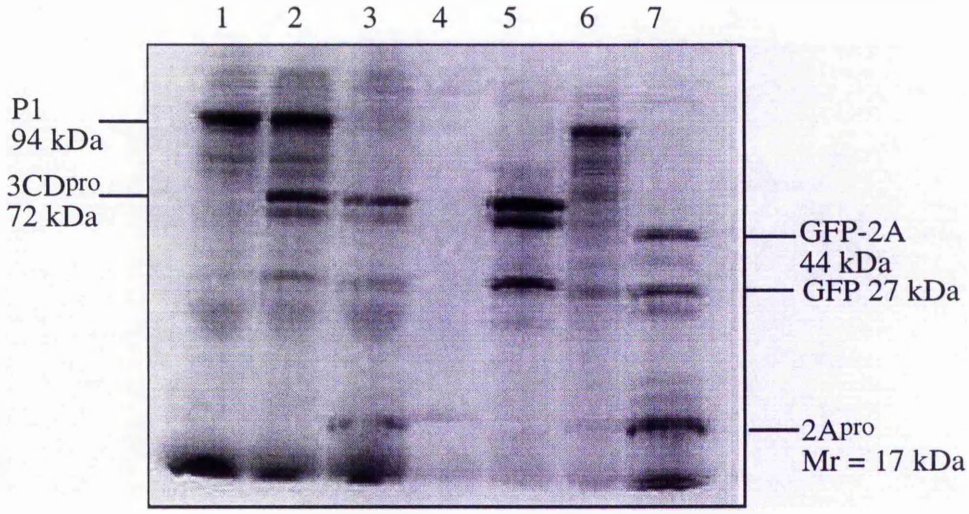
Samples of all the reactions were analysed by SDS-PAGE on a 10% polyacrylamide gel (Figure 5.4.4).

Incubation of $3CD^{pro}$ with P1 failed to result in *trans* cleavage to produce any capsid proteins. The translation products seen on the polyacrylamide gel (figure 5.4.4 lane 2) are the same as those of P1 and $3CD^{pro}$ in the control reactions showing that no processing of P1 had occurred. The protease source in this experiment was labelled due to the nature of the co-translation reaction and therefore shows up on the autoradiograph of the SDS-PAGE gel. The incubation of 3C' with P1 also failed to show any processing but this is not necessarily due to the non-proteolytic nature of 3C'. The fact that $3CD^{pro}$ also fails to show any processing of P1 indicates that the whole system is unsuitable and conclusions cannot be drawn about the processing activity of 3C'.

Incubation of $3CD^{pro}$ with GFP-2A does not show any cleavage at the alternative cleavage site. Both components were labelled with ^{35}S-methionine and so there are bands from both $3CD^{pro}$ and GFP-2A. There are not any extra bands indicating that if any cleavage products are present they are of a similar size to the original components. The product of cleavage at the tyr-gly pair has a predicted molecular weight of 31 kDa which is around the same size as GFP. As no cleavage has been seen in previous incubations it seems unlikely that there is a 3C' cleavage product being masked by the GFP band in this translation.

As even the control reactions failed to show any signs of *trans* processing it was concluded that the system is unsuitable for this type of analysis. It may be that separate transcription and translation reactions would enable *trans* processing to proceed. The

Figure 5.4.4 Co-translation reactions. Co-translations of the pairs of plasmids were carried out in an RRL TnT system. Equal amounts of each plasmid DNA were used to program a 20 µl reaction and then incubated at 30 °C for 90 minutes. Reactions were then stopped by the addition of cycloheximide and RNaseA and the reactions were overlaid with mineral oil. To allow *trans* processing to take place the reacstions were incubated at 30 °C overnight.



Lane 1 P1 + 3C'
Lane 2 P1 + 3CD
Lane 3 GFP-2A + 3CD
Lane 4 3C'
Lane 5 3CD
Lane 6 P1
Lane 7 GFP-2A

length of incubation should have been sufficient to observe *trans* processing. If the reactions are left for any longer degradation starts to occur and all products are destroyed.

The control reaction for GFP-2A on this gel (Figure 5.4.3) shows some uncleaved GFP-2A. This is different to other translations of this plasmid. The ratios of GFP to $2A^{pro}$ are not constant in other translations (Table 5.1) but this is the only experiment which shows uncleaved GFP. Densitometric analysis shows the percentage of uncleaved GFP-2A to be approximately 18%.

Figure	5.2.1 10%			5.2.1 15%			5.3.1			5.4.3		
	GFP-2A	GFP	2A	GFP-2A	GFP	2A	GFP-2A	GFP	2A	GFP-2A	GFP	2A
PSL-BG	-	928.8	1640	-	841.6	1666.6	-	231.0	1031.0	575.0	647.8	526.3
No. of mets	-	6	3	-	6	3	-	6	3	9	6	3
PSL-BG / No. of mets	-	154.8	546.7	-	140.3	555.5	-	38.5	343.6	63.8	108.0	526.3
Ratio	-	1	3.53		1	4.0	-	1	8.92	1	1.69	8.26

Table 5.1 Densitometric analysis of the cleavage products of GFP-2A.

The four translations of GFP-2A used in these experiments were analysed to see whether the ratios of cleavage products were constant. In one case (Figure 5.4.3) the cleavage was incomplete and the percentage of uncleaved GFP-2A was calculated. The analysis was performed on a phosphorimage visualisation of the SDS-PAGE gels. Analysis was performed using Mac Bas version 2 software. The PSL-BG values are the relative intensity values after a background count has been subtracted. As labelling is *via* radiolabelled methionine the number of methionine codons in each protein must be taken into account. To correct for the differing numbers of methionine codons in each protein product the PSL-BG values are divided by the number of methionine codons in each protein product. This gives a value proportional to the number of moles of each protein product. Ratios can then be calculated to show the amount of each of the protein products relative to one another.

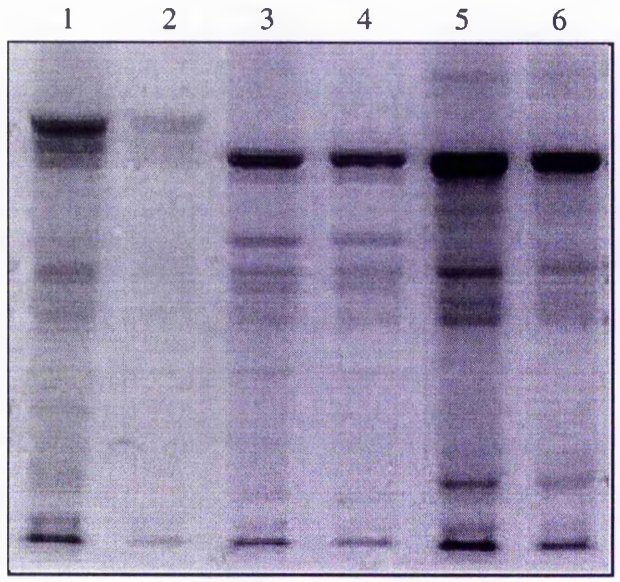
5.5 Processing *in trans* of reporter gene constructs.

In order to study the effect of processing *in trans* the coupled transcription translation reactions were incubated overnight at 30°C. Reactions were carried out as described in the protocol provided by the manufacturer and then the reaction was stopped by the addition of RNaseA (to a final concentration of 0.8 μg/ml) and cycloheximide (to a final concentration of 0.6 μg/ml). The reaction mixture was overlaid with mineral oil and incubated overnight at 30 °C. SDS-PAGE loading buffer was added to the reaction after the overnight incubation to stop any further degradation. The products were then analysed by SDS-PAGE. The first 90 minute incubation allows *cis* processing to take place and the overnight incubation gives time for *trans* processing reactions to take place between the products of the *cis* processing reactions.

It can be seen from the translation products in figure 5.5.1 that no further processing occured in the overnight reactions. There appears to have been some degradation of products especially in the overnight incubation of pEB4. Further experiments were not carried out to try and improve the quality of the overnight reactions as the system was found to be unsuitable for this type of analysis from other experiments.

The addition of a reducing agent was not tried and the incubation times and concentrations of reactants were not varied. It may be more successful if tried in the RRL system rather than the WGE.

Figure 5.5.1 Processing of reporter polyprotein constructs *in trans*.. TnT wheatgerm extract systems were programmed with plasmid DNA and incubated at 30°C for 90 minutes. The reaction was then stopped and overlaid with mineral oil. The reactions were then incubated overnight to allow *trans* processing to occur.



Lane 1 pEB4
Lane 2 pEB4 overnight
Lane 3 pEB15
Lane 4 pEB15 overnight
Lane 5 pEB25
Lane 6 pEB25 overnight

5.6 Conclusions.

The plasmids constructed to perform the *trans* processing experiments were all translated separately and gave products of the expected sizes with the exception of 3C' pGEM-T.

- **GFP-2A cleaves *in cis* but not *in trans*.** The artificial polyprotein GFP-2A cleaves *in cis* to produce GFP and $2A^{pro}$. The ratio of products is not constant and there is also some uncleaved GFP-2A in some cases. The reason for the difference in the processing ability of $2A^{pro}$ is not known and may be a result of the system used. However, the artificial polyprotein does produce significant amounts of the cleaved products in all cases and the amount of uncleaved polyprotein is less than 10% when it is seen. The 2A protease is capable of cleaving itself from the GFP fairly efficiently *in cis*. This provides evidence that a functional $2A^{pro}$ is present and is capable of cleaving itself in the RRL TnT system. The 2A protease from HRV14 cleaves itself at its N-terminus in the virus and it does the same in the artificial polyprotein GFP-2A.

- **$2A^{pro}$ does not cleave $3CD^{pro}$ *in trans* in the RRL TnT system.** The use of GFP-2A to provide the $2A^{pro}$ for *trans* processing meant that a functional GFP was known to be in the reaction. However, even though $2A^{pro}$ was known to be active *in cis* it did not process HRV14 $3CD^{pro}$ *in trans*. This is a reaction known to occur in human rhinovirus 1A and the non-occurrence in this experiment seems to be more likely to be due to the system used than the lack of processing of this type in HRV14. The original strategy of translating the plasmids separately was designed to allow the accumulation of protease and substrate before *trans* processing to mimic the situation in the virus. This also allowed the protease to be synthesised in an unlabelled reaction which meant that the cleavage products visualised on the gel would only be those resulting from cleavage of the substrate and the products would be more easily identifiable.

- **Varying the amounts of substrate and protease does not affect *trans* processing in this system.** The variation of the amounts of protease and substrate in the overnight incubations was intended to abolish any inhibitory effects that an excess of one component may have. Initially an excess of the unlabelled protease was added to ensure that processing would take place but when this was unsuccessful equal volumes of protease and substrate were used. Lowering the amount of protease present did not have any effect on the processing of the substrate.

- **Addition of a reducing agent did not affect the lack of *trans* processing in this system.** The addition of DTT as a reducing agent did have any effect on the *trans* processing reaction but again the system was not ideal for this type of processing reaction.

- **Co-translation of plasmids did not stimulate *trans* processing in this system.** The two plasmids were also co-translated to see if this would allow *trans* cleavage. This strategy is not as close to the strategy employed by the virus but may allow *trans* processing to occur when lower levels of protein are present. As the translation products of both reactions are labelled in this case it is more difficult to distinguish cleavage products of the substrate.

- **3CD and 3C' do not process P1 *in trans.* in RRL TnT system.** The observations made above about the *trans* processing of $3CD^{pro}$ by $2A^{pro}$ also apply to the processing of P1 by $3CD^{pro}$ and 3C'. Processing did not occur when the separate synthesis strategy was applied or when the plasmids were co-translated. The addition of DTT as a reducing agent was also unsuccessful. Again the fault was thought to be with the system used not the plasmid DNA used to program the system.

  The lack of translation product from the plasmid encoding 3C' was thought to be due to an error in the promoter sequence of the vector used. The nucleotide sequence of

the 3C' insert was checked after ligation into the vector and found to be correct so the lack of translation products were not due to any mutations in the nucleotide sequence.

- **The reporter polyprotein system does not show any signs of *trans* processing products after an overnight incubation. The** incubation of the reporter polyprotein constructs overnight to allow *trans* processing reactions to occur was not successful. The translation products were the same as in the sample from the initial incubation which was to allow *cis* processing to occur. The only difference is that overnight incubations appear to have been degraded to some degree. This is probably due to the prolonged incubation which would allow the accumulation of proteases from the reaction components.

In general, the incubation of the plasmids in pairs to observe *trans* processing was not very successful. As all the plasmids had been translated successfully separately the fault was thought to be with the system used rather than with the quality of the DNA used to program the reactions.

Summary.

- HRV14 $2A^{pro}$ cleaves GFP-2A *in cis* to give GFP and $2A^{pro}$.
- N-terminal cleavage by HRV14 $2A^{pro}$.
- Cleavage is total in most cases but there is evidence of some uncleaved GFP-2A in some reactions. In this case cleavage has been calculated to be 91%.
- Whether the protease is in excess or at the same volume as the substrate *trans* processing does not occur.
- Addition of DTT does not improve cleavage *in trans*.
- Co-translation does not allow *trans* cleavage.
- *Trans* processing of $3CD^{pro}$ to give the alternative cleavage 3C' product does not work in this system.
- This system is not suitable for monitoring *trans* processing between parts of the HRV14 genome.
- *Trans* processing is not seen in any of the reporter gene constructs. This is probably due to the unsuitability of the system for this type of analysis as stated earlier.

5.7 Discussion.

The lack of *trans* processing observed in the above experiments leads to the conclusion that the coupled RRL system is unsuitable for this type of analysis. Even with the addition of a reducing agent to maintain the oxidation of the protease no *trans* processing was seen. Using separate transcription and translation systems may be more successful.

The protease genes used in these experiments will be cloned into bacterial expression vectors to allow higher concentrations of the proteolytic enzymes to be added to the system. The TnT system produces very small amounts of protein and it may be the case that this is not enough to process the substrates *in trans*. The substrates produced are in these experiments should be suitable for processing if the enzymes are present in higher concentrations. As there are no positive controls working the conclusions drawn from these experiments are minimal and merely serve to highlight the unsuitability of the system for this kind of analysis.

6. Discussion

The replication of a virus requires proteins to be produced in different amounts and at different times in the replication cycle. There are various strategies which viruses use to accomplish this. Some produce subgenomic RNAs which encode proteins required at a certain point. For example, the capsid proteins are usually required at a later stage in replication to package the genome and encoding these in a subgenomic RNA means that they will be produced when required.

In picornaviruses the single polyprotein strategy means that all proteins should be produced in a 1:1 ratio and at approximately the same time. The regulation of the proteolytic enzymes means that some sort of differential control of processing is present. The capsid proteins in rhinovirus and poliovirus are know to be processed by $3CD^{pro}$ *in trans*. This requires the release of $3CD^{pro}$, as opposed to alternative processing products, from the polyprotein before P1 processing can begin, and, therefore some sort of regulation is operating. The viral polyprotein has one processing pathway which yields $3C^{pro}$ and $3D^{pol}$ and an alternative pathway to give $3CD^{pro}$.

The *in vitro* system used in these experiments is limited by the lack of viral and, perhaps, cellular factors which will not allow the same regulation mechanisms to operate as in the virus. At this stage, therefore, direct comparisons with virus-infected cells are not possible. The reporter polyprotein system has been designed to allow the addition of other factors back into it which will allow more direct comparisons. Certain inferences can, however, still be made as to the implications of these results for processing of the viral polyprotein.

6.1 HRV14 $3C^{pro}$ cleavage in a reporter polyprotein system.

The aim of this project was to construct artificial reporter polyproteins which would allow cleavage by HRV14 $3C^{pro}$ to be monitored without the influence of viral and cellular factors. The reporter genes used would also allow investigation into the monitoring of cleavage by antibiotic resistance.

cellular factors. The reporter genes used would also allow investigation into the monitoring of cleavage by antibiotic resistance.

In the reporter polyprotein [$Kan^R3C^{pro}Tet^R$] cleavage by $3C^{pro}$ was monitored to establish a "default" $3C^{pro}$ processing pathway. That is to say that the reporter polyprotein system will show how $3C^{pro}$ cleaves when it is not under the influence of any viral or cellular factors, *i.e.* a "default" pathway. Results showed that $3C^{pro}$ cleaves preferentially at its N-terminus in this system to yield [Kan^R] and [$3C^{pro}Tet^R$] (Figure 6.1). There is a lesser degree of cleavage at the C-terminus of $3C^{pro}$ to yield [Kan^R3C^{pro}] and [Tet^R]. If this state of affairs was directly applied to the virus polyprotein situation then N-terminal cleavage to yield $3CD^{pro}$ would be preferential to C-terminal cleavage to ultimately yield $3C^{pro}$. However, there are likely to be other factors affecting cleavage by $3C^{pro}$ in virus polyprotein processing. These results indicate which cleavage pathway $3C^{pro}$ would follow if it was not influenced by other viral and cellular factors - exactly what the artificial polyprotein [$Kan^R3C^{pro}Tet^R$] was designed to show. Further experiments with addition of viral and cellular factors will show whether these have an influence on cleavage by $3C^{pro}$ and may be used by the virus to 'switch' production from $3C^{pro}$ to $3CD^{pro}$ or *vice versa*.

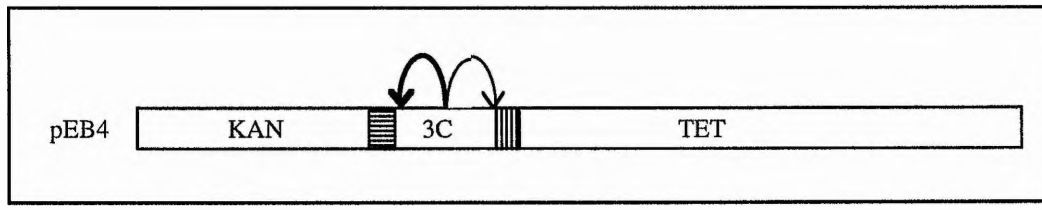


Figure 6.1 Summary of cleavages in pEB4. The bold arrow indicates the major (preferential) cleavage and the other minor cleavage. The shaded areas indicate additional sequences included to maintain $3C^{pro}$ cleavage sites.

6.1.1 Expression of self-processing antibiotic resistance polyproteins in *E. coli.*

Expression of antibiotic resistance in *E. coli* was investigated as a possible means for monitoring cleavage in the reporter polyprotein system. It was known that the kanamycin resistance gene would not function with a C-terminal extension (M. Ryan, pers comm) and it was hoped that the tetracycline resistance gene product would not function with an N-terminal extension so that cleavage at either end of $3C^{pro}$ would allow expression of antibiotic resistance to one or other of the antibiotics used. This did not seem to be the case, with all constructs showing tetracycline resistance phenotype (see section 4.1.5). This may imply that the polyprotein is either; not cleaving at all in this system, the alterations made to the C-terminus of the kanamycin resistance gene having rendered it inactive, or, cleavage is occurring preferentially at the C-terminus of $3C^{pro}$. If the latter is true then these results conflict with those from the TnT system.

Further work needs to be carried out to ascertain whether this strategy is suitable for determining the site, (N- or C-terminal), of cleavage . The lack of a positive control for tetracycline resistance needs to be remedied and more constructs need to be assessed in the system to see whether the results are conclusive. If this can be achieved then the system could be used to monitor cleavage by proteases in artificial polyproteins rather than the TnT system. The genetic screening system would also be useful in the analysis of large numbers of mutations. Screening of mutations using the TnT system is a laborious process and screening using antibiotic resistance of polyproteins would be less time consuming and less expensive.

6.1.2 Endogenous processing of $3CD^{pro}$ and 3ABC

Plasmids containing $3CD^{pro}$ and 3ABC were constructed as controls for the reporter system but were also used to monitor endogenous processing. $3CD^{pro}$ seems to be stable in the TnT reaction and does not cleave to give $3C^{pro}$ and $3D^{pol}$. The comparison of the translation products with the appropriate controls led to the identification of the other translation products from the plasmid encoding $3CD^{pro}$ as internal initiation products. This lack of endogenous processing is consistent with results seen previously *in vitro*

(Harris *et al.*, 1992) and *in vivo* (Porter, 1993). In the systems studied previously proteolysis of $3CD^{pro}$ to $3C^{pro}$ and $3D^{pol}$ is slow compared to processing at other Q-G sites in the polyprotein but does occur, indicating that $3CD^{pro}$ has a role distinct from that of its cleavage products. This did not seem to be the case in this system studied here. Previous studies have been made in virus-infected cells where a higher degree of *trans* processing is seen. This may explain the lack of endogenous $3CD^{pro}$ cleavage in TnT systems.

There is quite a high degree of internal initiation within $3CD^{pro}$ on translation but altering the kozak consensus sequence of the methionine codon or the full length transcript may alter this. The salt concentration of the TnT system is known to affect initiation. A higher salt concentration (Mg^{2+} and K^{+}) and lower concentration of RNA is known to improve translation efficiency from the initial methionine and should reduce level of internal initiation products.

Intramolecular *(in cis)* processing in 3ABC results in 3AB and $3C^{pro}$, $-3C^{pro}$ cleaving at its N-terminus. 3ABC is seen as a product of processing in the virus but it is not particularly stable and does not have any known function apart from that of a protein precursor . 3AB however, is more stable and cleavage to 3A and 3B is rarely seen in *in vitro* systems unless all other processing routes are blocked (Parks *et al.*, 1989). This cleavage is performed by an unknown mechanism in the virus which is obviously not active in *in vitro* systems. There is a small amount of processing to yield 3BC and 3A.

6.1.2.1 Further work

In order to investigate the possibility that the nature of the reporter sequences used affects cleavage, construction of a plasmid encoding $[Tet^{R}3C^{pro}Kan^{R}]$ would be necessary. This would, of course, not be suitable for using in the *E. coli* antibiotic expression system as there would be no possible C-terminal extension for Kan^{R}. However use of this plasmid may indicate whether Tet^{R} will be disabled by the addition of sequences to its C-terminus.

The introduction of mutations into the $3C^{pro}$ sequence would allow their effect on cleavage to be studied. Some mutations have been made in the $3C^{pro}$ of poliovirus which cause 'over-cleavage', *i.e.* cleavage occurs at either the N- or the C-terminus of $3C^{pro}$ preferentially. Such mutations could be introduced into $3C^{pro}$ within the reporter polyprotein system either as $3C^{pro}$ mutants or $3CD^{pro}$ mutants and the cleavage products analysed by SDS-PAGE. The addition of bacterially expressed $3C^{pro}$ or $3CD^{pro}$ to these mutant reporter systems would show whether the bacterially expressed protease could rescue the mutant and process the reporter polyprotein.

6.2 Processing with 3CD^{pro} in the reporter polyprotein leads to an unexpected cleavage downstream.

6.2.1 Processing in pEB15-Cleavage downstream of 3D^{pol} by 3C^{pro}

The artificial polyprotein [Kan^{R}3CD^{pro}Tet^R], encoded by the plasmid pEB15, has been shown to cleave to produce the products Kan^R, 3C^{pro}, 3D^{pol} and Tet^R (Figure 6.2) and possibly [Kan^{R}3C^{pro}]. There is no product representing [3D^{pol}Tet^R] which is consistent with cleavage at the C-terminus of 3D^{pol}. The cleavage between 3D^{pol} and Tet^R was unexpected as it does not mimic a cleavage which occurs in the virus. There is no coding sequence downstream of 3D^{pol} in the rhinoviral polyprotein and therefore there are no 3C^{pro}-mediated cleavage events downstream of 3D^{pol}. The cleavage of 3D^{pol} from Tet^R in the reporter polyprotein is therefore a novel event and it is surprising that 3C^{pro} will cleave at this junction. Inspection of the 3D^{pol} amino acid sequence has shown that there is a putative 3C^{pro} cleavage site near the C-terminus of 3D^{pol} (Figure 6.3). Alignment of other picornavirus 3D^{pol} amino acid sequences has shown that this site is conserved (Figure 6.4).

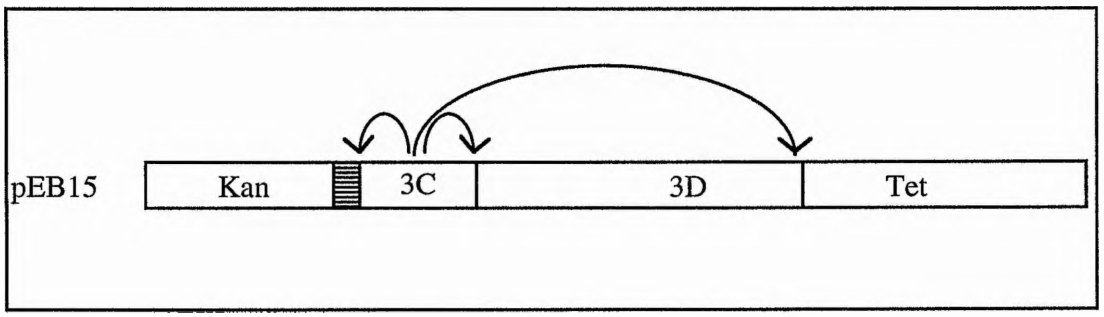
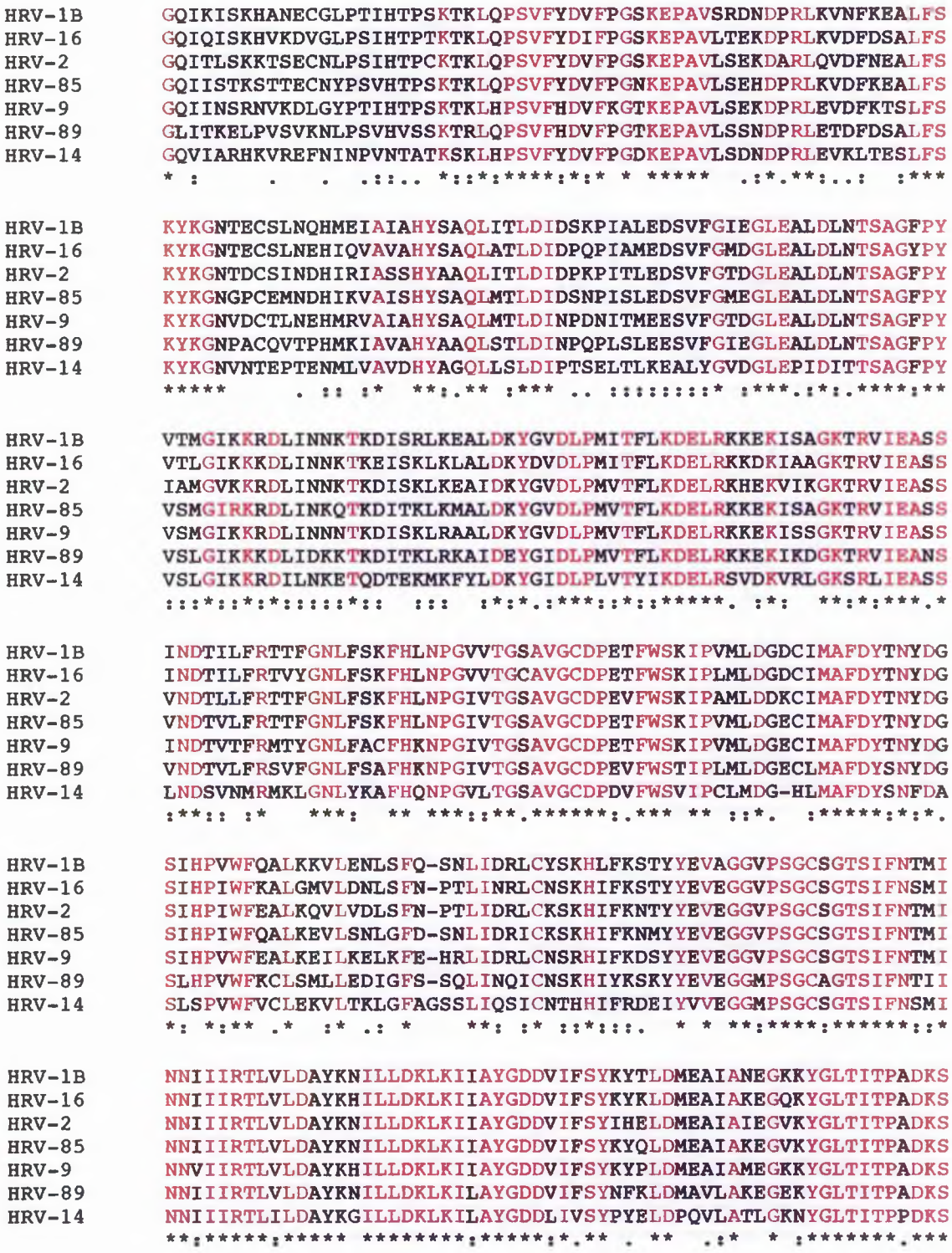


Figure 6.2 Summary of cleavages in pEB15. The shaded region indicates 3B sequences which have been incorporated to maintain the 3C^{pro} 3B/C cleavage site.

Figure 6.3 Amino acid sequence of HRV14 3CD^{pro}. The residues shown in blue are the 3C/D cleavage site, those shown in green are the alternative 3C'/3D' cleavage site and those shown in red are possible cleavage sites at the C-terminus of 3D^{pol} and in the Tet^{R} sequence.

```
->3Cpro
GPNTEFALSL LRKNIMTITT SKGEFTGLGI HDRVCVIPTH AQPGDDVLVN

GQKIRVKDKY KLVDPENINL ELTVLTLDRN EKFRDIRGFI SEDLEGVDAT

LVVHSNNFTN TILEVGPVTM AGLINLSSTP TNRMIRYDYA TKTGQCGGVL
                                   ->3Dpol
CATGKIFGIH VGGNGRQGFS AQLKKQYFVE KQGQVIARHK VREFNINPVN

TATKSKLHPS VFYDVFPGDK EPAVLSDNDP RLEVKLTESL FSKYKGNVNT

EPTENMLVAV DHYAGQLLSL DIPTSELTLK EALYGVDGLE PIDITTSAGF

PYVSLGIKKR DILNKETQDT EKMKFYLDKY GIDLPLVTYI KDELRSVDKV

RLGKSRLIEA SSLNDSVNMR MKLGNLYKAF HQNPGVLTGS AVGCDPDVFW

SVIPCLMDGH LMAFDYSNFD ASLSPVWFVC LEKVLTKLGF AGSSLIQSIC

NTHHIFRDEI YVVEGGMPSG CSGTSIFNSM INNIIIRTLI LDAYKGIDLD

KLKILAYGDD LIVSYPYELD PQVLATLGKN YGLTITPPDK SETFTKMTWE

NLTFLKRYFK PDQQFPFLVH PVMPMKDIHE SIRWTKDPKN TQDHVRSLCM
                                                 3D<-->TET
LAWHSGEKEY NEFIQKIRTT DIGKCLILPE YSVLRRRWLD LFLQKSNNAL

IVILGTVTLD AVGIGLVMPV LPGLLRDIVH SDSIASHYGV LLALYALMQF
```

Figure 6.4 CLUSTAL X (1.5b) multiple sequence alignment of Rhinovirus 3D^{pol} amino acid sequences. Conserved residues are shown in red and the putative 3C^{pro} cleavage site in blue



```
HRV-1B      TEFKKLDYNNVTFLKRGFKQDEKHTFLIHPTFPVEEIYESIRWTKKPSQMQEHVLSLCHL
HRV-16      SEFKELDYGNVTFLKRGFRQDDKYKFLIHPTFPVEEIYESIRWTKKPSQMQEHVLSLCHL
HRV-2       NTFVKLDYSNVTFLKRGFKQDEKYNFLIHPTFPEDEIFESIRWTKKPSQMHEHVLSLCHL
HRV-85      SEFKQLNYNNVTFLKRGFKQDDKYQFLIHPTFPIEEIQESIRWTKKPSQMQEHVLSLCHL
HRV-9       DTFKKLDYDSVTFLKRGFKQDSKYPFLIHPTFPVNEIHESIRWTKKPSQMQEHVLSLCHL
HRV-89      DVFQELTYKNVTFLKRGFRADERHSFLIHPTFPVAEIHDSIRWTKNPSCMQEHVLSLCHL
HRV-14      ETFTKMTWENLTFLKRYFKPDQQFPFLVHPVMPMKDIHESIRWTKDPKNTQDHVRSLCML
              * :: : .:***** *: *.:. **:**.:*  :* :******.*.  ::** *** *

HRV-1B      MWHNGRKVYEDFSSKIRSVSAGRALYIPPYDLLKHEWYEKF
HRV-16      MWHNGPEIYKDFETKIRSVSAGRALYIPPYELLRHEWYEKF
HRV-2       MWHNGRDAYKKFVEKIRSVSAGRALYIPPYDLLLHEWYEKF
HRV-85      MWHNGKDVYKQFEQKIRSVSAGRALYIPPYELLLHEWYEKF
HRV-9       MWHNGRDVYKEFESGIRSVS-GRALYIPPYELLLHEWYEKF
HRV-89      MWHNGRHAYQEFIKGIRSVSAGRALYIPAYEVLEHEWYEKF
HRV-14      AWHSGEKEYNEFIQKIRTTDIGKCLILPEYSVLRRRWLDLF
             **.* . *:.*   **:.. *:.* :* *.:* :.* : *
```

The cleavage between $3D^{pol}$ and Tet^R suggests that $3C^{pro}$ in some way surveys the C-terminal region of $3D^{pol}$. The fact that it does not normally perform a cleavage at this position does not seem to matter. The gln-lys site in the C-terminal region of $3D^{pol}$ (FIQKTR) is probably not the site used in the 3D/TetR cleavage since this site is present in the virus seqeunces. There are other possible sites at the C-terminus of $3D^{pol}$ such as the Q-K site created at the $3D^{pol}/Tet^R$ junction. *Cis* cleavage sites are very resistant to mutagenesis whereas *trans* sites are much more susceptible. This explains the deviation of the sequence of this site from the pre-determined (E, Q), (G, S, A). The artificial polyprotein provides a situation for the protease to cleave at these sites, not usually utilised in the virus, by incorporating a coding sequence downstream of $3D^{pol}$. This is similar to the genome organisation found in another virus family, the caliciviruses.

6.2.2 Caliciviruses

There are two different genome structures present in the caliciviridae (Figure 6.5). Rabbit haemorrhagic disease virus (RHDV) has one long ORF (7 kb) encoding both structural and non-structural proteins and a second short ORF (351 nucleotides) encoding a single protein of unknown function. The other type of virus has two ORFs of approximately 5 kb and 2.2 kb which encode the putative non-structural proteins and capsid proteins respectively. They also have a third ORF in the 3' part of the genome which is analogous to ORF2 in RHDV. Feline calicivirus (FCV) and two human caliciviruses, Norwalk virus and Southampton virus (SV), have this type of genome organisation. Both types have a 3C-like cysteine protease which cleaves at various sites throughout the polyprotein to release the mature protein products. Unlike picornaviruses there is a single proteolytic entity in the caliciviruses which cleaves at all proteolytic sites in the polyprotein. Due to their similarity to picornaviruses and related viruses (potyviruses, nepoviruses and comoviruses) caliciviruses have been proposed as members of the picornavirus superfamily (Goldbach and Wellink, 1988) but this has recently been revised.

The calicivirus protease has similarities to both the 2A and 3C proteases of picornaviruses. The position of the protease in the genome, *i.e.* upstream of the viral polymerase is similar to that of $3C^{pro}$ in picornaviruses and major determinants of $3C^{pro}$ cleavage specificity are conserved in the protease sequence. However, amino acid sequence comparisons show the calicivirus protease to be shorter than the picornavirus $3C^{pro}$ and closer in size to the picornavirus $2A^{pro}$. The protease in caliciviruses acts at sites which are C-terminal to it to cleave the capsid protein precursor in both genome organisations. Cleavage occurs at a range of sites but the preferred dipeptide in SV is Q-G which is the same as in picornaviruses. In RHDV cleavage of the protease occurs at a Q-G site at its N-terminus and at a Q-T site at its C-terminus. Cleavage is dependent on the amino acid residue at position 1. Only glutamate, aspartate and glutamine are tolerated at this position. ORF1 is processed by the protease at seven cleavage sites to give eight final protein products (Wirblich*et al.*, 1996). ORF2 was found to encode a structural protein of 10 kDa. Other cleavage sites recognised by 3C in RHDV include E-G and E-T.

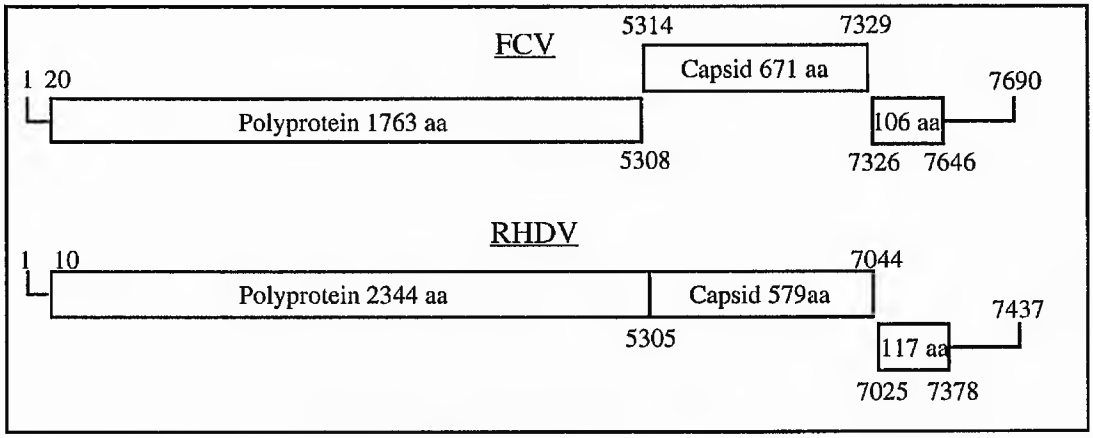


Figure 6.5 Diagrammatic representation of the two genome organisations found in caliciviruses. Boxes indicate regions translated from each ORF.

The cleavage between $3C^{pro}$ and Tet^R in the polyprotein [$Kan^R3C^{pro}Tet^R$] is similar to that carried out by the caliciviral 3C protease to separate the structural and non-structural proteins.

The caliciviral protease cleaves between the viral polymerase and the capsid protein precursor at a glutamic acid-glycine (E-G) site (Figure 6.6). It has been shown to be

cysteine protease and to have similarities with the picornaviral proteases (see section 6.2.2).

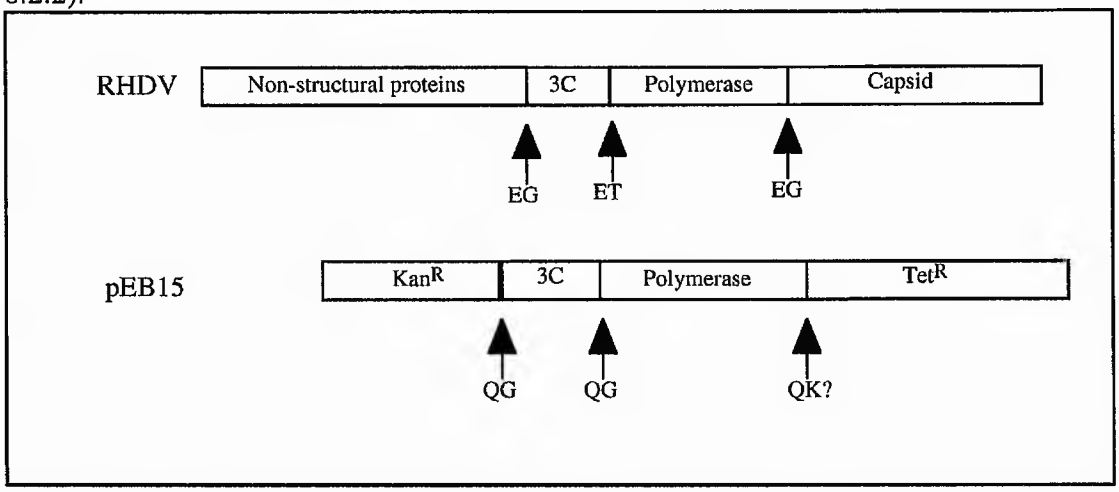


Figure 6.6 Comparison of cleavages downstream of $3C^{pro}$ in RHDV, a calicivirus, and pEB15 [$Kan^R3CD^{pro}Tet^R$]. Arrows indicate cleavage sites and the amino acid sequence of these sites is shown below.

6.2.3 Comparison of picornavirus and calicivirus genomes.

The genome organisation of the caliciviruses and picornaviruses is similar in many ways, as can be seen from the diagram below (Figure 6.7). The organisation of the part of the genome coding for the non-structural protein is similar; putative helicase, protease, polymerase. There is only one proteolytic entity in caliciviruses compared to the two, $2A^{pro}$ and $3C^{pro}$, in picornaviruses. The caliciviral protease does cleave at sites similar in nature to those of the $3C^{pro}$ and has been categorised as a cysteine protease. It is similar in size though to the $2A^{pro}$. The position of the structural proteins is one of the major differences between the genomes of the two classes but it would appear from results of the artificial polyprotein studies that the picornaviral $3C^{pro}$ could be capable of cleaving proteins from the C-terminus of the polyprotein as the caliciviral protease does.

The comparison shown in the diagram below is only to one type of calicivirus but the organisation of the non-structural proteins in two types of calicivirus genomes are similar. The major difference is in the encoding of the structural proteins (Figure 6.7). In the type of calicivirus not shown in comparison with poliovirus, exemplified by FCV,

there is a frameshift between the termination codon for the first ORF encoding the non-structural and the initiation codon of the second ORF encoding the structural proteins. The structural proteins are however, still downstream of the structural proteins in the genome.

In both picornaviruses and caliciviruses there are possible $3C^{pro}$ cleavage sites which remain uncleaved during processing of the viral polyprotein. This is also the case in the artificial reporter polyprotein. This implies that there are other factors influencing cleavage by the proteolytic enzyme rather than just the amino acid sequence of the cleavage site. It has been demonstrated that sequences distal to the cleavage sites have an effect on cleavage and this is probably due to the conformation of the protein due to these residues. There must be a preferred structure at which the protease will cleave.

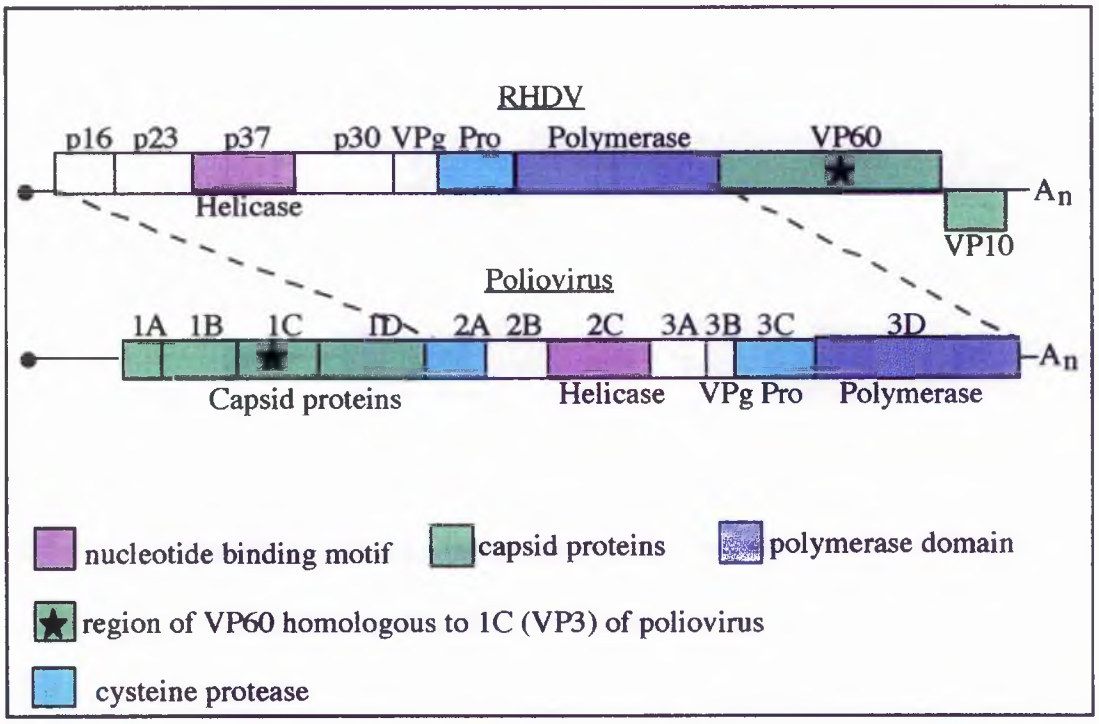


Figure 6.7 Comparison of genome structure of RHDV and poliovirus. The open reading frames are shown by boxes. The two genomes are not to scale.

6.2.4 Implications of the cleavage by $3C^{pro}$ in the downstream region of the polyprotein.

The downstream cleavage between $3D^{pol}$ and Tet^R in the polyprotein carried out by $3C^{pro}$ may have implications for virus evolution. The similarity to the cleavage performed by the caliciviral protease may point to a common ancestry between these two groups of viruses. The similarity in their genome organisation has already been discussed and this cleavage seen in the reporter polyprotein indicates that the picornavirus genome could encode the (structural) protein(s) downstream of the non-structural proteins. Perhaps an ancestral form evolved to have its structural proteins upstream of the non-structural proteins but the protease has still retained its ability to cleave at suitable sites in coding sequences downstream of $3D^{pol}$.

If $3C^{pro}$ is capable of cleaving at sites downstream of itself regardless of whether they are situated in the same position in the virus there may be implications for biotechnological uses of $3C^{pro}$ in the production of multiple proteins from artificial polyproteins. Current strategies allow for the production of two proteins from the same polyprotein but $3C^{pro}$ may be capable of cleaving at multiple sites to release multiple 'mature' proteins. One of the advantages of using polyproteins for expression of in biotechnology is that proteins are produced in equal amounts, as in the virus. This is a problem with the use of separate promoters for each gene especially where the introduction of a trait relies on the expression of two or more independent genes.

It would appear that $3C^{pro}$ does not need to have viral sequences to cleave as long as the cleavage site is one which is recognises, *i.e.* of the pattern (E, Q), (G, S, A). $3C^{pro}$ will cleave itself away from the proteins at its C-terminus but also process at subsequent sites downstream to release proteins. The number of downstream sites which the protease will process is not known and needs to be investigated.

The problem with this, if it is found to be applicable to biotechnological uses, is that $3C^{pro}$ is known to target cellular sequences which may be a problem. Processing, is however, *in cis* so it will be quick and efficient if can be used.

6.2.5 Further work

In order to see if this strategy works *in vivo* sequences could be engineered between the end of 3Dpol and the 3'NCR and then the recombinant could be transfected into cells to observe whether viable virus is obtained. Insertions which have previously been made in the 3' NCR have been unstable and recent work has shown that poliovirus is capable of replicating (at low levels) without the 3' NCR . To further establish the link between picornaviruses and caliciviruses the structural proteins could be moved to the 3' end of the genome to see if they are cleaved and produce viable virus particles. This may not be possible as there are still unidentified packaging signals which may be interrupted by this process. It has been demonstrated that recombinant genomes lacking P1 can be packaged by helper viruses which supply P1 i*n trans* so the position of the capsid precursor in the genome may not be important as long as the proteins are processed correctly.

6.2.6 Processing in pEB25

In order to monitor the effect of upstream sequences on cleavage by $3C^{pro}$ the viral precursor protein 3ABC was ligated between the reporter genes. On translation the resulting artificial polyprotein [Kan^R3ABCTet^R] gave rise to [Kan^R3AB], Kan^R, 3AB and [$3C^{pro}Tet^R$] (Figure 6.8). This indicates that $3C^{pro}$ is cleaving at its N-terminus and at the N-terminus of 3AB. The cleavage at the N-terminus of 3AB mimics the primary cleavage in the virus between P2 and P3. The cleavage at the N-terminus of $3C^{pro}$ and lack of cleavage at the C-terminus imply that 3AB has a function in up-regulating cleavage at the N-terminus of $3C^{pro}$. In [$Kan^R3C^{pro}Tet^R$] the N-terminal cleavage predominated over cleavage at the C-terminus, as is the case when 3AB is present. The presence of *any* extension at its N-terminus regulates N-terminal cleavage by $3C^{pro}$ but 3AB seems to have an even greater effect than Kan^R.

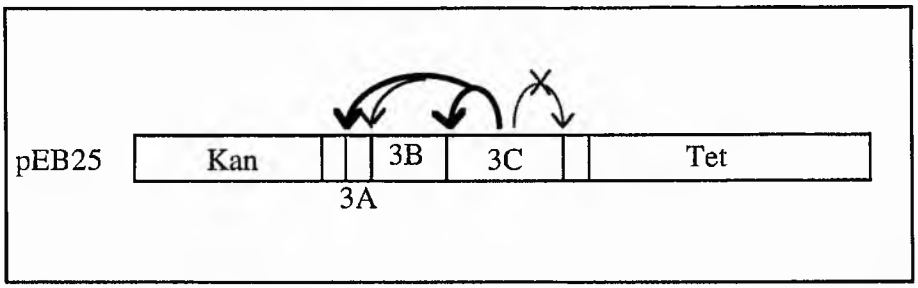


Figure 6.8 Summary of cleavages in pEB25. Shaded regions indicate additional viral sequences incorporated in the polyprotein to maintain cleavage sites

6.2.7 Effect of N- or C-terminal extension on cleavage by $3C^{pro}$.

Inclusion of the proteins 3AB and $3D^{pol}$ had a regulatory effect on cleavage by $3C^{pro}$. The position of these regulatory sequences relative to $3C^{pro}$ may be important and their position in the viral part of the reporter polyprotein should be investigated. It may be that $3D^{pol}$ would also function upstream of $3C^{pro}$ and 3AB would also function downstream but this has not yet been investigated.

Cleavage at the C-terminus of $3C^{pro}$ seems to function very efficiently without any influence from viral sequences but cleavage at the N-terminus is more reliant on the influence of other factors. The construction of an artificial polyprotein encoding [Kan^R3ABCDTet^R] would be useful to observe the effect of both putative regulatory sequences together. It would also show if 3AB has any effect on the additional downstream cleavages by $3C^{pro}$ observed in [Kan^R3CD$^{pro}Tet^R$].

The addition of 3AB and/or $3D^{pol}$ *in trans* to [$Kan^R3C^{pro}Tet^R$] would be an experiment which would provide further insight into the role of these sequences in up-regulating cleavage by $3C^{pro}$.

6.3 Processing *in trans*

The coupled transcription translation system proved to be unsuitable for monitoring processing *in trans* in these experiments. Previous work has shown the system to function adequately but this was not the case here. It may be improved if separate transcription and translation experiments are carried out using an uncoupled system. There are known to be variations in the efficiency of the systems, both coupled and uncoupled, from batch to batch. This may account for the lack of activity in this set of experiments as opposed to previous successful ones. The addition of DTT to the *trans* reactions was not successful in enhancing the processing but this may have been as processing was not taking place at all due to the general unsuitability of the system. DTT prevents formation of disulphide bridges in proteins and this alteration in structure may affect active site of protein so that it becomes inactive. This may be another factor which affects the usefulness of DTT in this situation.

The improvement of this system would allow a great deal more to be elucidated about the interaction of $3C^{pro}$ with other viral and cellular factors and their effect on cleavage. It has been seen from reporter protein experiments that 3AB and $3D^{pol}$ appear to have some kind of regulatory function on cleavage by $3C^{pro}$. It would be useful to see if these viral factors could be added *in trans* to observe their effects and see if they still have the same up-regulatory effect.

The interaction of the 5' NCR with P1 and $3CD^{pro}$ is also one which could be investigated when the *trans* processing system is functional. As mentioned in the section above, the addition of other viral sequences *in trans* to the reporter polyprotein would also be informative.

The low levels of protease produced in the TnT system are not sufficient to allow *trans* processing to occur. The levels of substrate are good and can be used with proteases produced by an alternative method. The production of bacterially expressed enzymes will allow greater concentrations of the enzymes to be added to the system and cleavage can be monitored by synthesis of radiolabelled substrate.

A summary diagram of cleavages in the artificial reporter polyproteins constructed during this work is shown below (Figure 6.9).

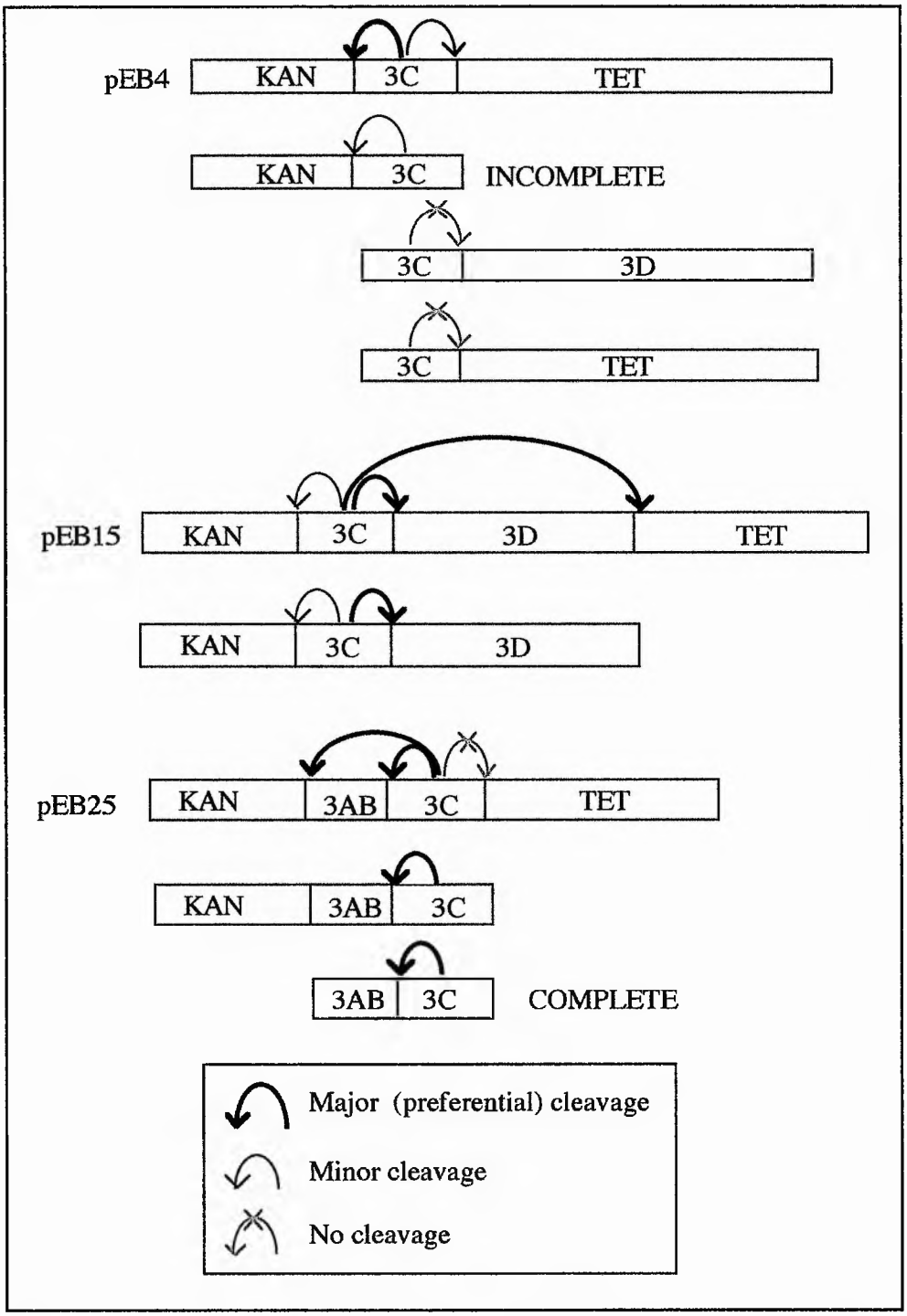


Figure 6.9 Summary diagram to show cleavages by $3C^{pro}$ in artificial reporter polyproteins

6.4 Concluding remarks

The construction of a reporter polyprotein system has been achieved and a default pathway for $3C^{pro}$ cleavage has been established. It appears that 3AB and $3D^{pol}$ have a regulatory effect on the "choice" of $3C^{pro}$ terminus to be cleaved.

The unexpected cleavage by $3C^{pro}$ at the C-terminus of $3D^{pol}$ in the reporter system has led to the postulation of possible evolutionary links with caliciviruses and plant viruses, may indicate another site which could be used in the construction of recombinant picornavirus genomes and could lead to possible uses for $3C^{pro}$ in artificial polyproteins in biotechnology.

Several new avenues have been opened from this preliminary study of cleavage by $3C^{pro}$ which will hopefully provide further information about cleavage by $3C^{pro}$ and elements which have an influence on this process.

Appendix

Table of plasmids constructed

Plasmid	Insert	Vector	AUG	STOP
pEB1	Tet^R	pGEM-T	√	√
pEB1.1	Tet^R	pGEM5zf(+)	√	√
pEB1.2	Tet^R	pGEM5zf(+)	X	√
pEB2	Kan^R	pGEM5zf(+)	√	√
pEB2.1	Kan^R	pGEM5zf(+)	√	X
pEB3	HRV14 $3C^{pro}$	pGEM5zf(+)	√	√
pEB3.1	HRV143C^{pro}	pGEM5zf(+)	X	X
pEB4	[$Kan^R3C^{pro}Tet^R$]	pGEM5zf(+)	√	√
pEB4.2	[$Kan^R3C^{proC146A}$ Tet^R	pGEM5zf(+)	√	√
pEB5	[$3C^{pro}Tet^R$]	pGEM5zf(+)	√	√
pEB6	HRV14 $3C^{pro}$	pRSETA	√	√
pEB7	HRV14 $3CD^{pro}$	pRSETA	√	√
pEB8	HRV14 P1	pMF16c	√	√
pEB13	[Kan^R3C^{pro}]	pGEM5zf(+)	√	X
pEB13.1	[Kan^R3C^{pro}]	pGEM5zf(+)	√	√
pEB14	[Kan^R3C^{pro}]	pGEM5zf(+)	√	√
pEB15	[$Kan^R3CD^{pro}Tet^R$]	pGEM5zf(+)	√	√
pEB15.2	[Kan^R3CD^{pro}]	pGEM5zf(+)	√	√
pEB16	HRV14 3AB	pGEM5zf(+)	√	√
pEB19	HRV14 3ABC	pGEM5zf(+)	√	√
pEB19.1	HRV14 3ABC	pGEM5zf(+)	X	X
pEB20	HRV14 $3CD^{pro}$	pGEM5zf(+)	√	√
pEB20.1	HRV143CD^{pro}	pGEM5zf(+)	X	√
pEB20.2	HRV14 $3CD^{pro}$	pGEM5zf(+)	X	X
pEB21	[$Kan^R3C^{pro}Tet^R$]	pBSIIKS+	√	√
pEB22	Tet^R	pBSIIKS+	√	√
pEB23	$3C^{pro}$	pBSIIKS+	√	√
pEB24	Kan^R	pBSIIKS+	√	√
pEB25	[$Kan^R3ABCTet^R$]	pGEM5zf(+)	√	√
pEB27	HRV14 $3D^{pol}$	pGEM5zf(+)	√	√

References

Allaire, M., Chernaia, M.M., Malcolm, B.A. and James, M.N.G. (1994) Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteases. *Nature 369*, 72-76.

Allaway, G.P. and Burness, A.T.H. (1986) Site of attachment of Encephalomyocarditis virus on human erythrocytes. *J.Virol. 59*, 768-770.

Alnakib, W., Higgins, P.G., Barrow, I., Batstone, G. and Tyrrell, D.A.J. (1987) Prophylaxis and treatment of rhinovirus colds with zinc gluconate lozenges. *J.Antimicrob. Chemother. 20*, 893-901.

Alsaadi, S., Hassard, S. and Stanway, G. (1989) Sequences in the 5' non-coding region of Human Rhinovirus 14 that affect *in vitro* translation. *J.Gen.Virol. 70*, 2799-2804.

Altman, S. (1984) Aspects of biochemical catalysis. *Cell 36*, 237-239.

Andino, R., Rieckhof, G.E., Achacoso, P.L. and Baltimore, D. (1993) Poliovirus RNA synthesis utilises an RNP complex formed around the 5' end of viral RNA. *EMBO J.12*, 3587-3598.

Andino, R., Rieckhof, G.E. and Baltimore, D. (1990a) A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell 63*, 369-380.

Andino, R., Rieckhof, G.E., Trono, D. and Baltimore, D. (1990b) Substitutions in the protease ($3C^{pro}$) gene of poliovirus can suppress a mutation in the 5' noncoding region. *J.Virol. 64*, 607-612.

Andries, K., Dewindt, B., Snoeks, J., Wouters, L., Moereels, H., Lewi, P.J. and Janssen, P.A. (1990) Two groups of rhinoviruses revealed by a panel of antiviral compounds present sequence divergence and differential pathogenicity. *J.Virol. 64*, 1117-1123.

Barbero, J.L., Buesa, J.M., Gonzalez de Buitrago, G., Mendez, E., Pez-Aranda, A. and Garcia, J.L. (1986) Complete nucleotide sequence of the penicillin acylase gene from *Kluyvera citrophila. Gene 49*, 69-80.

Bazan, J.F. and Fletterick, R.J. (1988) Viral cysteine proteases are homologous to the trypsin-like serine proteases:Structural and functional implications. *P.N.A.S.USA. 85*, 7872-7876.

Beadle, G.W. and Tatum, E.C. (1941) Genetic control of biochemical reactions in *Neurospora. P.N.A.S.USA 27*, 499-506.

Bella, J., Kolatkar, P.R., Marlor, C., Greve, J.M. and Rossmann, M.G. (1998) The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. *P.N.A.S. USA 95*, 4140-4145.

Bianco, A. and Spiteri, M.A. (1998) A biological model to explain the association between human rhinovirus respiratory infections and bronchial asthma. *Monaldi Arch. Dis. 53*, 83-87.

Bienz, K., Egger, D. and Pasamontes, L. (1987) Association of poliovirus proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography. *Virology 160*, 220-226.

Böck, A., Wirth, R., Schmid, G., Schumacher, G., Lang, G. and Buckel, P. (1983) The two subunits of penicillin acylase are processed from a common precursor. *FEMS Microbiology Letters 20*, 141-44.

Brown, P.K. and Taylor-Robinson, D. (1966) Respiratory virus antibodies in sera of persons living in isolated communities. *Bull WHO 34*, 895-900.

Calhoun, W.J., Jordan, W.S.J. and Gwaltney, J.M.J. (1974) Rhinovirus infections in an industrial population. V. Change in distribution of serotypes. *Am.J.Epidemiol. 99*, 58-64.

Callahan, P., Mitzutani, S. and R. J. Colonno, R.J. (1985) Molecular cloning and complete sequence determination of RNA complete genome of HRV14. *P.N.A.S.USA 82*, 732-6.

Casasnovas, J.M., Bickford, J.K. and Springer, T.A. (1998) The Domain Structure of ICAM-1 and the Kinetics of Binding to Rhinovirus. *J.Virol. 72*, 6244-6246.

Casasnovas, J.M., Stehle, T., Liu, J.-H., Wang, J.-H. and Springer, T.A. (1998) A dimeric crystal structure for the N-terminal two domains of ICAM-1. *P.N.A.S. USA 95*, 4134-4139.

Casasnovas, J.M. and Springer, T.A. (1994) Pathway of rhinovirus disruption by soluble intercellular-adhesion molecule-1 (ICAM-1) - an intermediate in which ICAM-1 is bound and RNA is released. *J. Virol. 68*, 5882-5889.

Chen, P., Tsuge, H., Almassy, R.J., Gribskov, C.L., Katoh, S., Vanderpool, D.L., Margosiak, S.A., Pinko, C., Matthews, D.A. and Kan, C.C. (1996) Structure of the human cytomegalovirus protease catalytic domain reveals a novel serine-protease fold and catalytic triad. *Cell 86*, 835-843.

Chen, H.-H., Kong, W.-P., Zhang, L., Ward, P.L. and Roos, R.P. (1995) A picornaviral protein synthesised out of frame with the polyprotein plays a role in a virus-induced immune related demyelinating disease. *Nature Medicine 1*, 927-931.

Chen, Z., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kramer, G., Shanks, M., Lomonossoff, G. and Johnson, J.E. (1989) Protein-RNA interactions in an icosahedral virus at 3.0Å resolution. *Science 245*, 154-159.

Clark, M.E., Lieberman, P.M., Berk, A.J. and Dasgupta, A. (1993) Direct cleavage of Human TATA-binding protein by poliovirus protease 3C *in vivo* and *in vitro. Mol. Cell.Biol. 13*, 1232-1237.

Clark, M.E., Hammerle, T., Wimmer, E. and Dasgupta, A. (1991) Poliovirus proteinase 3C converts an active form of transcription factor IIIC to an inactive form : a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO J. 10*, 2941-2947.

Clarke, B.E., Sangar, D.V., Burroughs, J.N., Newton, S.E., Rowlands, D.J. and Carroll, A.R. (1985) Two initiation sites for foot-and -mouth disease virus polyprotein *in vivo. J.Gen.Virol. 66*, 2615-2626.

Colonno, R.J., Condra, J.H., Mizutani, S., Callahan, P.L., Davies, M.E. and Murcko, M.A. (1988) Evidence for direct involvement of the rhinovirus canyon in receptor binding. *P.N.A.S. USA 85*, 5449-5453.

Cooney, M.K., Fox, J.P. and Kenny, G.E. (1982) Antigenic groupings of 90 rhinovirus serotypes. *Infect.Immun. 37*, 642-7.

Cordingley, M.G., Callahan, P.L., Sardana, V.L., Garsky, V.M. and Colonno, R.J. (1990) Substrate requirements of Human Rhinovirus 3C Protease for peptide cleavage *in vitro*. *J.Biol.Chem. 265*, 9062-65.

Devaney, M.A., Vakharia, V.N., Lloyd, R.E., Ehrenfeld, E. and Grubman, M.J. (1988) Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. *J Virol. 62*, 4407-4409.

Dewalt, P.G., Blair, W.S. and Semler, B.L. (1990) A genetic locus in mutant poliovirus genomes involved in overproduction of RNA polymerase and 3C proteinase. *Virology. 174*, 504-14.

Dewalt, P.G., Lawson, M.A., Colonno, R.J. and Semler, B.L. (1989) Chimeric picornavirus polyproteins demonstrate a common 3C proteinase substrate specificity. *J.Virol. 63*, 3444-3452.

Dewalt, P.G. and Semler, B.L. (1987) Site-directed mutagenesis of proteinase 3C results in a poliovirus deficient in synthesis of viral RNA Polymerase. *J.Virol. 61*, 2162-2170.

Dinakarpandian, D., Shenoy, B., PusztaiCarey, M., Malcolm, B.A. and Carey, P.R. (1997) Active site properties of the 3C proteinase from hepatitis A virus (a hybrid cysteine/serine protease) probed by Raman spectroscopy. *Biochem. 36*, 4943-4948.

Doedens, J.R. and Kirkegaard, K. (1995) Inhibition of cellular protein secretion by poliovirus proteins 2B and 3A. *EMBO J. 14,* 894-907.

Dougherty, W.G., Carrington, J.C., Cary, S.M. and Parks, T.D. (1988) Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J. 7*, 1281-1287.

Etchison, D. and Fout, S. (1985) Human rhinovirus14 infection of HeLa cells results in the proteolytic cleavage of the p220 cap-binding complex subunit and inactivates globin messenger-RNA translation *in vitro*. *J.Virol. 54*, 634-638.

Farr, B.M., Conner, E.M., Betts, R.F., Oleske, J., Minnefor, A. and Gwaltney, J.M. (1987) 2 Randomized controlled trials of zinc gluconate lozenge therapy of experimentally induced rhinovirus colds. *Antimicrob.Agents Chemother. 31*, 1183-1187.

Farr, B.M. and Gwaltney, J.M. (1987) Zinc gluconate lozenge therapy of experimentally induced rhovirus colds., in *Clinical Research* vol. 35.

Fout, G.S., Medappa, K.C., Mapoles, J.E. and Rueckert, R.R. (1984) Radiochemical determination of polyamines in poliovirus and human rhinovirus 14. *J.Biol.Chem. 259*, 3639-3643.

Fox, J.P., Cooney, M.K., Hall, C.E. and Foy, H.M. (1985) Rhinoviruses in Seattle families, 1975-1979. *Am.J.Epidemiol. 122*, 830-846.

Gessert, S.F., Kim, J.H., Nargang, F.E. and Weiss, R.L. (1994) Polyprotein precursor of 2 mitochondrial-enzymes in *Neurospora crassa* - gene structure and precursor processing. *J.Biol. Chem. 269*, 8189-8203.

Giachetti, C. and Semler, B.L. (1991) Role of a viral membrane polypeptide in strand-specific initiation of poliovirus RNA synthesis [published errata appear in J Virology 1991 Jul;65(7):3972 and 1991 Oct;65(10):5653]. *J Virol. 65*, 2647-2654.

Goldbach, R. and Wellink, J. (1988) Evolution of plus-strand RNA viruses. *InterVirology 29*, 260-7.

Greve, J.M., Davis, G., Meyer, A.M., Forte, C.P., Yost, S.C., Marlor, C.W., Kamarck, M.E. and McClelland, A. (1989) The major human rhinovirus receptor is ICAM-1. *Cell 56*, 839-47.

Greve, J.M., Forte, C.P., Marlor, C.W., Meyer, A.M., Hoover-Litty, H., Wunderlich, D. and McClelland, A. (1991) Mechanisms of receptor-mediated rhinovirus neutralization defined by two soluble forms of ICAM-1. *J.Virol.65*, 6015-6023.

Guterman, L. (1998) Colds frozen out, *New Scientist 2154* pp 17.

Hämmerle, T., Akhteruzzaman, M. and Wimmer, E. (1992) Mutational analysis of the proposed FG loop of Poliovirus Proteinase 3C identifies amino acids that are necessary for 3CD cleavage and might be determinants of a function distinct from proteolytic activity. *J.Virol. 66*, 6028-6034.

Hamparian, V.V., Colonno, R.J. and Cooney, M.K. (1987) A collaborative report:Rhinoviruses-extension of the numbering system from 89 to 100. *Virology. 159*, 191-192.

Hanecak, R., Semler, B.L., Anderson, C.W. and Wimmer, E. (1982) Proteolytic processing of poliovirus polypeptides - antibodies to polypeptide p3-7c inhibit cleavage at glutamine-glycine pairs. *P.N.A.S.USA 79*, 3973-3977.

Hanecak, R., Semler, B.L., Ariga, H., Anderson, C.W. and Wimmer, E. (1984) Expression of a cloned gene segment of poliovirus in *Escherichia coli* - evidence for autocatalytic production of the viral proteinase. *Cell 37*, 1063-1073.

Hansen, J.L., Long, A.M. and Schultz, S.C. (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure 5*, 1109-1122.

Harris, K.S., Reddigari, S.R., Nicklin, M., Hammerle, T. and Wimmer, E. (1992) Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA-polymerase. *J.Virol. 66*, 7481-7489.

Harris, K.S., Xiang, W., Alexander, L., Lane, W.S., Paul, A.V. and Wimmer, E. (1994) Interaction of Poliovirus Polypeptide $3CD^{pro}$ with the 5' and 3' termini of the Poliovirus genome. *J.Biol.Chem. 269*, 27004-27014.

Hayden, F.G. and Gwaltney, J.M.J. (1984) Intranasal interferon-2 treatment of experimental rhinoviral colds,. *J.Infect.Dis. 150*, 174-180.

Hayman, M.J., Royer-Pokora, B. and Graf, T. (1979) Defectiveness of avian erythroblastosis virus: synthesis of a 75K gag- related protein. *Virology. 92*, 31-45.

Hofer, F., Gruenberger, M., Kowalski, H., Machat, H., Huettinger, M., Kuechler, E. and Blaas, D. (1994) Members of the low-density-lipoprotein receptor family mediate cell entry of a minor-group common cold virus. *P.N.A.S. USA 91*, 1839-1842.

Hogle, J.M., Chow, M., Frick, C.E., Minor, P.D. and Filman, D.J. (1987) in *Protein structure,folding and design 2.* (Oxender, D. L., Ed.) pp 505-19, Alan R. Liss, New York.

Horsnell, C., Gama, R.E., Hughes, P.J. and Stanway, G. (1995) Molecular relationships between 21 human rhinovirus serotypes. *J.Gen.Virol.. 76*, 2549-2555.

Huguenel, E.D., Cohn, D., Dockum, D.P., Greve, J.M., Fournel, M.A., Hammond, L., Irwin, R., Mahoney, J., McClelland, A., Muchmore, E., Ohlin, A.C. and Scuderi, P.

(1997) Prevention of rhinovirus infection in chimpanzees by soluble intercellular adhesion molecule-1. *American Journal Of Respiratory And Critical Care Medicine 155*, 1206-1210.

Jacobson, M.F. and Baltimore, D. (1968) Morphogenesis of poliovirus. *J.Mol.Biol. 33*, 369-378.

Jacobson, S.J., Konings, D.A.M. and Sarnow, P. (1993) Biochemical and Genetic Evidence for a Pseudoknot Structure at the 3' Terminus of the Poliovirus RNA Genome and Its Role in Viral RNA Amplification. *J.Virol. 67*, 2961-2971.

James, M.N.G. (1994) Serine proteinases and the convergence of active-site geometries among the 4 classes of proteolytic-enzymes. *J.Cell.Biochem.*, 118-118.

Joachims, M., Harris, K.S. and Etchison, D. (1995) Poliovirus protease 3C mediates cleavage of microtubule associated protein 4. *Virology 211*, 451-461.

Jore, J., De Geus, B., Jackson, R.J., Pouwels, P.H. and Enger-Valk, B.E. (1988) Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 *in vitro*. *J.Gen.Virol. 69*, 1627-1636.

Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Mol. Biol. 35*, 993-1001.

Kean, K.M., Agut, H., Fichot, O., Wimmer, E. and Girard, M. (1988) A poliovirus mutant defective for self-cleavage at the COOH-terminus of the 3C protease exhibits secondary processing defects. *Virology. 163*, 330-40.

Kean, K.M., Howell, M.T., Grunert, S., Girard, M. and Jackson, R.J. (1993) Substitution mutations at the putative catalytic triad of the poliovirus 3C protease have differential effects on cleavage at different sites. *Virology. 194*, 360-364.

Kennedy, M.W., Britton, C., Price, N.C., Kelly, S.M. and Cooper, A. (1995) The DvA-1 Polyprotein of the Parasitic Nematode *Dictyocaulus viviparus*. *J. Biol.Chem. 270*, 19277-19281.

Kirchweger, R., Ziegler, E., Lamphear, B.J., Waters, D., Liebig, H.D., Sommergruber, W., Sobrino, F., Hohenadl, C., Blaas, D., Rhoads, R.E. and Skern, T. (1994) Foot-and-mouth-disease virus leader proteinase - purification of the Lb form and determination of its cleavage site on eIf4-gamma. *J.Virol. 68*, 5677-5684.

Kitamura, N., Adler, C. and Wimmer, E. (1980) Structure and expression of the picornavirus genome. *Annals Of The New York Academy of Sciences 354*, 183-201.

Kong, W.-P. and Roos, R.P. (1991) A;ternative translation initiation site in the DA strain of Theiler's murine encephalomyelitis virus. *J.Virol. 65*, 3395-3399.

Korant, B.D., Kauer, J.C. and Butterworth, B.E. (1974) Zinc ions inhibit replication of rhinoviruses. *Nature 248*, 588-590.

Kozak, M. (1986) Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *P.N.A.S.USA 83*, 2850-4.

Kurjan, J. and Herskowitz, I. (1982) Structure of a yeast pheromone gene (MFα): a putative α factor precursor contains four tandem copies of mature α factor. *Cell 30*, 933-943.

Laemmli, U.K., (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227* 680-685.

Lama, J. and Carrasco, L. (1992) Expression of poliovirus nonstructural proteins in *Escherichia coli* cells: Modification of membrane permeability induced by 2B and 3A. *J.Biol.Chem.* 267, 15932-15937.

Lama, J., Paul, A.V., Harris, K.S. and Wimmer, E. (1994) Properties of purified recombinant poliovirus protein 3AB as substrate for viral proteinases and as co-factor for RNA-polymerase 3D(pol). *J.Biol.Chem. 269*, 66-70.

Lawson, M.A. and Semler, B.L. (1991) Poliovirus thiol proteinase-3C can utilize a serine nucleophile within the putative catalytic triad. *P.N.A.S. USA. 88*, 9919-9923.

Lee, C.K. and Wimmer, E. (1988) Proteolytic processing of poliovirus polyprotein - elimination of $2A^{pro}$-mediated, alternative cleavage of polypeptide 3CD by *in vitro* mutagenesis. *Virology 166*, 405-414.

Leong, L. E. Walker, P. A. and Porter, A. G. (1993) Human rhinovirus-14 protease 3C ($3C^{pro}$) binds specifically to the 5'- noncoding region of the viral RNA. Evidence that $3C^{pro}$ has different domains for the RNA binding and proteolytic activities., *J.Biol.Chem 34*, 25735-25739.

Lesk, A.M. and Fordham, W.D. (1996) Conservation and variability in the structures of serine proteinases of the chymotrypsin family. *J.Mol.Biol 258*, 501-537.

Lòpez-Otîn, C., Simòn-Mateo, C., Martinez, L. and Viñuela, E. (1989) Gly-Gly-X, a novel consensus sequence for the proteolytic processing of viral and cellular proteins. *J.Biol.Chem. 264*, 9107-9110.

Love, R.A., Parge, H.E., Wickersham, J.A., Hostomsky, Z., Habuka, N., Moomaw, E.W., Adachi, T. and Hostomska, Z. (1996) The crystal-structure of hepatitis-C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc-binding site. *Cell 87*, 331-342.

Lumry, R. (1997) The famous serine protease "catalytic triad" is a pair. *Biophys.J. 72*, MP318-MP318.

Macadam, A.J., Ferguson, G., Fleming, T., Stone, D.M., Almond, J.W. and Minor, P.D. (1994) Role for poliovirus protease 2A in cap independent translation. *EMBO J. 13*, 924-927.

Markland, W., Petrillo, R.A., Fitzgibbon, M., Fox, T., McCarrick, R., McQuaid, T., Fulghum, J.R., Chen, W.Y., Fleming, M.A., Thomson, J.A. and Chambers, S.P. (1997) Purification and characterization of the NS3 serine protease domain of hepatitis C virus expressed in *Saccharomyces cerevisiae. J.Gen.Virol. 78*, 39-43.

Marlin, S.D., Staunton, D.E., Springer, T.A., Stratowa, C., Sommergruber, W. and Merluzzi, V. (1990) A soluble form of intercellular adhesion molecule-1 inhibits rhinovirus infection. *Nature 344*, 70-72.

Martin, S., Casasnovas, J.M., Staunton, D.E. and Springer, T.A. (1993) Efficient neutralization and disruption of rhinovirus by chimeric ICAM-1/immunoglobulin molecules. *J.Virol. 67*, 3561-3568.

Matsuda, A. and Komatsu, K.-I. (1985) Molecular-cloning and structure of the gene for 7-Beta-(4-carboxybutanamide) cephalosporanic acid acylase from a *Pseudomonas* strain. *J.Bacteriol. 163*, 1222-28.

Matthews, D.A., Smith, W.W., Ferre, R.A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J.E., Janson, C.A., McElroy, H.E., Gribskov, C.L. and Worland, S. (1994) Structure of Human Rhinovirus 3C Protease Reveals a Trypsin-like Polypeptide Fold, RNA-Binding Site, and Means for Cleaving Precursor Polyprotein. *Cell 77*, 761-771.

Mayer, H., Collins, J. and Wagner, F. (1979) in *Plasmids of Medical, Environmental and Commercial Importance* (Timmis, K. N., & Pühler, A., Ed.) pp 459-69, Elsevier/North Holland Biochemical Press, Amsterdam.

McClelland, A., DeBear, J., Yost, S.C., Meyer, A.M., Marlor, C.W. and Greve, J.M. (1991) Identification of monoclonal antibody epitopes and critical residues for rhinovirus binding in domain 1 of intercellular adhesion molecule 1. *P.N.A.S. USA 88*, 7993-7997.

McKinlay, M.A., Pevear, D.C. and Rossmann, M.G. (1992) Treatment of the picornavirus common cold by inhibitors of viral uncoating and attachment. *Annu.Rev.Microbiol. 46*, 635-654.

McKnight, K.L. and Lemon, S.M. (1996) Capsid coding sequence is required for efficient replication of human rhinovirus-14 RNA. *J.Virol. 70*, 1941-1952.

McKnight, K.L. and Lemon, S.M. (1998) The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. *RNA 4*, 1569-1584.

McLean, C., Matthews, T.J. and Rueckert, R.R. (1976) Evidence of ambiguous processing and selective degradation in the noncapsid proteins of Rhinovirus 1A. *J.Virol.. 194*, 355-359.

Menard, R., Chatel, H., Dupras, R., Plouffe, C. and Laliberte, J.F. (1995) Purification of turnip mosaic potyvirus viral protein genome-linked proteinase expressed in *Escherichia coli* and development of a quantitative assay for proteolytic activity. *Eur. J. Biochem. 229*, 107-112.

Miller, R.H. and Purcell, R.H. (1990) Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *P.N.A.S.USA 87*, 2057-2061.

Minor, T.E., Dick, E.C., DeMeo, A.N., Ouellette, J.J., Cohen, M. and Reed, C.E. (1974) Viruses as Precipitants of Asthma Attacks in Children. *JAMA 227*, 292-298.

Monto, A.S. and Cavallaro, J.J. (1972) The Tecumseh study of respiratory illness. IV. Prevalence of rhinovirus serotypes, 1966-1969. *Am.J.Epidemiol. 96*, 352-360.

Mosser, A.G. and Rueckert, R.R. (1993) WIN-51711-dependant mutants of poliovirus type 3: evidence that virions decay after release from cells unless drug is present. *J.Virol. 67*, 1246-1254.

Nakanishi, S., Inoue, A., Taii, S. and Numa, S. (1977) Cell-free translation product containing corticotrophin and β-endorphin encoded by messenger RNA from anterior lobe and intermediate lobe of bovine pituitary. *FEBS Letters 84*, 105-109.

Newman, J.F., Rowlands, D.J. and Brown, F. (1973) A physico-chemical sub-grouping of the mammalian picornaviruses. *J.Gen.Virol. 18*, 171-180.

Novick, S.G., Godfrey, J.C., Godfrey, N.J. and Wilder, H.R. (1996) How does zinc modify the common cold - clinical observations and implications regarding mechanisms of action. *Medical Hypotheses 46*, 295-302.

Oliveira, M.A., Zhao, R., Lee, W.M., Kremer, M.J., Minor, I., Rueckert, R.R., Diana, G.D., Pevear, D.C., Dutko, F.J., Mckinlay, M.A. and Rossmann, M.G. (1993) The structure of human rhinovirus-16. *Structure 1*, 51-68.

Olson, N.H., Kolatkar, P.R., Oliveira, M.A., Cheng, R.H., Greve, J.M., McClelland, A., Baker, T.S. and Rossmann, M.G. (1993) Structure of a human rhinovirus complexed with its receptor molecule. *P.N.A.S.USA 90*, 507-511.

Olson, N.H., Smith, T.J., Kolatkar, P.R., Oliveira, M.A., Rueckert, R.R., Greve, J.M., Rossmann, M.G. and Baker, T.S. (1992) Cryoelectron microscopy of complex of human rhinovirus with a monoclonal Fab and the viral cellular receptor. *Proc.Elect.Microsc.Soc.Am. 50*, 524-525.

Olson, S.T., Bock, P.E., Bjork, I., Kvassman, J., Shore, J.D., Lawrence, D. and Ginsburg, D. (1994) Role of the catalytic serine in the interactions of serine proteinases with inhibitors of the kunitz and serpin families. *J.Cell.Biochem.*, 159-159.

Orr, D.C., Long, A.C., Kay, J., Dunn, B.M. and Cameron, J.M. (1989) Hydrolysis of a series of synthetic peptide substrates by the Human Rhinovirus 14 3C proteinase, cloned and expressed in *Escherichia coli*. *J.Gen.Virol. 70*, 2931-2942.

Pallansch, M.A., Kew, O.M., Semler, B.L., Omilianowski, D.R., Anderson, C.W., Wimmer, E. and Rueckert, R.R. (1984) Protein processing map of poliovirus. *J.Virol.49*, 873-880.

Palmenberg, A.C. (1990) Proteolytic processing of picornaviral polyprotein. *Annu.Rev.Microbiol. 44*, 603-623.

Palmenberg, A.C. and Rueckert, R.R. (1982) Evidence for intramolecular self-cleavage of picornaviral replicase precursors. *J.Virol. 41*, 244-249.

Parks, G.D., Baker, J.C. and Palmenberg, A.C. (1989) Proteolytic cleavage of encephalomyocarditis virus capsid region substrates by precursors to the 3C enzyme. *J.Virol. 63*, 1054-1058.

Parks, G.D., Duke, G.M. and Palmenberg, A.C. (1986) Encephalomyocarditis virus-3C protease - efficient cell-free expression from clones which link viral 5' noncoding sequences to the P3-region. *J.Virol. 60*, 376-384.

Parry, N.R., Ouldridge, E.J., Barnett, P.V., Clarke, B.E., Francis, M.J., Fox, J.D., Rowlands, D.J. and Brown, F. (1989) Serological prospects for peptide vaccines against foot-and-mouth disease virus. *J.Gen.Virol. 70*, 2919-2930.

Paul, A.V.,Boom, D.V., Filipov,D. and Wimmer, E. (1998) Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature 393*, 280-284.
Pearl, L.H. and Taylor, W.R. (1987) A structural model for retroviral proteases. *Nature 329*, 351-354.

Pevear, D.C., Fancher, F.J., Feloc, P.J., Rossmann, M.G., Miller, M.S., Diana, G., Treasurywala, A.M., McKinlay, M.A. and Baker, T.S. (1989) Conformational change in the floor of the human rhinovirus canyon blocks adsorption to HeLa cell receptors. *J.Virol. 63*, 2002-2007.

Philpotts, R.J., Wallace, J., Tyrrell, D.A. and Tagart, V.B. (1983) Therapeutic activity of enviroxime against rhinovirus infection in volunteers. *Antimicrob Agents Chemother 23*, 671-675.

Piccone, M.E., Zellner, M., Kumosinski, T.F., Mason, P.W. and Grubman, M.J. (1995) Identification of active site residues of the L-proteinase of foot-and-mouth disease virus. *J.Virol. 69*, 4950-4956.

Pilipenko, E.V., Maslova, S.V., Sinyakov, A.N and Agol, V.I. (1990) Towards identification of *cis*- acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. *Nuc.Acids.Res. 20*, 1739-1745.

Porter, A.G. (1993) Picornavirus nonstructural proteins - emerging roles in virus-replication and inhibition of host-cell functions. *J.Virol. 67*, 6917-6921.

Porter, D.C., Ansardi, D.C., Lentz, M.R. and Morrow, C.D. (1993) Expression of poliovirus P3 proteins using a recombinant vaccinia virus results in proteolytically active 3CD precursor protein without firther processing to $3C^{pro}$ and $3D^{pol}$. *Virus Res 29*, 241-54.

Press release-Agouron Pharmaceuticals, I. (1998) in *PRNewswire*, La Jolla, California.

Qiu, X.Y., Culp, J.S., DiLella, A.G., Hellmig, B., Hoog, S.S., Janson, C.A., Smith, W.W. and AbdelMeguid, S.S. (1997) Human cytomegalovirus protease: A novel serine protease fold, a unique catalytic triad and an important drug target. *Protein Engineering 10*, 52-52.

Register, R.B., Uncapher, C.R., Naylor, A.M., Lineberger, D.W. and Colonno, R.J. (1991) Human-murine chimeras of ICAM-1 identify amino acid residues critical for rhinovirus and antibody. *J.Virol. 65*, 6589-6596.

Roberts, P.J. and Belsham, G.J. (1995) Identification of critical amino acids within the foot-and-mouth disease virus leader protein, a cysteine protease. *Virology. 213*, 140-146.

Roos, R.P., Kong, W. and Semler, B.L. (1989) Polyprotein processing of Theiler's Murine Encephalomyelitis virus. *J.Virol. 63*, 5344-5353.

Rossmann, M.G., Arnold, E., Erickson, J.W., Frankenberger, E.A., Griffith, J.P., Hecht, H.-J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G., Rueckert, R.R., Sherry, B. and Vriend, G. (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature 317*, 145-153.

Rossmann, M.G. and Palmenberg, A.C. (1988) Conservation of the putative receptor attachment site in picornaviruses. *Virology. 164*, 373-382.

Rothberg, P.G., Harris, T.J.R., Nomoto, A. and Wimmer, E. (1978) O4-(5'uridylyl) tyrosine is the bond between the genome linked protein and the RNA of poliovirus. *P.N.A.S.USA 75*, 4868-4872.

Rueckert, R.R. (1996) in *Fields Virology* (Fields, B. N., Knipe, D. M., Howley, P. M., & et al., Eds.) pp 609-654, Lippincott-Raven, Philadelphia.

Rueckert, R.R., Matthews, T.J., Kew, O.M., Pallansch, M.A., McLean, C. and Omilianowski, D. (1979) in *The Molecular Biology of Picornaviruses* (Perez-Bercoff, R., Ed.) pp 113-125, Plenum, New York.

Ryan, M.D. and Drew, J. (1994) Foot and Mouth Disease Virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO 13*, 928-933.

Ryan, M.D. and Flint, M. (1997) Virus-encoded proteinases of the picornavirus super-group. *J.Gen.Virology. 78*, 699-723.

Ryan, M.D., Monaghan, S. and Flint, M. (1998) Virus-encoded proteinases of the *Flaviviridae*. *J.Gen.Virol. 79*, 947-959.

Samo, T.C., Greenberg, S.B. and Couch, R.B. (1983) Efficacy and tolerance of intranasally applied recombinant leukocyte A interferon in normal volunteers. *J.Infect.Dis. 148*, 535-42.

Sangar, D.V., Clark, R.P., Rowlands, D.J. and Clarke, B.E. (1988) Modification of the leader protein (Lb) of foot-and-mouth disease virus. *J.Gen.Virol. 69*, 2327-2333.

Sangar, D.V., Newton, S.E., Rowlands, D.J. and Clarke, B.E. (1987) All FMDV serotypes initiate protein synthesis at two separate AUGs. *Nuc.Acids Res. 15*, 3305-3315.

Sato, Y., Oguchi, M., Menjo, N., Imai, K., Saito, H., Ikeda, M., Isobe, M. and Yamashita, O. (1993) Precursor polyprotein for multiple neuropeptides secreted from the subesophageal ganglion of the silkworm *Bombyx mori* - characterization of the cDNA-encoding the diapause hormone precursor and identification of additional peptides. *P.N.A.S. USA 90*, 3251-3255.

Schechter, I. and Berger, A. (1967) On the size of the active site in proteases.I.papain. *Biochem.Biophys.Res.Comm. 27*, 157-162.

Schlegel, A., Giddings, J.R., Ladinsky, M.S. and Kirkegaard, K. (1996) Cellul;ar origin and ultrastructure of membranes induced during poliovirus infection. *J.Virol. 70,* 6576-6588.

Schneemann, A., Gallagher, T.M. and Rueckert, R.R. (1994) Reconstitution of flock house provirions - a model system for studying structure and assembly. *J.Virol. 68*, 4547-4556.

Schober, D., Kronenberger, P., Prchla, E., Blaas, D. and Fuchs, R. (1998) Major and Minor Receptor Group Human Rhinoviruses Penetrate from Endosomes by Different Mechanisms. *J.Virol. 72*, 1354-1364.

Seemüller, E., Lupas, A. and Baumeister, W. (1996) Autocatalytic processing of the 20S proteasome. *Nature 382*, 468-70.

Seemüller, E., Lupas, A., Stock, D., Löwe, J., Huber, R. and Baumeister, W. (1995a) Proteasome from *Thermoplasma acidophilum* : A Threonine Protease. *Science 268*, 579-582.

Seemüller, E., Lupas, A., Zühl, F., Zwickl, P. and Baumeister, W. (1995b) The proteasome from *Thermoplasma acidophilum* is neither a cysteine nor a serine protease., *FEBS Letters 359*, 173.

Semler, B.L., Johnson, V.H., Dewalt, P.G. and Ypma-Wong, M.F. (1987) Site-specific mutagenesis of cDNA clones expressing a Poliovirus proteinase. *J.Cell.Biochem. 33*, 39-51.

Sherry, B. and Rueckert, R.J. (1985) Evidence for at least two dominant neutralization antigens on human rhinovirus14., *Journal of Virology 53*, 137-143.

Sherry, B., Mosser, A.G., Colonno, R.J. and Rueckert, R.R. (1986) Use of monoclonal antibodies to identify four neutralization immunogens on a common cold picornavirus, human rhinovirus 14., *J.Virol. 57*, 246-257.

Simòn-Mateo, C., Andrès, G., Almazàn, F. and Viñuela, E. (1997) Proteolytic Processing in African Swine Fever Virus: Evidence for a New Structural Polyprotein, pp62. *J.Virol. 71*, 5799-5804.

Simòn-Mateo, C., Andrès, G. and Viñuela, E. (1993) Polyprotein processing in African Swine Fever Virus: a novel gene expression strategy for a DNA virus. *EMBO J. 12*, 2977-2987.

Sommergruber, W., Ahorn, H., Klump, H., Seipelt, J., Zoephel, A., Fessl, F., Krystek, E., Blaas, D., Kuechler, E., Liebig, H.D. and Skern, T. (1994a) 2A proteinases of coxsackievirus and rhinovirus cleave peptides derived from elf-4-gamma via a common recognition motif. *Virology. 198*, 741-745.

Sommergruber, W., Casari, G., Fessl, F., Seipelt, J. and Skern, T. (1994b) The 2A proteinase of human rhinovirus is a zinc-containing enzyme. *Virology. 204*, 815-818.

Sommergruber, W., Zorn, M., Blaas, D., Fessl, F., Volkmann, P., Maurer-Fogy, I., Pallai, P., Merluzzi, V., Matteo, M., Skern, T. and Kuechler, E. (1989) Polypeptide 2A of Human Rhinovirus type 2: Identification as a protease and characterization by mutational analysis. *Virology 169*, 68-77.

Sperber, S.J. and Hayden, F.G. (1988) Chemotherapy of rhinovirus colds. *Antimicrob.Agents Chemother. 32*, 409-419.

Sperber, S.J., Levine, P.A., Sorrentino, J.V., Riker, D.K. and Hayden, F.G. (1989) Ineffectiveness of recombinant interferon-serine nasal drops for prophylaxis of natural colds. *J.Infect.Dis. 160*, 700-705.

Stanway, G. (1990) Structure, function and evolution of picornaviruses. *J.Gen.Virol. 71*, 2483-2501.

Stanway, G., Hughes, P.J., Mountford, R.C., Minor, P.J. and Almond, J.W. (1984) The complete nucleotide sequence of a common cold virus : human rhinovirus 14., *Nucleic Acids Res 12*, 7859-7875.

Stanway, G., Kalkkinen, N., Roivainen, M., Ghazi, F., Khan, M., Smyth, M., Meurman, O. and Hyypiä, T. (1994) Molecular and biological characteristics of echovirus 22, a representative of a new picornavirus group. *J.Virol. 68*, 8232-8238.

Staunton, D.E., Merluzzi, V.J., Rothlein, R., Barton, R., Marlin, S.D. and Springer, T.A. (1989) A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell 56*, 849-53.

Steitz , T.A. and Schulman, R.G. (1982) Crystallographic and NMR studies of the serine proteases., *Annu Rev Biophys Bioeng 11*, 419-44.

Strebel, K. and Beck, E. (1986) A second protease of foot-and-mouth disease virus. *J.Virol. 58*, 893-899.

Tautz, N., Elbers, K., Stoll, D., Meyers, G. and Thiel, H.J. (1997) Serine protease of pestiviruses: Determination of cleavage sites. *J.Virology. 71*, 5415-5422.

Thöny-Meyer, L., James, P. and Hennecke, H. (1991) From one gene to two proteins: The biogenesis of cytochromes b and c1 in *Bradyrhizobium japonicum*. *P.N.A.S.USA 88*, 5001-5005.

Todd, S. and Semler, B.L. (1996) Structure-infectivity analysis of the human rhinovirus genomic RNA 3' non-coding region. *Nucleic Acids Res 24*, 2133-2142.

Todd, S., Towner, J.S. and Semler, B.L. (1997) Translation and replication properties of the human rhinovirus genome *in vivo* and *in vitro. Virology. 229*, 90-97.

Toh, H., Kikuno, R., Hayashida, H., Miyata, T., Kugimaya, W., Inouye, S., Yuki, S. and Saigo, K. (1985) Close structural resemblance between putative polymerase of a *Drosophila* transposable genetic element 17.6 and *pol* gene product of Moloney murine leukaemia virus. *EMBO J. 4.*

Tomassini, J.E., Graham, D., DeWitt, C.M., Lineberger, D.W., Rodkey, J.A. and Colonno, R.J. (1989) cDNA cloning reveals that the major group rhinovirus receptor on HeLa cells is intercellular adhesion molecule 1. *P.N.A.S. U S A 86*, 4907-11.

Toyoda, H., Nicklin, M.J.H., Murray, M.G., Anderson, C.W., Dunn, J.J., Studier, F.W. and Wimmer, E. (1986) A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell 45*, 761-770.

Turner, R.B., Dutko, F.J., Goldstein, N.H., Lockwood, G. and Hayden, F.G. (1993) Efficacy of oral win-54954 for prophylaxis of experimental rhinovirus infection. *Antimicrob Agents Chemother. 37*, 297-300.

Tyrrell, D.A.J. and Channock, R.M. (1963) A description of rhinoviruses. *Science 141*, 152-153.

Uncapher, C.R., DeWitt, C.M. and Colonno, R.J. (1991) The major and minor group receptor families contain all but one human rhinovirus serotype. *Virology. 180*, 814-817.

van Kuppefeld, F.J.M., Hoenderop, J.G.JJ., Smeets, R.L.L., Willems, P.G.H.M., Dikman, H.B.P.M., Galama J.M.D. and Melchers, W.J.G. (1997a) Coxsackievirus protein 2B modifies endoplasmic retculum membrane and plasma membrane permeability and facilitates virus release. *EMBO J. 16,* 3519-3532.

van Kuppefeld, F.J.M., Melchers, W.J.G., Kirkegaard, K. and Doedens, J. (1997b) Structure-function analysis of coxsackie B3 virus protein 2B. *Virology 227,* 111-118.

Voss, T., Meyer, R. and Sommergruber, W. (1995) Spectroscopic characterization of rhinoviral protease 2A: Zn is essential for the structural integrity. *Protein Science 4*, 2526-2531.

Warshel, A., Naray-Szabo, G., Sussman, F. and Hwang, J.K. (1989) How do serine proteases really work? *Biochem.stry 28*, 3629-3637.

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

Wimmer, E. (1982) Genome-linked proteins of viruses. *Cell 28*, 199-201.

Wirblich, C., Thiel, H.-J. and Meyers, G. (1996) Genetic Map of the Calicicvirus Rabbit Haemorrhagic Disease Virus as Deduced from *In Vitro* Translation Studies. *J.Virol. 70*, 7974-7983.

Wlodawer, A. (1995) Proteasome: a complex protease with a new fold and a distinct mechanism. *Structure 3*, 417-20.

Woods, M.G., Diana, G.D., Rogge, M.C., Otto, M.J., Dutko, F.J. and McKinlay, M.A. (1989) *In vitro* and *in vivo* activities of WIN 54954, a new broad-spectrum antipicornavirus drug. *Antimicrob Agents Chemother 33*, 2069-2074.

Xiang, W.K., Harris, K.S., Alexander, L. and Wimmer, E. (1995) Interaction between the 5'-terminal cloverleaf and 3AB/3CD(pro) of poliovirus is essential for RNA replication. *J.Virol. 69*, 3658-3667.

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.

Ypma-Wong, M.F., Dewalt, P.G., Johnson, V.H., Lamb, J.G. and Semler, B.L. (1988) Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology. 166*, 265-70.

Ypma-Wong, M.F. and Semler, B.L. (1987) *In vitro* molecular genetics as a tool for determining the differential cleavage specificities of the poliovirus 3C proteinase. *Nuc.Acid.Res.*, 2069-2088.

Yu, S. and Lloyd, R.E. (1991) Identification of Essential Amino Acid Residues in the Functional Activity of Poliovirus 2A Protease.*Virology 182*, 615-625.

Zeichhardt, H., Wetz, K., Willingmann, P. and Habermehl, K.O. (1985) Entry of poliovirus type 1 and mouse Elberfeld (ME) virus into HEp-2 cells: receptor mediated endocytosis and endosomal or lysosomal uncoating. *J.Gen.Virol. 66*, 483-492.

Zhao, R., Pevear, D.C., Kremer, M.J., Giranda, V.L., Kofron, J.A., Kuhn, R.J. and Rossmann, M.G. (1996) Human rhinovirus-3 at 3.0-angstrom resolution. *Structure 4*, 1205-1220.