



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study,
without prior permission or charge

This work cannot be reproduced or quoted extensively from without first
obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any
format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author,
title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

**Functional analysis of viral RNA and protein-RNA
interactions involved in the replication of
Poliovirus type 3**

**A Thesis presented for the degree of Doctor of Philosophy in the
Faculty of Biomedical and Life Sciences at the University of
Glasgow**

Inga Ruth Dry

Division of Virology

Institute of Biomedical and Life Sciences

ProQuest Number: 10800557

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10800557

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

GLASGOW
UNIVERSITY
LIBRARY:

Acknowledgements

I would like to thank my supervisor Dr. David J. Evans for all constructive discussions (and criticism!) throughout the course of my PhD and the writing of this thesis.

A special thank you must go to Prof. P. Simmonds of Edinburgh University who carried out the SSSV analysis documented in this thesis, Dr. I. Goodfellow, for his time and helpful advice and to David Kerrigan for practical advice regarding tissue culture and infectivity assays! Also to Dr. H. Harvala for keeping me up to date with recent developments with regards to my work on RNA secondary structures and to all members of the lab past and present, in particular Dr. V. Cowton and Claire Blanchard for their ability to cheer me up on my down days and for their support during the writing of the thesis. I won't forget the cakes and coffee!

I would also like to thank the girls in the washroom, but particularly Ann Brownlee, for their sterling work with the autoclaves, which made my life easier during the course of my PhD.

Lastly, but not least, I would like to acknowledge my family and friends for their support and encouragement throughout the course of my PhD.

Abstract

RNA viruses are responsible for the vast majority of diseases affecting vertebrates and invertebrates. One of the largest single groupings of RNA viruses have been classified in the family *Picornaviridae*. The prototype member of the *Picornaviridae* is poliovirus, though rhinovirus (causative agent of the common cold) and foot and mouth disease virus have in recent years been more economically significant to developed countries. Since the early 1980's poliovirus has provided a useful model for understanding the replication of the *Picornaviridae* and less amenable RNA viruses, like the human caliciviruses and hepatitis C virus (HCV).

Within this study, an attempt was made to further the current understanding of picornavirus replication by investigating viral RNA and protein-RNA interactions involved in the replication of poliovirus. To achieve this, three approaches were used: (i) interactions between the viral replication proteins and the RNA were studied using the yeast three-hybrid system, (ii) regions of suppression of variation at synonymous sites identified, within aligned Human Enterovirus genomes using a bioinformatic package were studied using a subgenomic replicon system and (iii) an attempt was made to further the current understanding of the interaction between the viral polymerase and the 3'UTR by characterising the N18Y mutation previously documented by Meredith *et al* (Meredith *et al.*, 1999).

To investigate the regions of the genome and anti-genome with which the viral proteins, interacted, using the yeast three-hybrid system, a cDNA library was created that enabled expression of hybrid RNAs that contained fusions of 100-200bp fragments of the poliovirus genome with the MS2 coat protein-binding sites. In addition to the hybrid RNA, the yeast three-hybrid system involves the expression of two fusion proteins (i) a fusion of the GAL4 binding domain and the MS2 coat protein and (ii) the viral protein to be investigated (3CD^{pro}, 3D^{pol}, 3AB or 2C) fused with the GAL4 activation domain. Transformants were selected but analysis of these failed to provide evidence for any specific interaction between the viral proteins and the poliovirus RNA (although expression of the non-structural proteins could be detected by western blotting).

The second approach used to identify RNA sequences or structures of functional importance was more fruitful. A bioinformatics package that recognises phylogenetically conserved sequences identified a region of the poliovirus genome (nt 6768-7148) that showed suppression of variation at synonymous sites. Predictions about the structure of the RNA at this specific location were used to plan mutagenesis studies. Mutations that disrupted the structure of this region, which lies within the region encoding the viral polymerase, were introduced using asymmetric and overlapping PCR. The effect of introducing these mutations on viral replication was investigated using a luciferase-based subgenomic replicon system. The availability of a reverse genetics system for poliovirus also enabled the effect of the mutations on the phenotype of the virus to be studied within the context of a complete replication cycle. The reduction in the ability of one of the mutants investigated to replicate led to the identification of a possible species-specific replication element. Although it is possible that this sequence functions as an antagonist of the immune system or as a membrane-targeting signal sequence it was not possible to verify either of these possibilities within the confines of this study. Identification of the function of this sequence provides an exciting avenue of future work and may provide valuable insight into differences between picornaviruses at the species level.

Lastly, 3D^{wt} and 3D^{N18Y} polymerases were expressed in *Escherichia coli*, purified and their relative activity was investigated and compared using a variety of biochemical assays (uridylation, polymerisation and terminal transferase). Although preliminary analysis of 3D^{N18Y} indicated that it led to an increase in the intrinsic polymerisation activity of the polymerase, subsequent analysis has suggested this may relate to differences between the preparations of 3D^{wt} and 3D^{N18Y}.

Amino Acid Table

Amino acid table showing single letter codes for proteins and codons.

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

	U	C	A	G	
U	UUU F UUC UUA L UUG	UCU UCC S UCA UCG	UAU UAC Y UAA Stop(ochre) UAG Stop(Amber)	UGU UGC C UGA Stop(opal) UGG W	U C A G
C	CUU CUC L CUA CUG	CCU CCC P CCA CCG	CAU H CAC CAA CAG Q	CGU CGC R CGA CGG	U C A G
A	AUU AUC I AUA AUG M	ACA ACC T ACA ACG	AAU AAC N AAA AAG K	AGU AGC S AGA AGG R	U C A G
G	GUU GUC V GUA GUG	GCU GCC A GCA GCG	GAU GAC D GAA GAG E	GGU GGC G GGA GGG	U C A G

Abbreviations

AD	Activation domain
A2RE	hnRNP A2 response element
ATP	Adenosine tri-phosphate
3-AT	3-aminotriazole
BD	Binding domain
BEV	Bovine enterovirus
CAT	Chloramphenicol acetyl transferase
CNS	Central nervous system
CRE	<i>cis</i> -acting replication element
cDNA	DNA complementary to RNA
DI	Defective interfering
DTT	Dithiothreitol
dNTP	Deoxyribonucleoside 5' triphosphate
DNA	Deoxyribosenucleic acid
<i>E.coli</i>	<i>Escherichia coli</i>
eIF	Eukaryotic initiation factor
EF	Elongation factor (Prokaryotic)
EMCV	Encephalomyocarditis virus
ER	Endoplasmic reticulum
ERAV	Equine rhinitis virus
FISH	Fluorescent <i>in situ</i> hybridisation
FLC	Full length clone
FMDV	Foot and Mouth disease virus
GTP	Guanidine tri-phosphate
GuHCl	Guanidine hydrochloride
HAV	Hepatitis A virus
HCV	Hepatitis C virus
HEV	Human enterovirus
HIV	Human Immunodeficiency virus
HSV	Herpes simplex virus
HRPO	Horseradish peroxidase
HRVA	Major rhinovirus group
HRVB	Minor rhinovirus group
HRV14	Human rhinovirus 14
IPTG	Isopropylthiogalactosidase
IRES	Internal ribosome entry site
IRE	Iron response element
IRP	Iron response protein
kDa	Kilodaltons
kb	Kilobases
Leu RS	leucyl-tRNA synthetase
MHC	Major histocompatibility complex
M.O.I.	Multiplicity of infection
mRNA	messenger RNA
mM	millimolar
nPTB	neuronal form of PTB
NMR	Nuclear magnetic resonance
NPC	Nuclear pore complex

nt	nucleotides
NTR	non-translated region
ORF	Open reading frame
ONPG	<i>O</i> -nitrophenolgalactosidase
PTB	Polypyridimine tract binding protein
PV	Poliovirus
RNA	Ribonucleic acid
RdRp	RNA dependent RNA polymerase
PBS	Phosphate buffered saline
PABP	Poly (r) A binding protein
PCBP	Poly (r) C binding protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEV	Porcine enterovirus
P-num	Pairing number
PTB	Polypyrimidine tract- binding protein
PVR	Poliovirus receptor
RD	Rhadowyosarcoma
RF	Replication Intermediate (picornavirus)
SDS	sodium dodecyl sulphate
SEV	Simian enterovirus
<i>S.cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SSSV	Suppression of synonomous site variation
TAE	Tris-acetate EDTA
TBSV	Tomato bushy stunt virus
TBP	TATA-binding protein
TEMED	N, N, N', N'-Tetramethylethylenediamine
TNF	Tumour necrosis factor
TMEV	Theiler's murine encephalomyocarditis virus
tRNA	Transfer RNA
Tris	Tris (hydroxylmethyl) methylamine
UAS	Upstream activation sequence
UTR	Untranslated region
UV	Ultra violet
w/v	Weight by volume
v/v	Volume by volume

Table of Contents

Statement.....	i
Acknowledgements.....	ii
Abstract.....	iii
Abbreviations.....	vi
Table of Contents.....	viii
List of Figures and Tables.....	xi
1 Introduction.....	1
1.1 Positive-sense strand RNA viruses.....	1
1.1.1 Enveloped viruses with positive-sense RNA genomes.....	1
1.1.2 Non-enveloped viruses with positive-sense RNA genomes.....	2
1.1.3 Replication of positive-sense single-stranded viruses.....	2
1.2 Picornaviridae.....	3
1.2.1 Human enterovirus classification.....	4
1.3 Picornaviridae particle structure.....	8
1.4 Genome Organisation.....	9
1.5 Overview of the replication cycle.....	10
1.6 Viral attachment.....	11
1.6.1 Viral entry and uncoating.....	12
1.7 Genome translation.....	13
1.7.1 IRES.....	13
1.8 Processing determinants and Cleavage products.....	16
1.8.1 The structural proteins.....	17
1.8.2 Nonstructural proteins.....	18
1.8.3 2A.....	18
1.8.4 2B, 2C and 2BC.....	21
1.8.5 3A, 3B and 3AB.....	24
1.8.6 3C ^{pro} , 3CD ^{pro} and 3D ^{pol}	26
1.9 Genomic replication.....	28
1.9.1 Formation of the poliovirus replication complex.....	29
1.9.2 <i>Cis</i> -acting replication elements.....	30
1.10 Models of poliovirus RNA replication.....	35
1.10.1 Model of minus strand synthesis.....	35
1.10.2 Model of plus-strand synthesis.....	36
1.11 Project aims.....	37
2 Identification of protein-RNA interactions using the Yeast three-Hybrid system.....	39
2.1 Introduction.....	39
2.2 Creating a cDNA library.....	42
2.2.1 pIII ^A /MS2-1.....	42
2.2.2 Generation of cDNA fragments.....	43
2.2.3 Construction of the cDNA library.....	46
2.2.4 Identification of known RNA <i>cis</i> -acting replication elements in cDNA library.....	48
2.3 cDNA expressing viral non-structural proteins.....	49
2.3.1 Expression of the viral proteins in <i>S. cerevisiae</i> R40.....	50
2.4 Identification of a characterised viral protein-RNA interaction.....	52
2.5 Background activation of the β-galactosidase promoter.....	53
2.6 cDNA library screen.....	55
2.7 Discussion of results.....	57

2.8	Future work	64
3	Analysis of the functional significance of a region of suppression of RNA sequence variation in poliovirus	66
3.1	Introduction	66
3.2	Analysis of SSSV amongst the Enteroviruses	67
3.3	MFOLD analysis of poliovirus type 3 Leon nt 6768-7148	68
3.4	Disruption of the stem loop structures I and II	69
3.4.1	Disruption of structure I.....	69
3.4.2	Disruption of structure II.....	71
3.5	Site directed mutagenesis of the region nt 6768-7148	73
3.5.1	PCR mutagenesis using an overlap strategy.....	73
3.5.2	Mutagenesis using an asymmetric PCR strategy.....	74
3.6	Construction of subgenomic replicons bearing mutations in 3D^{pol} structure I and II	75
3.7	Introduction of mutations in 3D^{pol} structure I and II into an infectious cDNA	76
3.8	Functional studies on the effect of mutations in 3D^{pol} structure I and II	78
3.8.1	Analysis of the influence of structure I and II mutations on replication.....	78
3.8.2	Influence on translation and processing of mutations in 3D ^{pol} structure I and II.....	82
3.9	Infectivity assay	85
3.10	Discussion of results	87
3.10.1	Identification of a species-specific sequence.....	93
3.11	Future work	101
4	Preliminary characterisation of 3D^{N18Y}	102
4.1	Introduction	102
4.1.1	Crystal structure of 3D ^{wt}	102
4.2	HCV polymerase: structure-function model for 3D^{wt}	107
4.3	N18Y mutation	110
4.4	Construction of DNA vectors expressing 3D^{wt} and 3D^{N18Y}	112
4.5	Expression of 3D^{wt} and 3D^{N18Y}	113
4.6	Purification of the 3D^{wt} and 3D^{N18Y}	114
4.7	Preliminary functional characteristics of 3D^{wt} and 3D^{N18Y}	115
4.7.1	Uridylylation.....	115
4.7.2	Polymerisation.....	117
4.7.3	Terminal transferase activity.....	118
4.7.4	Oligomerisation.....	119
4.8	Discussion of results	127
4.8.1	Uridylylation.....	127
4.8.2	Oligomerisation.....	129
4.8.3	Polymerisation.....	131
4.8.4	Recycling of the polymerase.....	132
4.8.5	RNA binding.....	134
4.8.6	Conclusion.....	137
5	Discussion of results	138
5.1	A membrane-targeting hypothesis	141
5.1.1	Possible effect of disrupting the membrane-targeting of the genome: Scenario 1.....	142
5.1.2	Possible effect of disrupting the membrane-targeting of the genome: Scenario 2.....	144

5.1.3	<u>Possible effect of disrupting the membrane-targeting of the genome: Scenario 3</u>	144
5.2	<u>Motif X as an antagonist of the anti-viral response</u>	145
5.3	<u>Future work</u>	145
5.4	<u>Conclusion</u>	150
6	<u>Materials and Methods</u>	151
6.1	<u>Computer programs</u>	151
6.2	<u>Solutions and Chemical suppliers</u>	151
6.2.1	<u>Standard solutions</u>	151
6.3	<u>Antibodies</u>	153
6.4	<u>Nucleic acid manipulations</u>	154
6.4.1	<u>Nucleotide suppliers</u>	154
6.4.2	<u>Oligonucleotides</u>	154
6.4.3	<u>Polymerase chain reaction (PCR)</u>	154
6.4.4	<u>Agarose gel electrophoresis</u>	155
6.4.5	<u>Restriction digestion of DNA</u>	155
6.4.6	<u>DNA extractions</u>	155
6.4.7	<u>SAP treatment of DNA</u>	155
6.4.8	<u>Ligations</u>	157
6.4.9	<u>Strains and genotypes of <i>E. coli</i></u>	158
6.4.10	<u>Bacterial cell transformations</u>	159
6.5	<u>Antibiotic selection</u>	160
6.6	<u>Preparation of plasmid DNA</u>	160
6.7	<u>Sequencing</u>	160
6.8	<u>T7 RNA transcriptions</u>	160
6.8.1	<u>Preparation of linearised DNA template</u>	160
6.8.2	<u>Preparation of rU15 /poly (rA)₄₀₀ primer/template (x20 stock)</u>	161
6.8.3	<u>T7 RNA transcription reactions</u>	161
6.9	<u>Protein analysis</u>	161
6.9.1	<u>Preparation of dialysis tubing</u>	161
6.9.2	<u>Protein concentration determination</u>	162
6.9.3	<u>SDS-PAGE electrophoresis</u>	162
6.9.4	<u>Tris-Tricine SDS polyacrylamide gel electrophoresis</u>	162
6.9.5	<u>Urea-Acrylamide gel</u>	162
6.9.6	<u>Coomasie brilliant blue staining</u>	163
6.9.7	<u>Western blotting</u>	163
6.9.8	<u>Stripping Antibody from proteins immobilised on nitrocellulose membrane</u>	164
6.9.9	<u>Standard VPg Uridylylation assay</u>	164
6.9.10	<u>Poly (rU) polymerase assay</u>	164
6.9.11	<u>Terminal uridyl transferase assay</u>	164
6.9.12	<u>Luciferase assay</u>	165
6.9.13	<u>In Vitro translation and processing assay</u>	165
6.10	<u>Yeast Techniques</u>	166
6.10.1	<u>Strains of <i>S. cerevisiae</i> used</u>	166
6.10.2	<u>Yeast media and solutions</u>	166
6.10.3	<u>Yeast transformation</u>	167
6.10.4	<u>Extraction of protein samples from yeast</u>	168
6.10.5	<u>β-Galactosidase activity</u>	168

6.11	<u>Mammalian cell techniques</u>	170
6.11.1	<u>Cell culture</u>	170
6.11.2	<u>Cell techniques</u>	170
6.12	<u>Virological techniques</u>	170
6.12.1	<u>Passage of virus</u>	170
6.12.2	<u>Plaque assays</u>	171
7	<u>Bibliography</u>	172
Appendix 1	201
Appendix 2	204
Appendix 3	209

List of Figures and Tables

Figures		Page Number
1.1	Phylogeny of the human enteroviruses	5
1.2	Neuronal retrograde transport pathway	7
1.3	Structures of picornavirus virions	8
1.4	Picornavirus genome organisation	9
1.5	Overview of picornavirus replication cycle	10
1.6	Known picornavirus receptors	11
1.7	Model of genomic uncoating in picornaviruses	12
1.8	Structure of type I and type II IRESes	13
1.9	Primary cleavages of the picornavirus polyprotein	16
1.10	Secondary cleavages of the picornavirus polyprotein	16
1.11	Model for assembly of progeny virus	17
1.12	Model for translation of cellular mRNA	18
1.13	Proposed structure of virus protein 2B	22
1.14	Model for the release of virus progeny	22
1.15	Structural domains of virus protein 2C	23
1.16	Map of the interactions between proteins of the P2 region	24
1.17	Structure of the viral replication complex	29
1.18	Structure of the cloverleaf	30
1.19	Structure of the CRE	31
1.20	Model for the templated addition of uridylate residues to VPg	31
1.21	Structure of the 3'UTR	32
1.22	Model of negative-sense RNA synthesis	35
1.23	Model of positive-sense RNA synthesis	36
2.1	Diagram illustrating the yeast three-hybrid system	41
2.2	Vector map of cDNA pIII/MS2-1	42
2.3	Schematic of hybrid RNA expressed from pIII/MS2-1 cDNA	42
2.4	Location of restriction enzyme sites within pT7FLC/Rep3	43
2.5	PCR amplification of the cDNA library	47
2.6	PCR identification of known <i>cis</i> -acting replication elements present within the cDNA library	48
2.7	Vector map of cDNA pACTII	49
2.8	Western blot of fusion proteins expressed in <i>S.cerevisiae</i>	51
2.9	Protein-RNA interaction as detected for by qualitative β -galactosidase filter assay	52
2.10	Protein-RNA interaction as detected for by qualitative liquid β -galactosidase assay	52
2.11	cDNA library interactions tested by qualitative β -galactosidase filter assay	56
2.12	Diagram illustrating the effects of a G-C clamp on RNA structure	61
2.13	Diagram illustrating how the MS2 hairpin could block the protein binding site	61
2.14	Types of multi-protein interactions	62
2.15	Interactions and functions of 3AB with other viral proteins	62
2.16	The modified yeast three-hybrid system	63
3.1	SSSV analysis of enterovirus species	67
3.2	MFOLD analysis of nt 6768-7148	68

3.3	P-num plot of nt 6768-7148	68
3.4	MFOLD analysis of nt 6893-7100 in poliovirus serotypes 1-3	68
3.5	MFOLD analysis of affect of disrupting base stem of structure I (<i>Nhe I</i>)	70
3.6	MFOLD analysis of affect of disrupting structure I (<i>Bgl II</i>)	70
3.7	MFOLD analysis of affect of disrupting structure I (<i>Nde I</i>)	70
3.8	MFOLD analysis of effect of disrupting structure II (<i>Apo I</i>)	71
3.9	MFOLD analysis of effect of disrupting structure II (<i>Nsi I</i>)	71
3.10	Mutagenesis using overlapping PCR	73
3.11	Mutagenesis using asymmetric PCR	74
3.12	Vector map of pT7Luc+R	75
3.13	Vector map of pT7FLC	76
3.14A	Replication assay pT7Luc3D ^{<i>Nhe I</i>} +R	80
3.14B	Replication over time using pT7Luc3D ^{<i>Nhe I</i>} +R	80
3.15A	Replication assay pT7Luc3D ^{<i>Bgl II</i>} +R	80
3.15B	Replication assay pT7Luc3D ^{<i>Nde I</i>} +R	80
3.16A	Replication assay pT7Luc3D ^{<i>NsIII</i>} +R	81
3.16B	Replication assay pT7Luc3D ^{<i>Apo I</i>} +R	81
3.17	Translation of pT7Luc3D ^{<i>Nde I</i>} +R and pT7Luc3D ^{<i>NsIII</i>} +R	82
3.18	Translation of P3/Leon/37 ^{<i>Nhe I</i>} and P3/Leon/37 ^{<i>Apo I</i>}	84
3.19	Infectivity assay	86
3.20	Stability of mutations	86
3.21A	One step growth curves of P3/Leon/37 ^{<i>Nhe I</i>} and P3/Leon/37	86
3.21B	Titres of P3/Leon/37 ^{<i>Nhe I</i>} at 39 °C from a starting M.O.I. of 0.5, 1 and 5	86
3.22	Mechanism of promoter masking in Tombusviruses	92
3.23	3'UTR versus Structure I and Structure II	92
3.24	MFOLD analysis of HEV-C	94
3.25	Sequence alignment of region of motif X	94
3.26	SSSV analysis of enterovirus B and C species	94
3.27	Witwer analysis of Picornavirus genomes	94
3.28	Mechanism of replicative recombination	97
3.29	Putative model for hnRNP A2/A2RE-mediated in HIV-1	100
4.1	Structure and sequences of 3'UTR	102
4.2	Crystal structure of Poliovirus type 1 polymerase	103
4.3	The structure of the palm domain of 3D ^{pol}	104
4.4	Amino acid residues which are located within 5 Å of the metal ion in the active site of 3D ^{pol}	104
4.5	Putative Interactions between the N-terminal strand and the thumb sub-domain of poliovirus polymerase	105
4.6	Polymerase-polymerase interactions in the three-dimensional structure of poliovirus polymerase and location of the 3AB binding site.	105
4.7	Model of a poliovirus polymerase-dsRNA complex	106
4.8	Model of higher order polymerase structure.	106
4.9	The crystal structure of the HCV RdRp.	108
4.10	Model of internucleotide transfer in HCV RdRp	108
4.11	Interactions between molecules of the HCV RdRp in a crystal lattice.	109
4.12	Location of N18Y based on the crystal structure of NS5B	111
4.13	A one step growth curve to compare N18Y viruses.	111

4.14	Pet16b cloning/expression region	112
4.15A	SDS-electrophoresis gel analysis of fractions from the purification of 3D ^{pol}	114
4.15B	SDS-electrophoresis gel analysis of fractions from the purification of 3D ^{N18Y}	114
4.16	Visualised products of an <i>in vitro</i> uridylylation assay synthesised by 3D ^{pol} or 3D ^{N18Y}	116
4.17	Poly r(U) polymerase activity of the 3D ^{pol} and 3D ^{N18Y}	117
4.18	Terminal transferase activity of 3D ^{pol} and 3D ^{N18Y} measured at pH 7.5	118
4.19	Illustration of the yeast two-hybrid system	119
4.20	Cloning vector pGBKT7	121
4.21	Western blot analysis of <i>S.cerevisiae</i> Y187 extracts	123
4.22	Vector map of pHYBLEXzeo	124
4.23	Vector map of pYESTrp2	126
4.24	Turbidity as a measure of oligomerisation	130
4.25	Illustration of the process of recycling of 3D ^{pol}	133
5.1	Location of recombinant breakpoints within the genome of vaccine-derived poliovirus isolates (Hispaniola outbreak)	140
5.2	Summary of future work	146

List of Tables

1.1	Classification of positive-sense single-stranded RNA viruses	1
1.2	Classification of Picornaviruses	3
1.3	Clinical syndromes associated with the human enteroviruses	7
1.4	Known co-factors required for virus-receptor interactions	11
1.5	Cellular proteins known to interact with IRES	14
2.1	Representative sample of the cDNA library identified through sequence analysis	47
2.2	Predicted molecular weights of fusion proteins	49
2.3	Titration of 3-AT	54
2.4	Transformation efficiencies of cDNA library into <i>S.cerevisiae</i>	57
2.5	Transformation efficiencies of published yeast three-hybrid studies	57
2.6	Known sizes of binding sites of proteins	59
3.1	Table of constructs	72
3.2	Cellular nuclear proteins that accumulate in the cytoplasm of picornavirus-infected cells	89
4.1	Table of yeast two-hybrid constructs	120
6.1	Suppliers and dilutions of Antibodies used	153
6.2	Oligonucleotides	156/157
6.3	Genotypes of bacterial strains used	158
6.4	Antibiotic selection used for each plasmid	160
6.5	Genotypes of yeast strains used	166

1 Introduction

1.1 Positive–sense strand RNA viruses

RNA viruses containing positive-sense single stranded genomes include one-third of all virus genera and have evolved to replicate within prokaryote and eukaryotic cells.

Amongst the viruses which have a positive-sense single stranded genome are a number of pathogens of economic and medical importance which have made headlines in recent years. These viruses include SARS coronavirus, foot and mouth disease virus (FMDV), hepatitis C virus (HCV) and the causative agents of “winter vomiting disease” the *Caliciviridae*.

The RNA viruses are grouped into families based on common features such as genomic organisation and particle structure. Examples of families and the prototype virus associated with each family are shown in table 1.1. Families of viruses that contain a single-stranded positive-sense RNA genome can be divided into two groups: those viruses which surround their protein capsid with a lipid bilayer and those that do not.

1.1.1 Enveloped viruses with positive-sense RNA genomes

RNA viruses with a single-stranded positive-sense genome that have an enveloped particle include the *Togaviridae*, *Flaviviridae* and *Coronaviridae* families.

Viruses in this group contain a protein capsid that has been encased in a lipid bilayer derived from host-cell membranes. Viruses within this group encode glycoproteins, which form characteristic projections or 'spikes' in the envelope, when the particles are visualised by electron microscopy (EM). The glycoproteins interact with the cellular receptor and uncoating of viruses within this group occurs as a result of fusion between the lipid bilayer encasing the virus particle and the plasma membrane.

Virus family	Genus	Type species	Host
<i>Narnaviridae</i>	Narnavirus	Saccharomyces cerevisiae 20S narnavirus	Yeast
<i>Leviviridae</i>	Levivirus Allolevirus	Enterobacteria phage MS2 Enterobacteria phage QB	Bacteria
<i>Barnaviridae</i>	Barnavirus	Mushroom bacilliform virus	Fungi
<i>Picornaviridae</i>	Enterovirus Rhinovirus Hepatovirus Cardiovirus Aphthovirus Parechovirus Erbovirus Kobuvirus Teschovirus	Poliovirus 1 Rhinovirus 1A Hepatitis A virus Encephalomyocarditis virus Foot-and-mouth disease virus O HpeV-1 and 2 Equine Rhinitis virus Aichi virus Porcine Teschovirus	Vertebrates
<i>Thedthoviridae</i>	Cricket paralysis-like viruses	Cricket paralysis virus	Invertebrates
<i>Flaviviridae</i>	Hepacivirus Pestivirus Flavivirus	Hepatitis C virus Bovine diarrhea virus Yellow fever virus	Vertebrates
<i>Caliciviridae</i>	Vesiculovirus Norwalk-like viruses Lagovirus	Swine vesicular exanthema virus Norwalk virus Rabbit hemorrhagic disease virus	Vertebrates
<i>Togaviridae</i>	Alphavirus Rubivirus	Sindbis virus Rubella virus	Vertebrates
<i>Tombusviridae</i>	Tombusvirus Carmovirus	Tomato bushy stunt virus Carnation mottle virus	Plants
<i>Potyviridae</i>	Potyvirus Bymovirus	Potato virus Y Barley yellow mosaic virus	Plants
<i>Sequiviridae</i>	Sequivirus Waikavirus	Parsnip yellow fleck virus Rice tungro spherical virus	Plants
<i>Comoviridae</i>	Comovirus Nepovirus	Cowpea mosaic virus Tobacco ringspot virus	Plants

Table 1.1 Classification of Positive-stranded RNA viruses

Information obtained from www.ncbi.nih.gov/ICTV

1.1.2 Non-enveloped viruses with positive-sense RNA genomes

RNA viruses with single-stranded positive-sense genomes that are non-enveloped include the *Caliciviridae* and a large number of plant viruses. However viruses of this type are typified by the *Picornaviridae*, which will be described in more detail in the following sections.

Viruses within this group do not possess a lipid-containing envelope and as a consequence interaction of viruses within this grouping and their cognate cellular receptors is mediated directly by contact between the receptor and the protein capsid and not indirectly via the use of a glycoprotein.

1.1.3 Replication of positive-sense single-stranded viruses

All known RNA viruses encode an RNA-dependent RNA polymerase (RdRp) that replicates the input genome. Replication of the viral genome cannot begin until the polymerase and supplementary proteins involved in the replication process have been synthesised.

All RNA viruses replicate their genome on membrane structures derived from host-cell membranes and modified by viral proteins to form functional replication complexes. The membranous replication complexes function to protect the input genome, intermediate-replication structures and the progeny-genomes from cellular RNAases and from activating anti-viral cell-responses.

Synthesis of progeny RNA from the input genome occurs only following the cessation of translation. Thus RNA viruses must encode a “molecular switch” that alters the function of the genome from acting as mRNA to providing a template for the RdRp to synthesize a negative-sense RNA copy. Replication of the positive-sense single-stranded (+ve ss) genome from the negative-sense RNA copy is asymmetric in nature generating between 10 and 100 fold excess of progeny strands over the anti-genome. In all (+ve ss) RNA viruses the negative-sense strand is only ever found in double-stranded or partially double-stranded forms.

1.2 Picornaviridae

The picornaviruses are amongst the smallest ribonucleic acid containing viruses identified and were named to reflect this observation i.e. pico (a small unit of measurement $\{10^{-12}\}$ RNA virus). Currently the picornavirus family is subdivided into nine genera. These are the enteroviruses, rhinoviruses, cardioviruses, aphthoviruses, parechoviruses, hepatoviruses, erboviruses, kobuviruses and teschoviruses (Table 1.2). Classification is based on common genomic organisation and particle structure.

The enterovirus genus consists of a large group of human pathogens that cause a wide range of clinical conditions ranging from mild to severe (see section 1.2.1.1). As the name suggests they are associated with infection of the alimentary canal. On the basis of antibody neutralisation studies approximately 70 serotypes of enteroviruses have been identified so far (Pallansch & Roos, 2001). It has been proposed that the Enterovirus genus can be subdivided into nine species. These are poliovirus (PV), Human enterovirus A (HEV-A), Human enterovirus B (HEV-B), Human enterovirus C (HEV-C), Human enterovirus D (HEV-D), Human enterovirus E (HEV-E), Bovine enteroviruses (BEV), Porcine enterovirus (PEV) and Simian enteroviruses (SEV) (www.ncbi.nlm.nih.gov/ICTV). The classification of the human enteroviruses will be discussed further in section 1.2.1

Members of the rhinovirus genus are morphologically very similar to the enteroviruses however they have adapted to grow in nasopharyngeal tissue. Rhinoviruses can be distinguished from enteroviruses on the basis that the former are acid labile and are adapted to grow at 34 °C. With the exception of rhinovirus 87, which uses sialic acid as a receptor, the rhinovirus genus can be subdivided into two species; the major rhinovirus (HRVA) and the minor rhinoviruses (HRVB), on the basis of receptor usage. 74 different serotypes have been classified as members of the HRV-A species and 25 serotypes constitute the HRV-B species. Rhinoviruses are the predominant causative agent for the common cold being responsible for 50% of clinical cases investigated.

Parechoviruses cause mainly gastroenteritis and respiratory infections. There are two species of parechovirus, Ljungen virus (1 serotype) and human parechovirus (two serotypes).

Genus	Species	Serotypes
Enterovirus	Poliovirus	PV 1-3
	Human Enterovirus A	CVA 2-8, 10, 12, 14, 16, EV-71
	Human Enterovirus B	CVB 1-6, CVA9, E1-33, EV69-73
	Human Enterovirus C	CVA 1,11,13,15,17-24
	Human Enterovirus D	EV-68, EV- 70
	Bovine Enterovirus	BEV-1, BEV-2
	Porcine Enterovirus A	PEV-8
	Porcine Enterovirus B	PEV-9, PEV10
Simian Enterovirus A	SEV-A	
Rhinovirus	Human Rhinovirus A	HRV 1,2,7, 8, 9 –13 etc
	Human Rhinovirus B	HRV 3-6,14,17,26 etc
Parvovirus	Human Parvovirus	HpeV-1 and 2
	Ljungan Virus	LV
Hepadnavirus	Hepatitis A virus	HAV
	Avian Encephalomyelitis Virus	AEV
Cardiovirus	Encephalomyocarditis Virus	EMCV
	Theilovirus	TMEV, VHEV
Aphthovirus	Foot and Mouth Disease Virus	FMDV-O, A ,C ,Asia1, SAT1-3
	Equine Rhinitis A virus	ERAV
Erbovirus	Equine Rhinitis virus	ERBV 1 and 2
Kobuvirus	Aichi virus	AiV
	Bovine Kobuvirus	BKV
Teschovirus	Porcine Teschovirus	PTV 1-11

Table 1.2 Classification of Picornaviridae.

The cardiovirus predominantly infect mice though they can infect humans, monkeys, pigs, elephants and squirrels. There are two species of cardiovirus, Encephalomyocarditis virus (1 serotype) and Theilovirus (2 serotypes). Theiler's murine encephalomyelitis (TMEV), a member of the Theilovirus species, can produce infections in mice that clinically resemble paralytic poliomyelitis and multiple sclerosis.

The aphthovirus genus contains only two species, foot and mouth disease (FMDV) and equine rhinitis virus (ERAV). Seven serotypes of FMDV and only one of ERAV have been identified. FMDV shows a limited tropism infecting cloven-hooved animals, for example, cows, goats, sheep and horses.

Two species form the hepatoviruses - the hepatitis A viruses and avian encephalomyelitis virus. Although there is only one hepatitis A virus (HAV) serotype there are two strains infecting humans and primates. HAV is unusual amongst the Picornavirus family as it infects the liver.

Erbovirus, Kobuvirus and Teschovirus, relatively new genera in the picornaviridae, are known to infect horses, insects, cows and pigs.

In addition a number of viruses have been identified in invertebrates (Lommel et al., 1985, Muscio et al., 1988, Toriyama et al., 1992) and from the ocean (Culley et al., 2003) that are morphologically and physio-chemically similar to picornaviruses. These are commonly referred to as "picorna-like" and have not been officially assigned to the picornavirus family.

1.2.1 Human enterovirus classification

Human enteroviruses infect millions of people worldwide every year resulting in disease ranging from mild respiratory conditions to acute flaccid paralysis. Classification of the human enteroviruses was originally carried out on the basis of the virulence of the isolated viruses in suckling mice and the human disease they were associated with. The original four categories were (a) polioviruses (poliomyelitis in humans but mostly non-pathogenic to mice), (b) coxsackie A viruses (central nervous system disease and flaccid paralysis in mice), (c) coxsackie B viruses (myocarditis, central nervous system disease also cause spastic paralysis of mice) and (d) echoviruses (not known to be associated with human

disease or pathogenesis in mice) (Melnick, 1996b). The classification of the enteroviruses was changed with the identification that echoviruses could be associated with human disease. Currently the classification of enteroviruses is based on a mixture of molecular analysis and biological properties of the virus. The human enteroviruses constitute five of these species: (I) PV (II) HEV-A (III) HEV- B (IV) HEV- C (V) HEV- D. Phylogenetic mapping of the relationship between human enterovirus species is shown in figure 1.1. The previous classification that placed the echoviruses as a separate species has been altered and they are now classified within the HEV-B species. In addition phylogenetic analysis (Figure 1.1) has shown that the genetic relationship between poliovirus and the viruses that constitute the HEV-C species is so close as to make them indistinguishable at the species level (Brown et al., 2003, Hyypia et al., 1997, Poyry et al., 1996).

1.2.1.1 HEV pathogenesis

Viral diseases are characterised by a specific set of symptoms and host range. The described symptoms can often be correlated with infection of specific tissue types. While the majority of infections are mild or asymptomatic the enteroviruses can cause a wide range of syndromes. These are summarised in Table 1.3. Perhaps unsurprisingly, the diseases associated with infection of humans by the enteroviruses broadly correlate with the genetic relationship identified between viruses using molecular phylogeny, for example CAV-16 and EV-71 are associated with hand, foot and mouth disease and classification places both viruses in the HEV-A species.

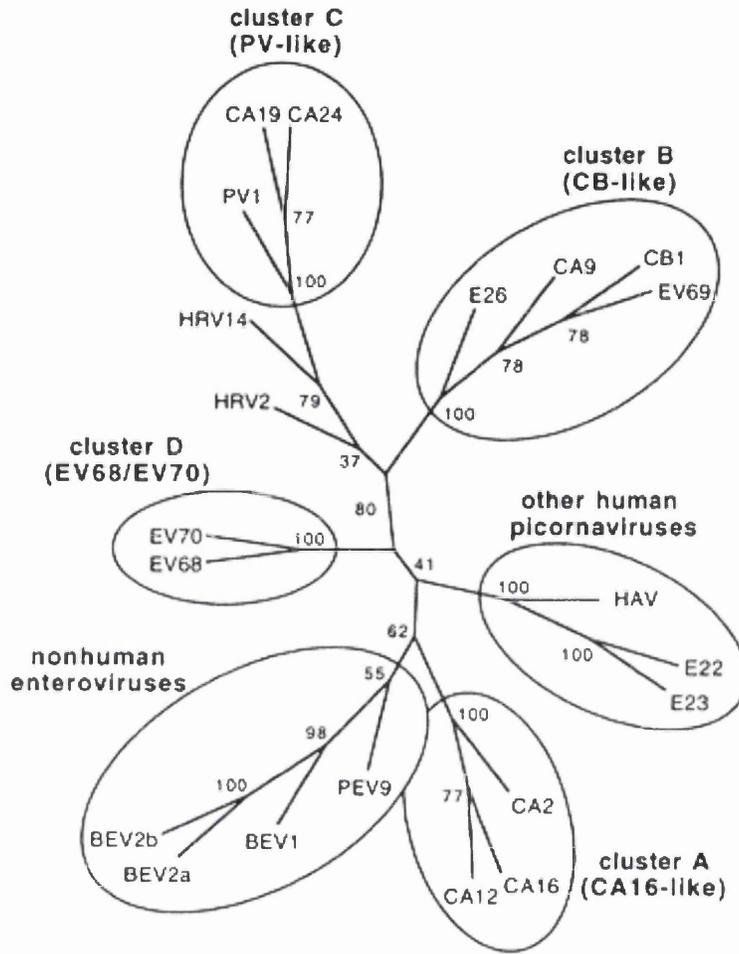


Figure 1.1. Consensus phylogenetic tree for representative human picornaviruses and nonhuman enteroviruses. Numbers at nodes represent the percentage of 100 bootstrap pseudoreplicates that contained the cluster distal to the node. Major clusters supported by bootstrap values of at least 67% are enclosed by circles. For clarity, branch lengths are not drawn to scale. Taken from Oberste et al., 1999.

Enteroviruses are transmitted by the faecal-oral route. While taking their name from the alimentary (enteric) tract that they predominantly inhabit, enteroviruses are associated clinically with diseases with CNS and cardiac involvement. Invasion of the central nervous system is a characteristic of the prototype picornavirus and human enterovirus, poliovirus, and results in acute flaccid paralysis. Coxsackie A viruses (HEV-C species) can also induce flaccid paralysis. In contrast, Coxsackie B viruses have been shown to have a propensity for heart muscle, in addition to nasopharyngeal cells. Coxsackie group B viruses (HEV-B species) are also associated with meningoencephalitis and myocarditis.

1.2.1.2 Transmission and replication of poliovirus *in vivo*

Infection with poliovirus is in over 90% of cases asymptomatic or marked by little more than a mild malaise. Paralytic poliomyelitis only occurs in approximately 1% of cases (Melnick, 1996a). As with other enteroviruses the natural poliovirus infection begins with ingestion of the virus. Following the digestive tract phase of infection, poliovirus and the other enteroviruses are found in the oropharynx and intestine. It is these cells that the virus uses to propagate itself.

Translocation of poliovirus across M-like cells found in the epithelial sheet of the Peyer patches has been demonstrated *in vitro* (Ouzilou et al., 2002). Translocation of the virus particles across M cells brings the virus into contact with cells of the immune system. This is believed to aid the spread of the virus to the cervical and mesenteric lymph nodes. The viraemia that follows the spread of the virus to the lymph nodes is believed to be an important step in the development of paralytic poliomyelitis. Neurovirulent strains of poliovirus type 1 have been shown to be able to replicate within peripheral blood mononuclear cells whilst vaccine strains cannot (Freistadt & Eberle, 1996). Following the onset of viraemia, the virus can invade the central nervous system (CNS) and replicate in neurons. Experimental evidence using transgenic mice suggests that poliovirus inoculated intravenously predominantly invades the CNS by crossing the blood brain barrier (Yang et al., 1997).

Paralytic poliomyelitis occurs as a result of the destruction of neurons following CNS invasion. The paralytic form of poliovirus has a high specificity for the anterior horn cells of the spinal cord. Although perhaps not the main pathway of dissemination a neural pathway for CNS infection has been reported in humans, monkeys and transgenic mice. Research using transgenic mice has shown that poliovirus particles could be transported by the fast retrograde axonal transport pathway via the sciatic nerve (Ohka et al., 1998). The recent identification of an interaction between PVR and the dynein-motor complex component Tctex-1 by Mueller *et al* has been offered as an explanation as to how poliovirus is transported from the axon to the neuronal cell body in a PVR-dependent manner (Mueller et al., 2002). A model for the invasion of the CNS using the neuronal retrograde transport pathway is shown in figure 1.2. A condition has been described which has been termed provocation poliomyelitis as it results following physical trauma during infection with poliovirus. Gromeier and Wimmer have managed to replicate the condition of provocation poliomyelitis experimentally and have shown that skeletal muscle injury stimulates the retrograde axonal transport of poliovirus thereby leading to CNS invasion and the development of poliomyelitis (Gromeier & Wimmer, 1998).

Enterovirus Serotype	Clinical diseases associated with Infection
<p align="center">Poliovirus</p> <p align="center">1-3</p>	<p align="center">Paralytic poliomyelitis and mild febrile illness</p>
<p align="center">Coxsackie A</p> <p align="center">2, 4, 7, 9, 10</p> <p align="center">5, 10, 16</p> <p align="center">21, 24</p>	<p align="center">Aseptic meningitis</p> <p align="center">Hand, foot and mouth disease</p> <p align="center">Common cold</p>
<p align="center">Coxsackie B</p> <p align="center">1-6</p> <p align="center">1-5</p> <p align="center">4, 5</p>	<p align="center">Aseptic meningitis</p> <p align="center">Myocarditis, pleurodynia, severe systemic disease in infants</p> <p align="center">Upper respiratory illness, pneumonia and post-viral fatigue syndrome</p>
<p align="center">Echovirus</p> <p align="center">All except 12,24,26,29</p> <p align="center">4,6,9,11,30</p> <p align="center">4,9,11,20,25</p> <p align="center">1,6,9,19</p> <p align="center">1,6,9</p> <p align="center">4,9</p>	<p align="center">Aseptic meningitis</p> <p align="center">Paralysis</p> <p align="center">Respiratory disease</p> <p align="center">Myocarditis</p> <p align="center">Epidemic myalgia</p> <p align="center">Hepatic disturbances</p>
<p align="center">Enterovirus</p> <p align="center">68</p> <p align="center">70</p> <p align="center">71</p>	<p align="center">Pneumonia</p> <p align="center">Acute haemorrhaging conjunctivitis</p> <p align="center">Hand, foot and mouth disease</p>

Table 1.3 Clinical syndromes associated with infection by enteroviruses (reviewed in Pallansch & Roos, 2001)

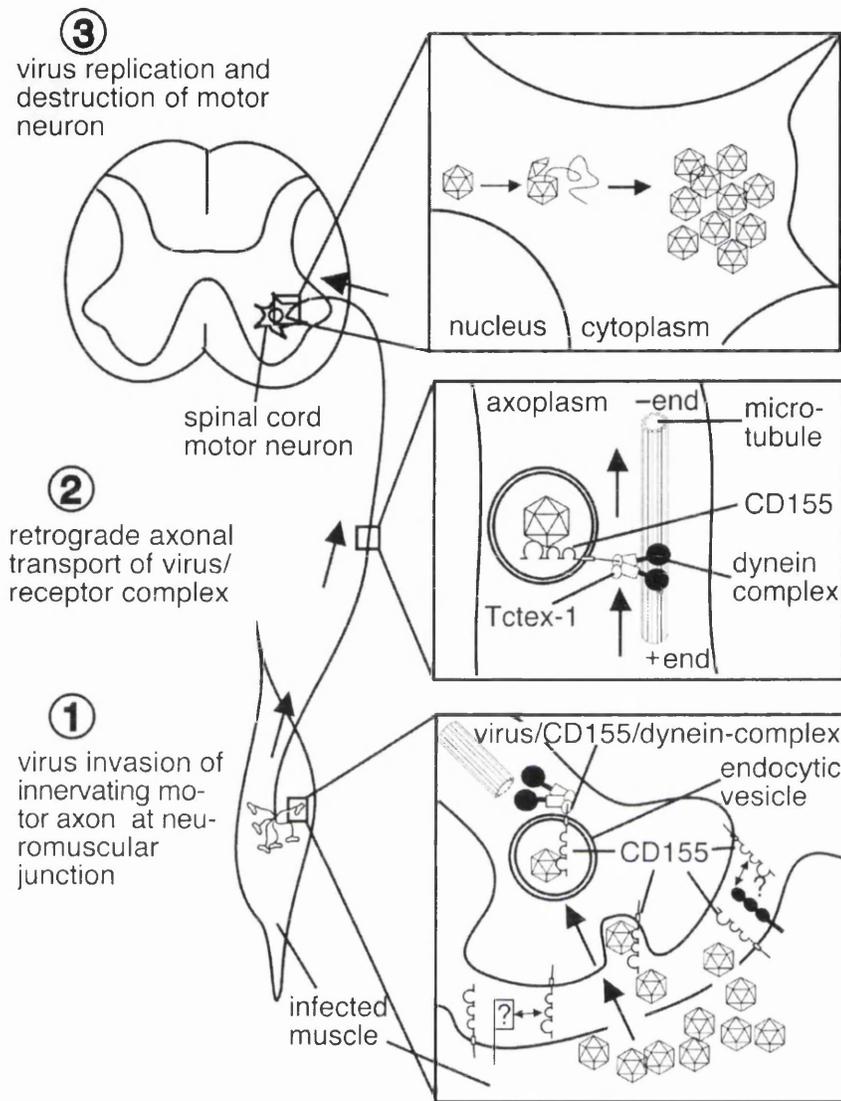


Figure. 1.2 Proposed model of CNS invasion by poliovirus. Taken from Mueller et al., 2002. 1, following virus replication in muscle, virions are released near or at the NMJ, and taken up at the presynaptic membrane of a motor axon by CD155-mediated endocytosis. Muscle injury or inflammatory responses due to the PV infection may facilitate this process by up-regulation of CD155 expression. The virus-receptor complex by interaction with Tctex-1 is targeted to the microtubular network of the axon. 2, the intact 160S virions complexed to the dynein motor by virtue of CD155 interaction with Tctex-1 are transported along microtubuli by fast retrograde axonal transport. 3, arriving at the motor neuron's cell body, the change in environment from axoplasm to cytoplasm triggers virus uncoating. Viral genomic RNA is released into the cytoplasm, and virus replication ensues, thereby killing the motor neuron. Paralysis of the muscle fiber formerly innervated by this motor neurons follows. Lateral spreading to neighboring spinal motor neurons may occur, which kills those neurons directly and independently of retrograde axonal transport.

1.3 *Picornaviridae* particle structure

Atomic resolution structures of the virus particle have been determined for all the picornavirus genera (reviewed within Hogle, 2002). The RNA genome is packed within an icosahedral capsid, 20-30 nm ($\sim 300\text{\AA}$) in diameter. The capsid is composed of sixty copies of each of the four viral structural proteins. VP1, VP2 and VP3 are exposed on the surface of the virion, VP4 is located beneath the surface of the virion and is found in association with the viral RNA. In entero- and rhinoviruses (Figure 1.3A) a deep gap ($\sim 20\text{\AA}$) can be observed surrounding the 3- fold axis (rosette-shape) and 5- fold axis (propeller shape). This deep gap is absent from the particles of aphthoviruses (Figure 1.3B) and cardioviruses..

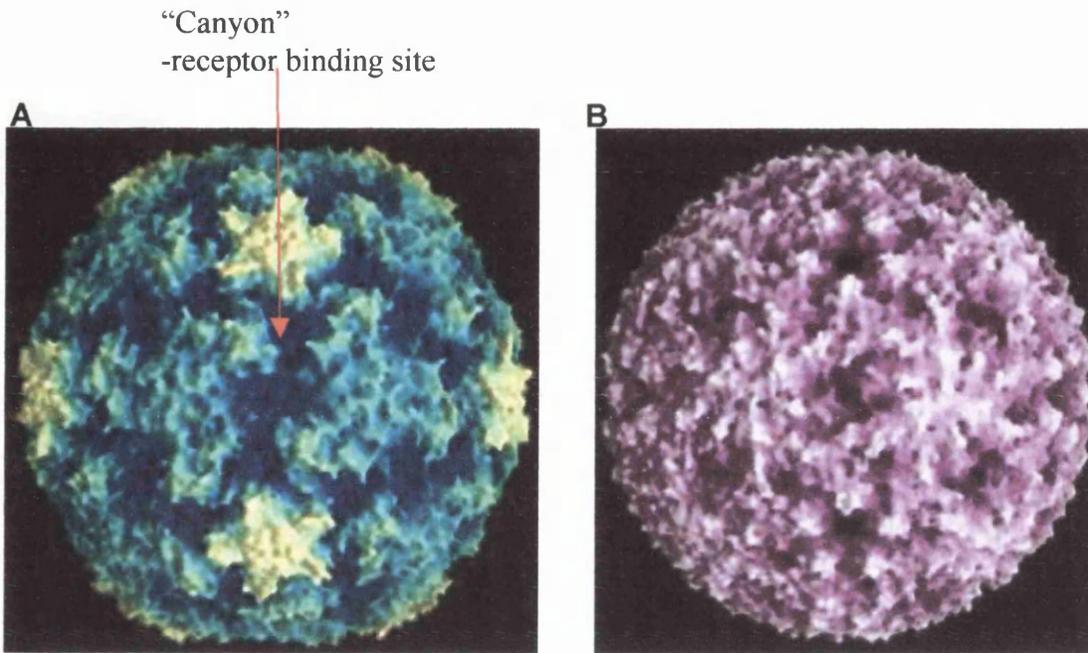


Figure 1.3 Representative structures of picornavirus virions

(A) Structure representative of the entero- and rhinovirus genera. The virion shown in this case is that of HRV14 (Rossman et al., 1985).

The “Canyon” that binds the cellular receptor is marked with an arrow. (B) Structure representative of aphtho- and cardiovirus genera. The virion shown in B is that of FMDV (Acharaya et al., 1989).

1.4 Genome Organisation

Sequencing has shown that the genomes of picornaviruses range in size from 7,209 bases (HRV14) to 8,450 bases (FMDV) in length. Despite this variation in length sequencing has shown that picornavirus genomes share a common organisation (Figure 1.4).

The picornavirus genome is unusual in that the 5' end of the genome lacks the 7-methylguanosine cap structure that is covalently linked via a 5'-5' linkage to the terminal base of the majority of cellular mRNA (Lee et al., 1977). Instead a small protein, VPg (Virion protein, genome), which varies only slightly in size between different picornaviruses, is covalently attached to the 5' terminal -pUpUp, of both genomic RNA and the negative sense replication intermediate, by a phosphodiester linkage to the phenolic hydroxyl group of a tyrosine residue (Rothberg et al., 1978b, Wimmer, 1982). VPg functions as a protein primer for the initiation of positive and negative-strand synthesis.

The protein-coding region is flanked by untranslated regions (UTR), whose sequences tend to be strongly conserved and carry *cis*-acting elements that are functionally involved in translation and RNA synthesis. Amongst the picornaviruses the 5'UTR varies in length from 624 nucleotides (nt) in length in rhinovirus 14 (HRV14) to approximately 1,200 nt in FMDV. The 5'UTR of cardioviruses and aphthoviruses also contains a poly (C) tract. The size of the poly (C) tract varies in length between viral isolates and ranges from 80-250 bases in length amongst the cardioviruses and 100-170 bases amongst the aphthoviruses.

In contrast to the 5'UTR the 3'UTR is short, ranging from 47 bases in length in HRV14 to 126 bases in length in EMCV. The 3'UTR is required for efficient replication of the viral genome. The genome is tailed by a genetically encoded poly (A) tract. The length of the poly (A) tract ranges from 35 (EMCV) to 100 (FMDV) residues in length.

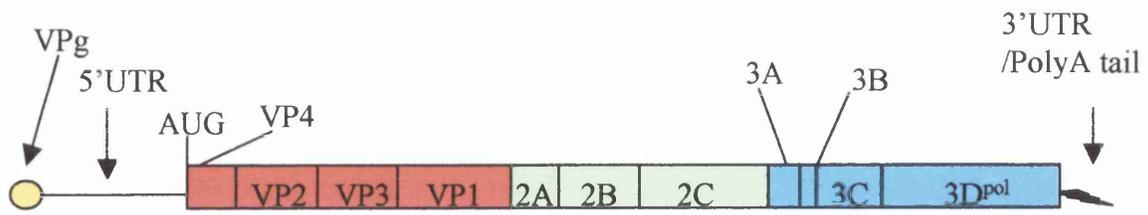
Figure 1.4 Common Picornavirus Genome Organization

For the purposes of this figure the genomic organisation of the Teschovirus and Erbovirus are not shown.

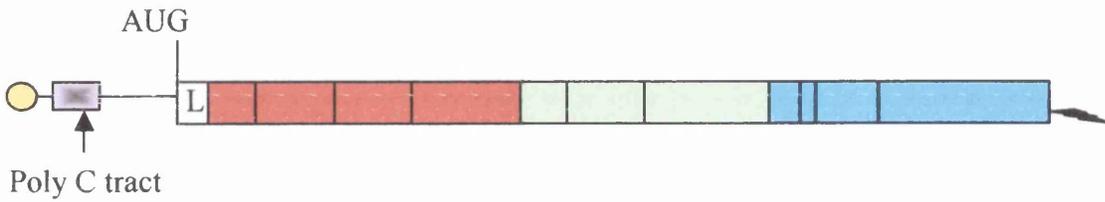
All genomes contain VPg a 22 amino acid protein attached to the proximal 5' end of the genome. The coding region is preceded by a long highly structured non-translated region (5'UTR) and at the 3'terminus of the genome by the short structured non-translated region (3'UTR) and a poly A tail. In addition Cardio- and Aphthoviruses contain a poly C tract that precedes the IRES element in the 5'UTR.

The main difference in genomic organization of the coding region of the genome are the L protein and 2A. At least four different types of L protein have been identified in the five genera that have a protein at this locus (Cardiovirus, Aphthovirus, Erbovirus, Teschovirus and Kobuvirus). There are also four structurally diverse protein types coded for by the 2A locus. Other differences of genomic organisation between the picornaviruses are the 3 copies of VPg (3B) carried by the Aphthoviruses and the substitution VP4/VP2 with VP0 in the Parecho- and Kobuviruses.

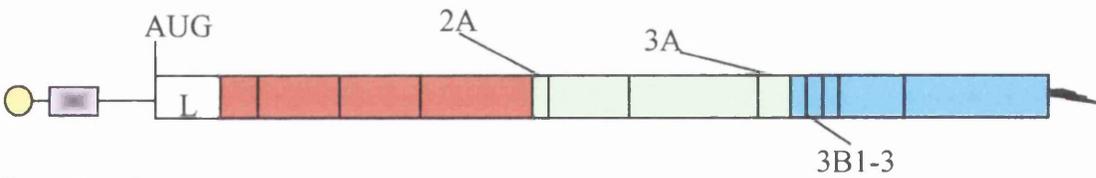
Enterovirus and Rhinovirus



Cardiovirus



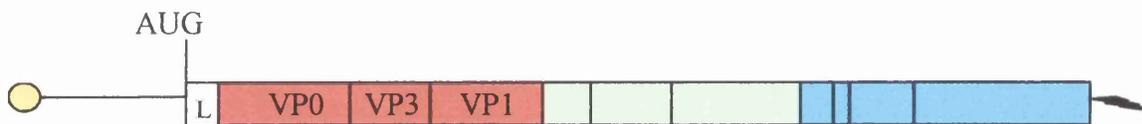
Aphthovirus (FMDV) and Erboviruses



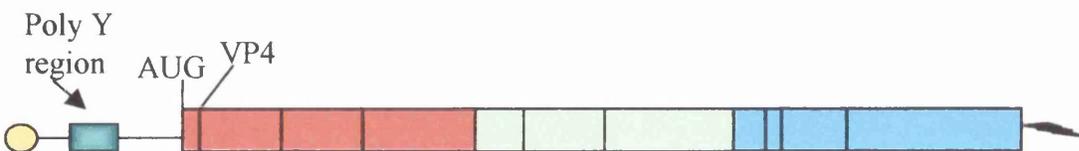
Parechovirus



Kobuvirus



Hepatovirus



1.5 Overview of the replication cycle

Multiplication of the picornaviruses occurs entirely in the cytoplasm. As a model for the picornavirus replication cycle that of poliovirus is illustrated (Figure 1.5). No significant differences from this model have so far been described for the other picornaviruses.

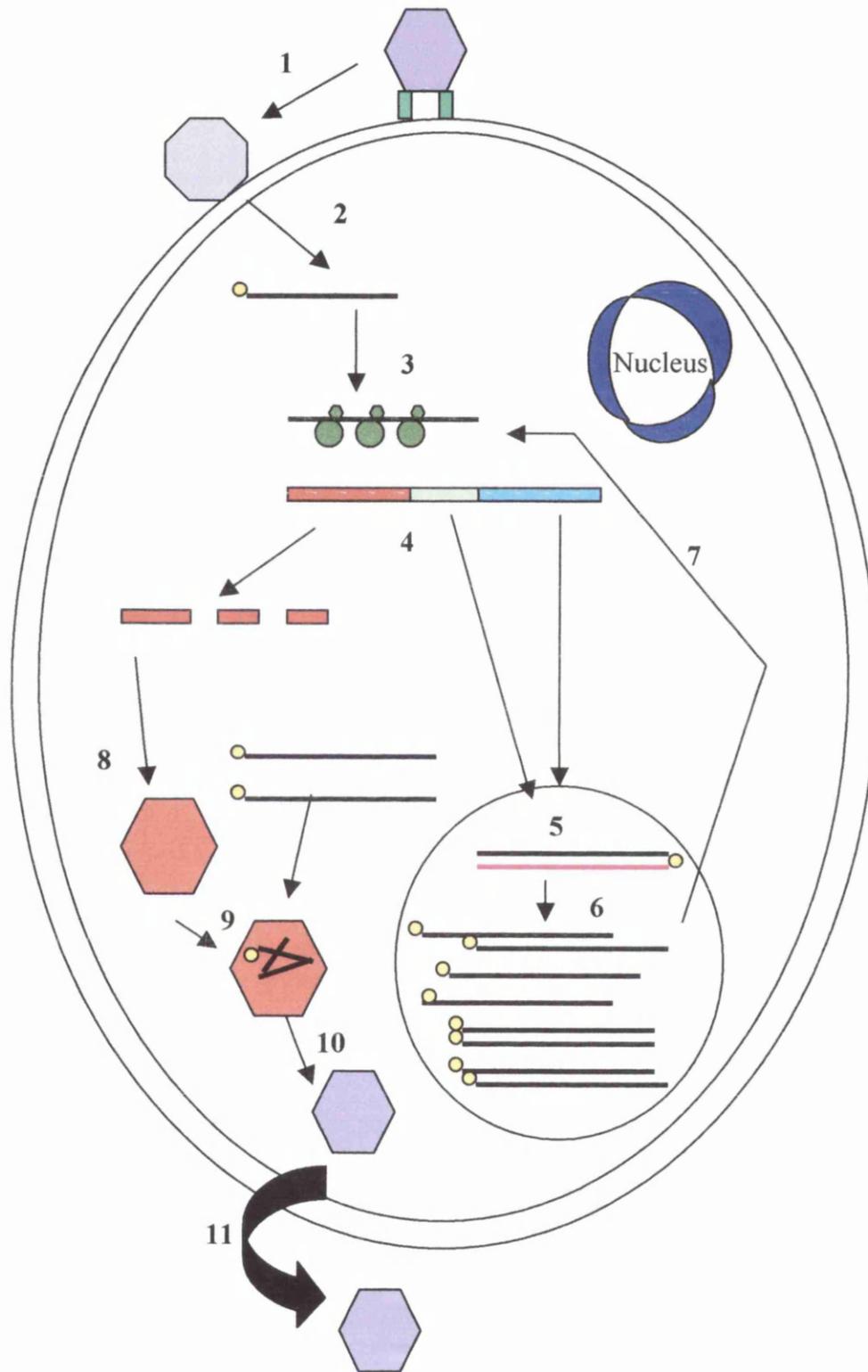
The initial event in infection is the interaction of the virion with its cellular receptor (1). The interaction of the virion with its receptor causes a conformational change. This displaces VP4 and releases the genomic RNA into the cytoplasm (2). Upon entry into the cytoplasm VPg is cleaved from the genomic RNA by a cellular enzyme, unlinking enzyme, and the genome is translated (3). Autocatalytic cleavage of the viral polyprotein releases the viral proteins (4). The first segment of the virus that is released is the P1 region, followed by P2 and P3.

Following the synthesis of the proteins encoded by the P2 region of the genome, the host cell is modified to form the viral replication complexes from the endoplasmic reticulum. The first step in the synthesis of genomic RNA is the initiation of synthesis of a minus-strand copy of the incoming RNA (5). The minus strand copy then serves as the template for synthesis of positive strands (6). In virus infected cells numerous positive strands are synthesised from every negative strand leading to the formation of a double stranded replication intermediate (RF). During the infection, newly synthesized RNA molecules are translated (7) to allow the formation of additional replication complexes. This results in a massive amplification of the number of viral genomes that are present within the cytoplasm (7 → 4 → 5 → 6).

Assembly of the virus capsid proteins into pentamers occurs spontaneously following cleavage of the P1 precursor (8). VPg-linked RNA is then packaged into the assembled pentamers by an unknown mechanism (9). The resultant provirions are not infectious. Production of infectious virions requires a maturation cleavage, of VP0 to VP4 and VP2, which occurs late on in infection. The infectious virions are released by the lysis of the host cell. The whole viral replication cycle takes between 5 and 10 hours.

Figure 1.5 Schematic representation of the poliovirus replication cycle.

Virus binds to cellular receptor (1) and the genome is delivered into the cytoplasm by uncoating of the viral particle (2). VPg is removed from the viral RNA by a host cell enzyme (unlinking enzyme) and the viral RNA is translated (3). Modification of the endoplasmic reticulum by proteins encoded by the virus results in the formation of replication vesicles where RNA synthesis occurs. (5) Viral +ve strands (Black strand) are copied by the RNA polymerase to form full length negative sense-strand (red strand). Numerous positive strands are amplified off the one negative sense template (6). An amplification cycle occurs where the newly synthesized positive strands are translated and replicated (7- 4 - 5 - 6). The capsid proteins are processed from the viral polyprotein and spontaneously assemble into non-infectious procapsids (8). The positive strand genomes are packaged into the preformed capsids (9). Cleavage of VP0 causes a conformation change in the capsid structure that results in the formation of fully functional virus (10). The infectious virions are released upon host cell lysis (11). The entire replication cycle takes between 5 and 10 hours.



Key:

Mature Virion		Anti-genome		Ribosomes	
Immature Virion		Genomic RNA		Cellular receptor	
"A" particle		VPg			

1.6 Viral attachment

The receptors for poliovirus, coxsackie B virus, echovirus and the major rhinoviruses have all been mapped to human chromosome 19 (Rueckert, 1996). Wide ranges of cellular proteins have been identified as receptors for the *Picornaviridae*. These include proteins of the integrin, SCR-like and immunoglobulin (Ig) lineages (Figure 1.6). Additionally, while polio-, rhino-, cardio- and hepatoviruses appear not to require additional factors for cellular infection to occur, this is not the case for all enteroviruses. The requirement for accessory factors amongst the *Picornaviridae* is summarised in table 1.4.

Attachment of poliovirus to cells is mediated solely by the poliovirus receptor (PVR). PVR is an 80kDA surface glycoprotein that belongs to the Ig-superfamily. Although PVR usage appears to be a defining characteristic of poliovirus, the receptor usage of the other picornaviruses does not always correlate with the phylogenetic relationship between the viruses. Cellular proteins used as receptors by more than one picornavirus genus include ICAM-1 (major group rhinoviruses and CAV 13, 18 and 21), $\alpha_v\beta_3$ vitronectin (FMDV, CAV 9 and echovirus 22) and sialic acid (rhinovirus 87 and bovine enterovirus).

Within individual genera it is the enteroviruses that show the most diversity in receptor choice. The use of proteins of the integrin, SCR-like and Ig-like lineages as the cellular receptor has been described for representatives of the enterovirus genera. The cellular protein identified as a receptor of enteroviruses with the most frequency is decay-accelerating factor (DAF), otherwise known as CD55. DAF has been identified as the receptor for all haemagglutinating strains of echovirus (Powell et al., 1997), CAV21 (Shafren et al., 1997), CBV 1, 3 and 5 (Bergelson et al., 1995, Shieh & Bergelson, 2002) and enterovirus 70 (Karnauchow et al., 1996). Recent work has shown that the binding sites on DAF of the haemagglutinating enteroviruses are different (Williams et al., 2003). This provides evidence that the use of DAF has arisen through a process of convergent evolution. This suggests that there must be a selective advantage in using DAF as a cellular receptor. If the haemagglutinating enterovirus-DAF interaction is reflective of all picornavirus-receptor interactions this would suggest that the cellular proteins used must positively influence virus infection in some way. It has been suggested that benefit to the virus may include enhancing virus entry via receptor internalisation or the recruitment of

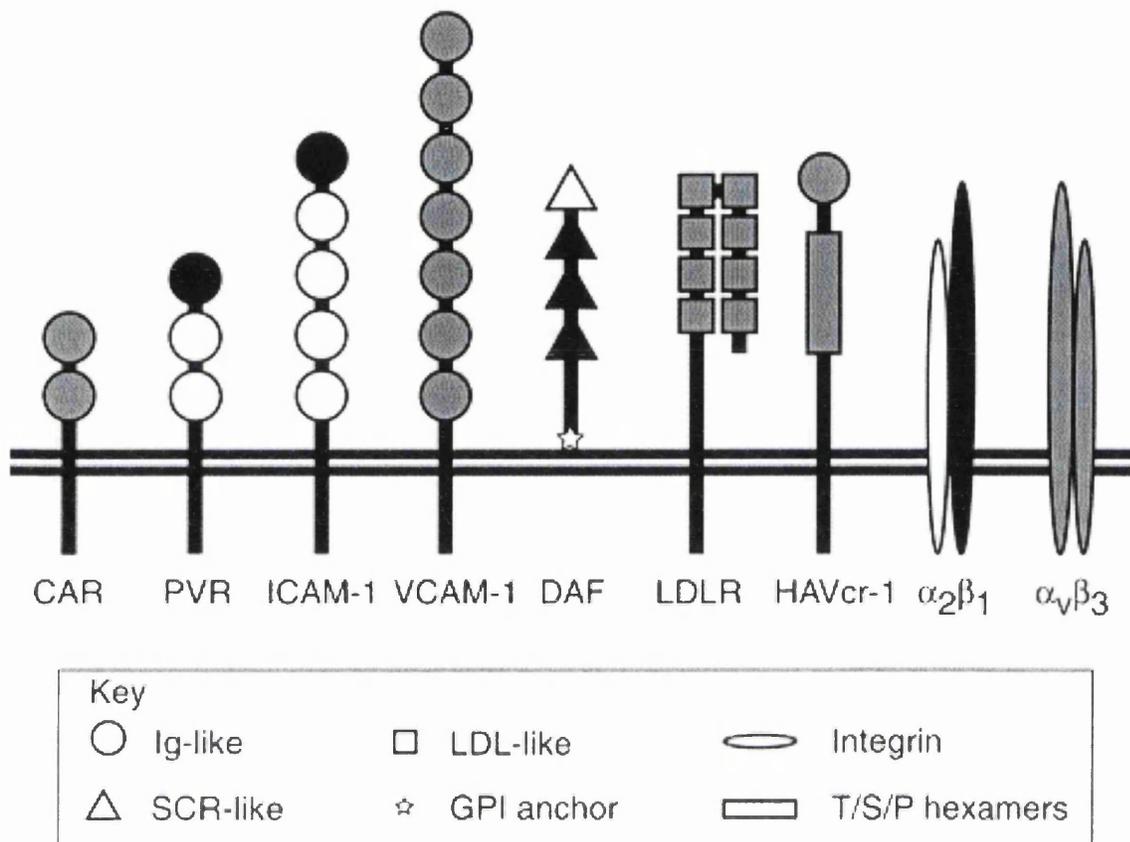


Fig. 1.6 Schematic representation of the proteins known to function as picornavirus receptors. Figure taken from Evans and Almond, 1998.

Domains (of Ig-like, SCR-like or LDL-like folds) implicated in virus binding are depicted in black. Domains that are not shaded are not involved in virus binding. Domains for which no information is available are shown in grey. Abbreviations: CAR, coxsackie–adenovirus receptor; DAF, decay-accelerating factor; GPI, glycosylphosphatidylinositol; HAVcr-1, hepatitis cellular receptor type 1; ICAM-1, intracellular adhesion molecule type 1; LDL, low density lipoprotein; LDLR, low density lipoprotein receptor; PVR, poliovirus receptor; SCR, short consensus repeat; T/S/P, threonine/serine /proline; VCAM-1, vascular cell adhesion molecule type 1.

Virus name and serotype	Receptor	Receptor lineage	Accessory factors
<u>Aphthovirus</u> Foot and Mouth disease virus	$\alpha_3\beta_5$ Vitronectin receptor	Integrin	Heparin sulphate proteoglycan
<u>Cardiovirus</u> Encephalomyocarditis virus	VCAM-1	Ig-like	N/A
Encephalomyocarditis virus	Sialylated glycoporphin A	Carbohydrate	N/A
<u>Enterovirus</u> Poliovirus 1-3	PVR	Ig-like	N/A
BEV	Sialic acid	Carbohydrate	N/A
CAV 13,18,21	ICAM-1	Ig-like	N/A
CAV 9	$\alpha_v\beta_3$ Vitronectin receptor	Integrin	N/A
CAV 21	DAF	SCR-like	ICAM-1
CBV 1,3,5	DAF	SCR-like	$\alpha_v\beta_6$ Integrin
CBV 1-6	CAR	SCR-like	N/A
EV 1	$\alpha_2\beta_1$ Integrin	Integrin	β_2m
EV 3,6,7,11-13,20,21,24,29,33	DAF	SCR-like	β_2m
EV22	$\alpha_v\beta_3$ Vitronectin receptor	Integrin	N/A
Enterovirus 70	DAF	SCR-like	N/A
<u>Hepatovirus</u> HAV	HAVcr-1	Ig-like and mucin-like	N/A
<u>Rhinovirus</u> Major group	ICAM-1	Ig-like	N/A
Minor group	LDLR	LDLR	N/A
Rhinovirus 87	Sialic acid	Carbohydrate	N/A

Table 1.4 Picornavirus receptors and accessory molecules involved in cell infection.

Table was taken from Evans and Almond, 1998. Abbreviations used: β_2m , β_2 -microglobulin; CAR, coxsackievirus and adenovirus receptor; DAF, decay-accelerating factor; HAVcr-1, hepatitis cellular receptor type 1; N/A, not applicable (no accessory molecules are currently implicated in virus infection); PVR, poliovirus receptor; SCR, short consensus repeat (s); VCAM-1, vascular adhesion molecule type 1.

naïve cells and activation of the apoptotic response to facilitate virus release (Evans & Almond, 1998).

1.6.1 Viral entry and uncoating

The interaction with the receptor ultimately must result in the release of the genome into the cytoplasm of the cell. Studies using poliovirus have shown that the interaction between the canyon region of the poliovirus particle and the PVR causes a conformational change in the virion. The virus particle, upon interaction with the cellular receptor, extrudes VP4 and the N-termini of VP1, that are located in the interior of the virion, near the 5-fold axis of symmetry (Hogle, 2002).

A model for poliovirus entry has been proposed where the myristylated VP4 and the N-termini of VP1 become embedded within the plasma membrane. In this model proposed by Belnap *et al* (Belnap *et al.*, 2000) symmetrical expansion results in the alteration of the location of VP3 within the particle. The viral genome is extruded from the particle through the transmembrane pore formed through these conformational changes (Figure 1.7).

Research into the uncoating abilities of mutants of VP4 by Danthi *et al* has recently provided supporting evidence for this model (Danthi *et al.*, 2003). Whilst poliovirus has no requirement for receptor-mediated endocytosis to mediate cell entry (Perez & Carrasco, 1993), receptor-mediated endocytosis has been demonstrated to occur in several rhinovirus serotypes (Nurani *et al.*, 2003, Schober *et al.*, 1998).

The conformational changes that occur following the virus-receptor interaction can be monitored by changes in the sedimentation coefficient of the particles. The poliovirus virion, which has a sedimentation coefficient of 160 S, shows a reduced sedimentation coefficient after undergoing conformational change (135 S). The 135 S particle is more commonly referred to as the A particle. The empty particle that sloughs off from the cell following release of genomic RNA has a sedimentation coefficient of 80S. The uncoating procedure is not efficient, it is estimated that between 50-90% of all attached particles are eluted or sloughed off abortively. As a consequence of this, picornaviruses are characterised by a high particle:infectivity ratio.

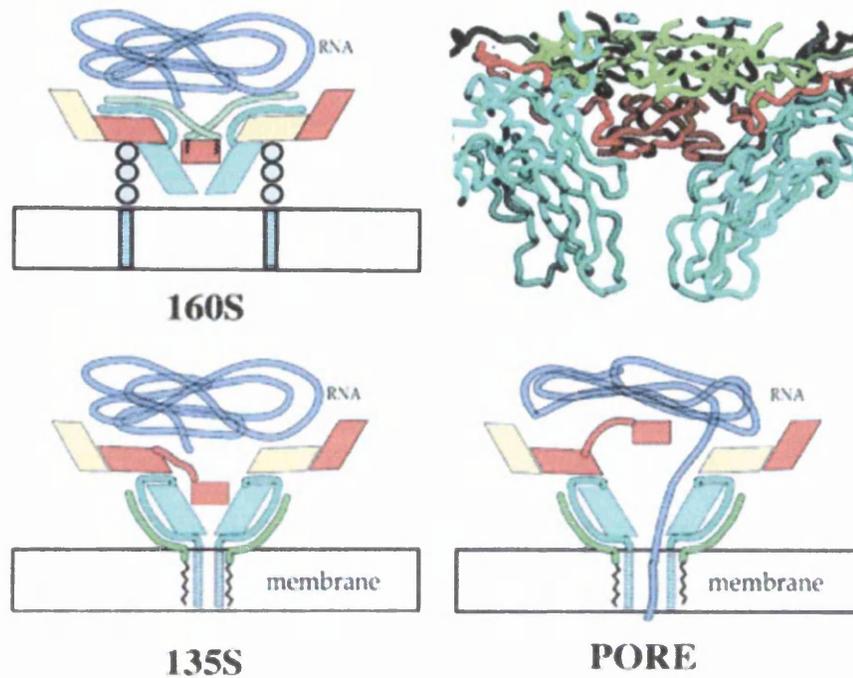


Figure 1.7. A possible mechanism for uncoating of poliovirus RNA.

Taken from Belnap et al., 2000.

VP2, VP3, and VP4 are colored cyan, yellow, red, and green, respectively. In the crystal structure of the virion (upper right), the beta-tube of VP3 (red) forms a plug at the fivefold axis that separates the virus interior from the outer surface. Attachment of the 160S particle (upper left) to the poliovirus receptor (three gray circles) triggers conversion to the 135S form (lower left). Upon conversion, cell attachment is mediated by externalized VP4 (green tubes) and the N termini of VP1 (blue tubes). The N termini emerge from the bottom of the canyon and extend along the sides of the fivefold mesa towards the apex. Once the N-terminal helices of VP1 have inserted into the membrane, they rearrange to form a pore (lower right). To permit the RNA (purple tube) to pass through the pore into the cytoplasm, it would be necessary for the VP3 beta-tube (red rectangle) to shift on its 40-residue tether (red tube) and for the VP1 barrels to splay farther apart.

1.7 Genome translation

Upon infection the uncapped picornavirus genome is directly used for translation. The use of a cap independent pathway for translation was first described in picornaviruses in 1988 (Pelletier et al., 1988, Pelletier & Sonenberg, 1988). Since this time the use of cap-independent pathways has been described in other RNA viruses (hepaciviruses and pestiviruses), cellular RNAs and the coding regions of some retroviruses.

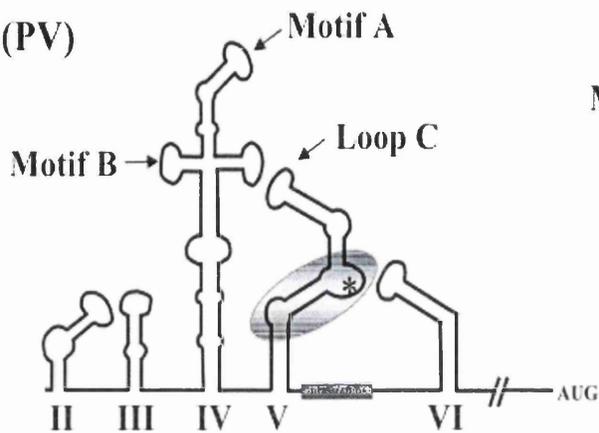
1.7.1 IRES

The ability of picornavirus genomes to bind the ribosomal subunits in the absence of a cap structure resides in the presence at the 5'UTR of a complex *cis*-acting element called the internal ribosome entry site (IRES).

The IRES element is a highly ordered structure of approximately 450 nucleotides in length that directs the assembly of the ribosomes close to the initiation codon. The IRES's of the *Picornaviridae* can be roughly subdivided into three groups on the basis of biochemical analyses. The entero- and rhinovirus genera show a type I IRES, whilst the aphtho- and cardiovirus genera contain a type II IRES. The IRES of hepatitis A virus is distinct and is referred to as a type III IRES. Despite biochemical and phylogenetic differences it has been shown recently that all picornaviruses IRES's are morphologically similar when analysed using transmission electron microscopy (Beales et al., 2003).

The type I IRES consists of six domains numbered I to VI. The structure predicted to form (Figure 1.8A) has been supported by chemical and enzymatic analysis. Mutational analysis of the poliovirus IRES has demonstrated that domains II, IV and V are essential for translation. In contrast, domain III can be removed in its entirety without any deleterious effect on translation (Dildine & Semler, 1989, Nicholson et al., 1991, Pilipenko et al., 1989b). Indeed BEV strains, which have a type I IRES, do not contain domain III.

**Type I
(PV)**



Type II

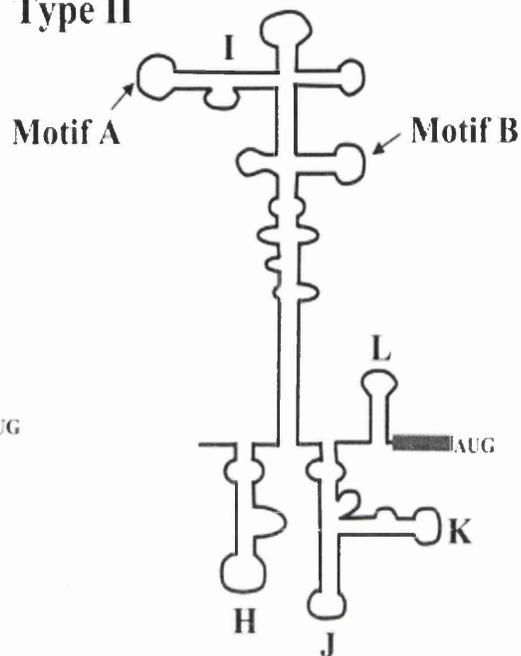


Figure.1.8 Schematic representation of the predicted secondary structure of type I and type II viral IRESes. Taken from Kean, 2003. Cartoons are shown of the entero/rhinovirus (type I) and cardio/aphthovirus (type II) IRESes. Letters or roman numerals identify individual domains. Grey rectangles represent pyrimidine-rich sequences and the initiation codons used are shown for each type of IRES. Arrows designate conserved motifs: A = GNRA tetraloop (GYGA for PV); B and C=C-rich loops (AACCA and AAACCA for PV). An attenuating mutation in the vaccine strains of each PV serotype lies within the grey oval. A star highlights a large lateral bulge-loop (UCGUAACGCGCAAG for PV type 1).

The consensus structure of the type II IRES (Figure 1.8B) typified by EMCV and FMDV consists of twelve stem-loops that have been termed domain A-L. This consensus structure has been confirmed using mutational analysis, chemical and enzymatic probing (Jang & Wimmer, 1990, Pilipenko et al., 1989a). Mutational analysis has shown the critical elements of the EMCV type II IRES to be I and J and the single-strand loops in the J and K domains. The IRES of parechovirus though classified as a type II IRES has no requirement for domain H and L.

Despite the differences in structure between the IRES a number of common structural features are shared between the type I, II and III IRES elements. The IRES elements all contain an Yn-Xm-AUG motif at the 5'UTR boundary (Pilipenko et al., 1992a). In this motif the Yn is a pyrimidine rich region and Xm is a spacer of 15-25 nt. Analysis of this motif has shown that the number of residues in the pyrimidine rich region and the length of the spacer region appear to be more important than the precise sequence (Jang & Wimmer, 1990). Both the type I and type II IRES's contain a GNRA tetraloop and A/C rich regions. No specific function has yet been identified for the motifs though it has been postulated that the GNRA tetraloop mediates protein /RNA interactions. It has also been suggested that the A/C rich region, in the type I IRES, may form a pseudoknot (Le et al., 1992). In addition to these common structural motifs a similarity exists, in the relative position of the structure to the AUG and the secondary structure of the remainder of the IRES, between domain V of the type I IRES and the Y shaped domain of the type II IRES. Nearly all the mutations affecting the structure of domain V abrogate translation (Dildine et al., 1991, Haller & Semler, 1992, Kuge & Nomoto, 1987, Pelletier et al., 1988).

In the majority of picornaviruses, host cell translation is shut-off by the virus through the cleavage of the cap binding protein, eIF4E, from eIF4G. This cleavage is mediated by the viral protease, 2A (Krausslich et al., 1987) or, the functional analog of 2A, the L protease. The cellular initiation factors, except eIF4E, but including the cleaved form of eIF4G appear to be an essential requirement for the initiation of translation at the IRES (reviewed in Belsham & Sonenberg, 2000, Pestova et al., 1996). In addition the picornavirus IRES appears to require additional cellular proteins, not generally associated with translation, for efficient translation. The requirement for these non-canonical initiation factors differs between IRES groups and appears to have some effect with respect to determining cellular

Cellular protein	Virus	Binding site within IRES	References
La autoantigen	Coxsackie B3		(Ray & Das, 2002)
La autoantigen	PV	Domain VI	(Meerovitch et al., 1993), (Meerovitch et al., 1989)
La autoantigen	EMCV		
PTB	EMCV		(Jang & Wimmer, 1990, Kaminski et al., 1995)
PTB	PV		
Unr	HAV		(Graff et al., 1998)
Unr	HRV		(Hunt et al., 1999)
Unrip	HRV		
PCBP2	PV	Stemloop B of cloverleaf and Domain IV of IRES	(Blyn et al., 1997) (Parsley et al., 1997)
eIF4G	PV	Domain V of IRES	(Ochs et al., 2003)
eIF4G	EMCV	Domain J-K-L	(Ochs et al., 2003)
eIF4G	FMDV	Domain IV	(Ochs et al., 2003)
eIF4B	PV	Domain V and Domain VI	(Ochs et al., 2002)
eIF4A	FMDV	Domain IV	(Ochs et al., 2002)
eIF4A	EMCV	Domain J-K-L	(Ochs et al., 2002)

Table 1.5 *Trans*-activating cellular proteins required for IRES function.

tropism and viral pathogenicity. Poliovirus appears to have more complex demands for non-canonical factors than the type II IRES of foot and mouth disease (FMDV) and encephalomyocarditis (EMCV). The requirement for non-canonical translation factors in IRES mediated translation amongst the *Picornaviridae* is summarized in Table 1.5. Due to the complexity of the requirements that appear to be necessary for the trans-activation of the IRES the way these proteins interact with the IRES has not been fully deduced. The pattern of binding of the required cellular factors is much more defined for the EMCV and FMDV, which appear to have substantially lower requirements.

The ability of the IRES to interact with cellular factors appears to be a major determinant of neurovirulence. Analysis of the reversions of the Sabin 3 strain identified a reversion at nt 472, of a uridine residue to a cytosine, that could be demonstrated to restore neurovirulence (Evans et al., 1985). The presence of a uridine at nt 472 was demonstrated to cause a 10-fold reduction in translation in neuroblastoma cells, though translation levels of the same strain were comparable with wildtype levels in HeLa cells (Svitkin et al., 1985). Recently this mutation has been demonstrated to abrogate the binding of the canonical initiation factor eIF4G (Ochs et al., 2003). Mutational experiments have also demonstrated that mutation of PTB binding sites also selectively reduces translation and replication efficiencies of poliovirus in neuronal cells but not in HeLa cells (Gutierrez et al., 1997). This effect is mediated by the reduced binding of the neuronal PTB isoform compared with the ability of cellular PTB form to bind the mutated structures.

1.8 Processing determinants and Cleavage products

The translated genome produces a large polyprotein ranging in length from 2,178 amino acid (aa) residues to 2,332 aa residues. The full-length polyprotein is seldom detected within infected cells as cleavage of the protein occurs co-translationally.

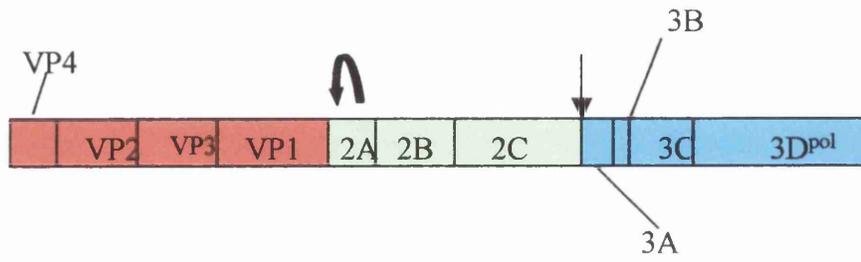
The picornavirus polyprotein can be divided into 3 regions, designated P1, P2 and P3 based on function. The P1 region corresponds to the capsid protein precursor, the middle region of the polyprotein, designated P2, contains the non-structural proteins 2A-2C, which are predominantly involved in alteration of the cellular architecture. The most C-terminal section of the polyprotein, designated P3, encodes the non-structural proteins 3A-3D^{pol} that are all involved in RNA synthesis.

Initial cleavage of the polyprotein (Figure 1.9) mediated by 2A^{pro} at the P1- 2A junction results in the separation of the structural (P1) and non-structural proteins (P23) (Toyoda et al., 1986). In enteroviruses and rhinoviruses this cleavage is proteolytic. In aphtho- and cardioviruses 2A this cleavage is not proteolytic and has been proposed to occur via a novel translational skip mechanism at the carboxy-terminus (Donnelly et al., 2001b) (Donnelly et al., 2001a).

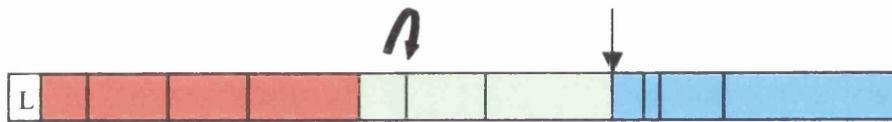
Further processing of the P23 ORF, by 3C^{pro} and 3CD^{pro}, results in the production of numerous intermediate and mature viral proteins that are required for a complete infectious cycle (Figure 1.10).

Figure 1.9 Primary processing of the picornavirus polyprotein. In cells infected with enteroviruses and rhinoviruses, the P1 region is cleaved from P2 by 2A^{pro}. In cells infected with cardiovirus and aphoviruses the P1/P2 junction is cleaved by 3C^{pro}. The 2A protein of cardioviruses and aphoviruses are not proteinases, but cause their own cleavage by a novel “translational skip” mechanism (Donnelly et al., 2001). In all viruses shown the P2/P3 junction is cleaved by 3C^{pro}. In hepatoviruses and parechoviruses, 3C^{pro} also carries out the primary cleavage at the 2A/2B junction. The L^{pro} proteinase autocatalytically releases itself from VP4. Unless otherwise stated the schematics are labelled according to the diagram of the entero- and rhinovirus polyprotein. Cleavages carried out by 3C^{pro} are shown by a straight arrow.

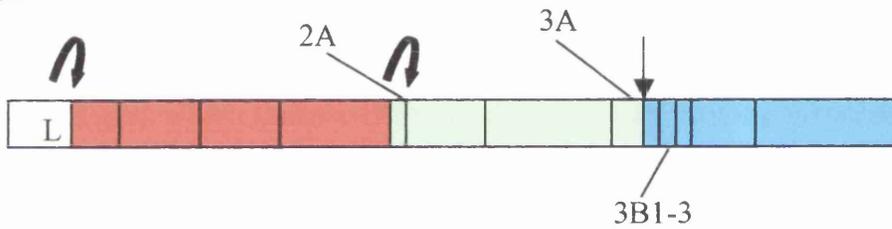
Enterovirus and Rhinovirus



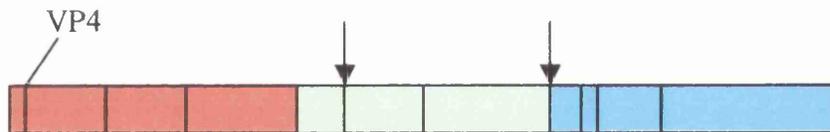
Cardiovirus



Aphthovirus



Hepatovirus



Parechovirus



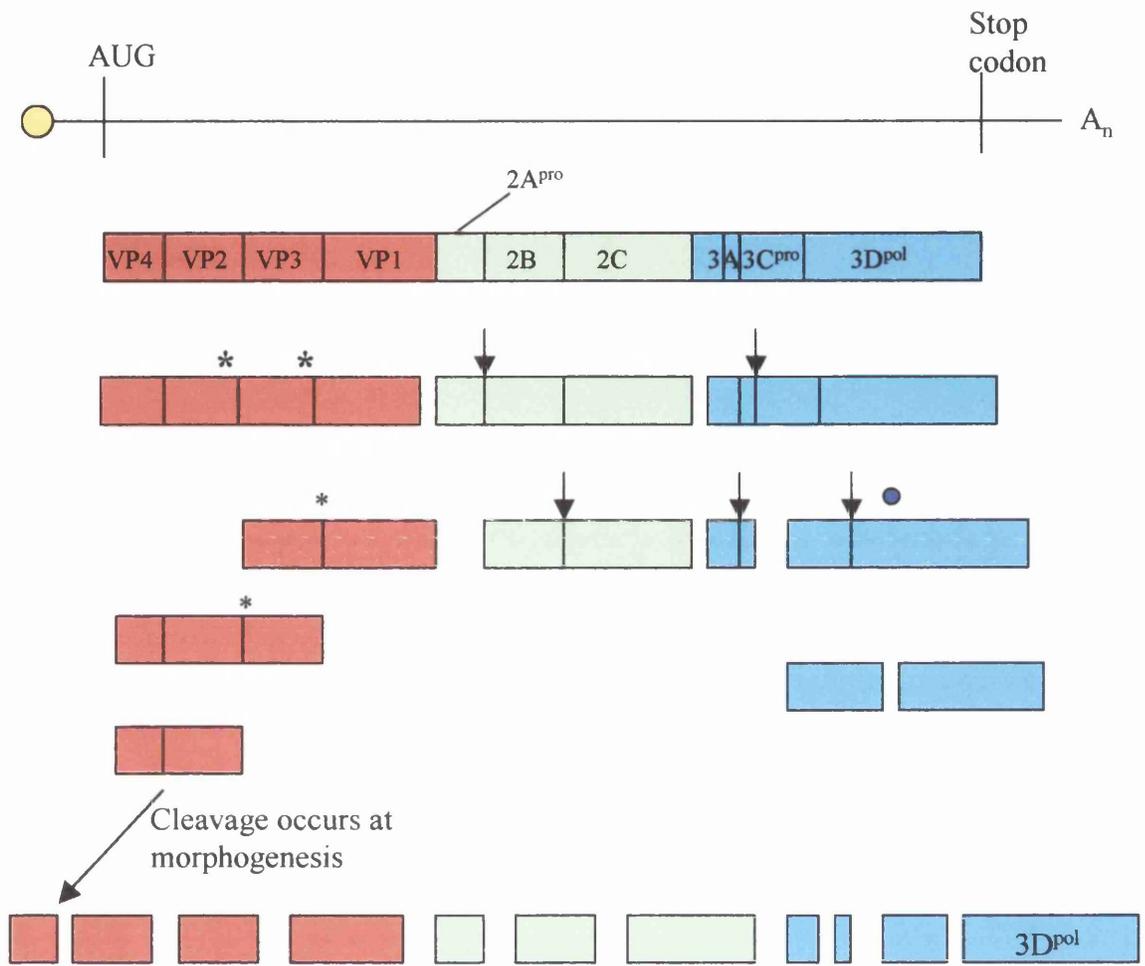


Figure 1.10. Secondary processing event of the poliovirus polyprotein

Secondary cleavages carried out by 3CD^{pro} are marked with a *. Alternative secondary cleavage carried out by 2A^{pro} to generate 3C' and 3D' is marked with a blue sphere. Cleavages carried out by 3C^{pro} are marked with a black arrow. In other picornaviruses, the secondary processing all secondary cleavages are mediated by 3C^{pro}.

1.8.1 The structural proteins

As previously outlined, upon release into the cytoplasm the viral RNA is translated and processed cotranslationally by viral proteases. Prior to processing, myristic acid is covalently linked to the N-terminus of the polypeptide (Chow et al., 1987). The presence of myristic acid covalently linked to the N-terminal glycine is required for the assembly of pentamers (14S subunit), an intermediate in the process of virion assembly (Figure 1.11). The covalent attachment of myristic acid to the N-terminal glycine requires the removal of a methionine residue from the polyprotein of entero- and rhinoviruses and the removal of the leader peptide among cardio- and aphthoviruses (Chow et al., 1987, Paul et al., 1987).

The myristoyl-P1 precursor is released from the polyprotein co-translationally by the mechanisms outlined previously in section 1.7. Virion assembly is controlled by the cleavage of the myristoyl-P1 polypeptide to myristoyl-VP0, VP1 and VP3. Myristoyl-P1 cleavage is slow during the early stages of infection as the levels of the myristoyl-P1 precursor and the viral protein 3CD^{pro} that processes the precursor are low. As the myristoyl-P1 precursor is cleaved, the structural proteins produced from the cleavage, myristoyl-VP0, VP3 and VP1 spontaneously aggregate to form a protomer (sedimentation coefficient of 5S). The aggregation of five protomer units results in the formation of a pentamer structure. The pentamer structure is often referred to as the 14S subunit. Association of twelve pentameric subunits results in the formation of capsids containing 60 copies each of myristoyl-VP0, VP3 and VP1. This process of viral assembly is illustrated in figure 1.11. Research carried out by Verlinden *et al* using a cell-free system suggests that the pentamers are the key intermediate in the encapsidation of poliovirus RNA (Verlinden et al., 2000). As a consequence, in the model shown in figure 1.9 the encapsidation of the viral RNA is shown to occur by association with the pentamer subunits during the formation of the capsid. However it has also been suggested that the viral RNA is encapsidated into preformed capsids. Whatever the exact mechanism of encapsidation of RNA it is known from studies that encapsidation is specific for viral genomic RNA and requires *de novo* synthesis of the genome (Barclay et al., 1998, Molla et al., 1991, Molla et al., 1993).

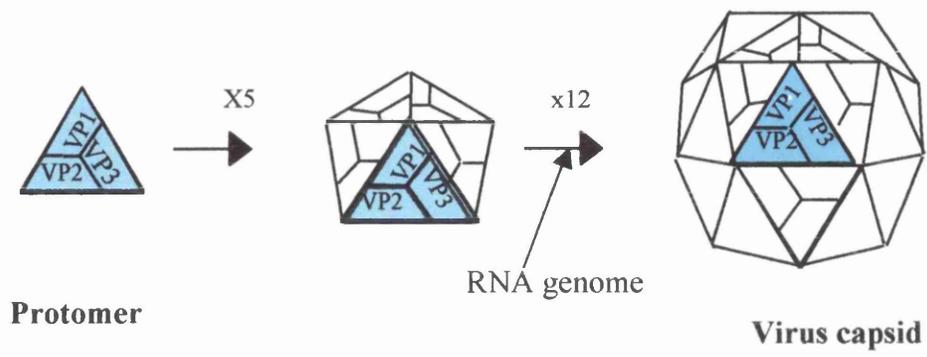


Figure 1.11 Model of particle assembly

Maturation of the virus particle is required for the virus to become infectious. Maturation requires the cleavage of myristoyl-VP0 to yield myristoyl-VP4 and VP2. The cleavage is believed to be autocatalytic and the subsequent protein rearrangement results in the formation of a more stable particle.

1.8.2 Nonstructural proteins

All non-structural proteins, have some role in RNA synthesis - only 2A is not resident in the membrane replication complexes that are the site of RNA synthesis *in vivo* (Girard & Baltimore, 1967).

1.8.3 2A

Despite a high level of conservation amongst the various genomes of the *Picornaviridae* the region encoding the viral 2A protein is the most divergent region of the genome.

In entero- and rhinoviruses 2A^{pro} is an endopeptidase that belongs to the serine protease family. 2A^{pro} is unusual among the serine protease family as the catalytic triad that makes up the protease's active site contains a cysteine residue rather than a serine. This has been confirmed by the crystal structure of HRV2 2A protease (Petersen et al., 1999) and site-directed mutagenesis studies (Yu & Lloyd, 1992). In poliovirus 2A^{pro} can cleave 3CD in an alternative pathway to generate 3C' and 3D' (Hanecak et al., 1982). The function of 3C' and 3D' proteins remains unclear, as at least in tissue culture, the 3C'/3D' cleavage site can be mutated without affecting the ability of the virus to propagate itself efficiently (Lee & Wimmer, 1988).

Aside from processing the P1/2A junction and the alternative cleavage of 3CD^{pro}, 2A^{pro} is an important determinant of cell viability. Expression of 2A^{pro} or the cardiovirus 2A in cells alone leads to cell death (referenced within Aminev et al., 2003). As a protease 2A^{pro} is responsible for the cleavage of eIF4G, a component of the translation factor eIF-4F (Krausslich et al., 1987, Lamphear et al., 1995, Lamphear et al., 1993, Ventoso et al., 1998). The initiation factor eIF-4F comprises eIF-4E, which is responsible for the recognition of the 7-methylguanosine-cap structure, eIF-4A an RNA helicase and eIF4G which functions as a bridge between the mRNA and the ribosome (Gingras et al., 1999).

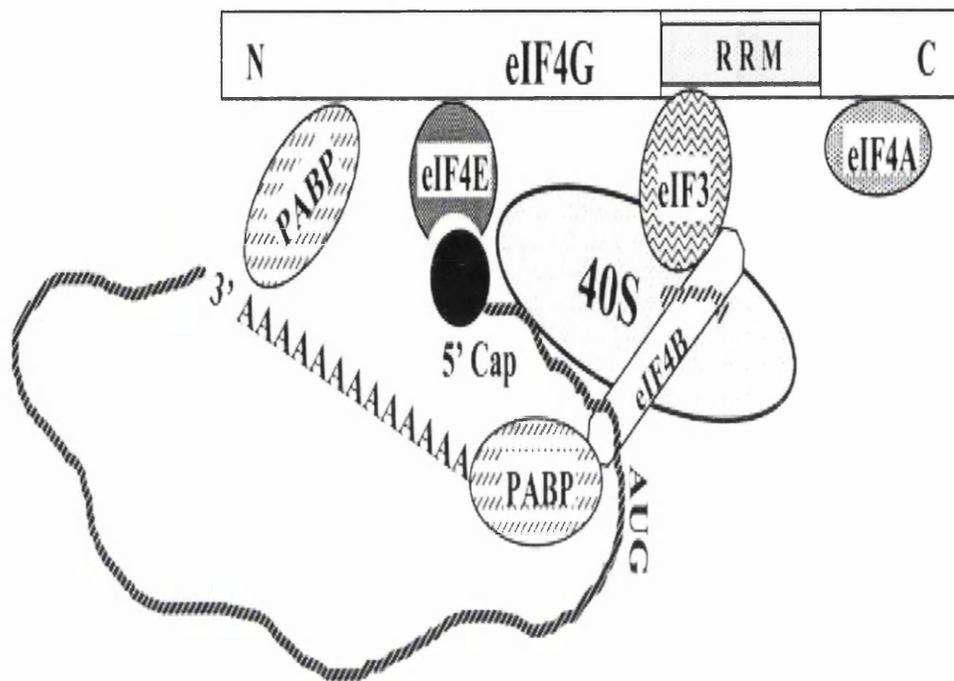


Figure 1.12 Eukaryotic closed-loop model of translation initiation.

Taken from Kean, 2003. Schematic representation of the role of eIF4G and the eIF4F holoenzyme complex in 40S ribosomal subunit recruitment on a capped polyadenylated cellular mRNA. The eIF4F complex interacts with both the capped 5' end of the RNA (black spot; via eIF4E) and the poly(A) tail (via PABP), and recruits 40S subunits via its association with eIF3. Stabilisation of this complex results from the simultaneous interaction of eIF4B with ribosomal RNA, mRNA and eIF3. For simplicity, other proteins described to interact with eIF4G have been omitted.

Two isoforms of eIF4G exist: eIF4GI and eIF4GII. Cleavage of the eIF4G isoforms prevents the recruitment of the methylguanosine-capped RNA to the ribosomes by preventing the interaction between eIF4E and eIF4G occurring (Ventoso et al., 1998). The cleavage of eIF4G leads to a significant decrease in the ability of the host to synthesize its own proteins and enables the virus to take over the cellular translational machinery for its own ends. The “shut-off” of host-cell translation by the cleavage of eIF4G by 2A^{pro}, supplemented by cleavages of other cellular proteins carried out by 2A^{pro} and 3C^{pro}, means the virus is the only possible winner in the outcome of the infection of the cell. The cleavage of eIF4G by 2A^{pro} in poliovirus-infected cells occurs rapidly post-infection. Gradi *et al* have suggested that it is the cleavage of the eIF4GII which coincides with the shut-off of host cell synthesis in picornavirus-infected cells and not cleavage of eIF4GI (Gradi et al., 1998).

Translation of the cellular mRNA may be further inhibited by the cleavage of poly-A binding protein (PABP). PABP interacts with the eIF4F complex during translation and it has been postulated that this enhances the translation of cellular mRNA through the formation of a circular structure (Figure 1.12) that allows ribosomes to reinitiate protein synthesis rapidly on the same mRNA (reviewed in Kean, 2003). PABP has been shown to be cleaved by both 2A^{pro} and 3C^{pro} *in vitro*. However if PABP cleavage contributes to the down regulation of host-cell mRNA translation it does so via a mechanism other than the disruption of the circular structure. Joachims and others have shown that the cleavage sites of both 2A^{pro} and 3C^{pro} map to the C-terminal region of PABP (Joachims et al., 1999, Kerekatte et al., 1999). The domain of PABP which is known to bind eIF4G is however located within the N-terminal region.

Expression of 2A^{pro} in COS-1 cells showed that, in addition to cap-mediated translation, cellular transcription was also inhibited (Davies et al., 1991). Inhibition of transcription by the 2A protease may be mediated by cleavage of cellular proteins. In addition to cleavage mediated by 3C^{pro} the TATA-binding protein (TBP) is targeted for cleavage by 2A^{pro} (Joachims et al., 1999, Yalamanchili et al., 1997a, Yalamanchili et al., 1996). All cellular RNA polymerases are positioned at the promoter by a transcription factor complex containing TBP. Cleavage of TBP by 2A^{pro} only removes the N-terminal residues of the

protein. It is known that this cleavage does not prevent the RNA polymerase II from initiating transcription but it may have a role in the down-regulation of transcripts from RNA polymerase I and III promoters.

Research into coxsackievirus B3 has implicated $2A^{pro}$ as a probable mechanism or contributing factor in the medical condition dilated cardiomyopathy. Badorff *et al* have demonstrated using *in vitro* and *in vivo* studies that $2A^{pro}$ cleaves two sites within the cytoskeleton protein dystrophin. The cleavage by $2A^{pro}$ of dystrophin separates the N-terminal actin-binding site from the rod domain and the C-terminal β -dystroglycan binding domain and in doing so disrupts the interaction that occurs naturally between the internal cytoskeleton and the external basement membrane (Badorff et al., 2000).

The contribution of $2A^{pro}$ to virulence has also been observed in swine vesicular disease virus (SVDV). Virulent strains of SVDV cause clinical symptoms indistinguishable from FMDV though antigenically and genetically it is more closely related to coxsackievirus B5 (Zhang et al., 1999). Sakoda *et al* have shown that the virulence of the virus maps to a single residue in $2A^{pro}$ adjacent to the catalytic triad (His21, Asp39 and Cys110). In non-pathogenic strains of the virus an isoleucine (Ile) residue is located at position 20 of $2A^{pro}$. In virulent strains this residue is an arginine (Arg). A comparative analysis of the virulent and avirulent strains of SVDV *in vitro* and *in vivo* showed that the $2A^{pro}$ containing a Ile residue at position 20 was significantly less effective at inducing cleavage of eIF4G and stimulating IRES-mediated translation *in trans* than $2A^{pro}$ containing an Arg residue at position 20. However no difference in the efficiency of cleavage of the VP1/2A junction was seen between the proteins. Sakoda *et al* engineered an Arg residue at position 20 of $2A^{pro}$ into the cDNA of an avirulent strain of the virus and reported that the phenotype associated with the virulent strains of the virus i.e. efficient cleavage of eIF4G and *trans*-activation of the viral IRES was restored (Sakoda et al., 2001).

In the cardio- and aphthoviruses genera 2A does not have proteolytic function. In aphthoviruses functions associated with the $2A^{pro}$ proteolytic cleavage of eIF4G are replaced by the Leader (L) protease. This is not the case in cardioviruses, which contain an L protein which is proteolytically inactive. Recently Aminev *et al* have shown that the

cardiovirus 2A, in contrast to the 2A^{pro} encoded by poliovirus, is localised to the nucleus. Aminev *et al* showed that deletions, which prevented nuclear localisation, resulted in a reduction in the effectiveness in which 2A shut-off host-cell protein synthesis. Aminev *et al* have proposed a model based on their observation that ribosomal RNA synthesis is up regulated in cells expressing the cardiovirus 2A and previous observations that the cardioviral 2A is found in association with ribosomes. In the proposed model the cardioviral 2A protein is incorporated into the ribosome structure during the assembly of the ribosome. When exported to the cytoplasm these modified ribosomes preferentially translate viral RNA over host-cell mRNA in a process directed by the 2A protein (Aminev *et al.*, 2003).

1.8.4 2B, 2C and 2BC

The final cleavage products of the P2 region of the genome are 2A^{pro}, 2B and 2C. The cleavage cascade also results in the formation of the stable intermediate protein 2BC. Using electron microscopy it has been determined that 2B, 2C and 2BC are all membrane-associated within cells (Bienz *et al.*, 1987, Bienz *et al.*, 1990). 2B, 2C and 2BC have all been implicated with the changes to the cellular architecture that are associated with infection of cells by picornaviruses.

Picornavirus 2B is approximately 100 amino acids in length. Two hydrophobic regions of the protein have been identified. Both hydrophobic regions (PV3 residues ³²VTSTITEKLLKNLIKIISSLVITG⁵⁵ and PV3 residues ⁶¹TTVLATLALLGCDVSPWQWL⁸¹) are believed to span the membrane bi-layer (reviewed in Nieva *et al.*, 2003). Experiments using the yeast and mammalian -two hybrid systems implied that 2B was capable of oligomerisation (Cuconati *et al.*, 1998b, de Jong *et al.*, 2002). Recent research using isolated unilamellar vesicles *in vitro* showed that oligomerisation of 2B occurred only in the presence of membranes (Nieva *et al.*, 2003). The oligomeric structure proposed for 2B (Figure 1.13) is a hairpin ‘ α -loop- α ’ motif (Agirre *et al.*, 2002, van Kuppeveld *et al.*, 1997b). Support for this structure arises from two observations. Firstly, Van Kuppeveld *et al* showed that both hydrophobic domains are required to increase membrane permeability (van Kuppeveld *et al.*, 1997a). Secondly, the structure proposed by Van Kuppeveld *et al* is structurally similar to the membrane inserting domain of the *Bacillus thuringiensis* Cry1A toxin (Masson *et al.*, 1999). In the *in vitro* system expression of 2B was shown to be able

to cause the efflux of small solutes (MW 660) at low protein to lipid ratios (Nieva *et al.*, 2003). The data provided by Nieva *et al* predicts that the functional oligomeric unit required to permeabilise the membrane is a tetramer. This is consistent with a role for 2B as a viroporin. The view that 2B is a viroporin was supported by earlier observations made by Van Kuppeveld *et al* that expression of the CB3 2B protein was sufficient to induce the influx into the cytoplasm of Ca^{2+} ions from ER-stores and extracellular sources (van Kuppeveld *et al.*, 1997a). The ability of 2B to increase the permeability of the cells was also observed using the poliovirus 2B protein in a study by Doedens and Kirkegaard (Doedens & Kirkegaard, 1995). In terms of the virus lifecycle Van Kuppeveld *et al* have shown that coxsackievirus requires a functional 2B protein to allow the efficient release of the progeny virus (van Kuppeveld *et al.*, 1997a). In this sense, the function of 2B protein appears to be similar to that of the influenza M2 protein and the togavirus 6K protein (Liljestrom *et al.*, 1991, Ruigrok *et al.*, 1991). A model for how 2B mediates the effective release of virus progeny from the cell is shown in figure 1.14.

Expression of 2B has been associated with the inhibition of protein secretion (Doedens & Kirkegaard, 1995) and the disassembly of the Golgi complex (Sandoval & Carrasco, 1997). Sandoval and Carrasco propose that 2B, through an interaction with the Golgi membrane, promotes fusion of the Golgi with the ER. This fusion is responsible for the swelling of the ER observed in poliovirus cells, and results in the accumulation of glycoproteins in the ER (Sandoval & Carrasco, 1997). Previous mutagenic studies investigating the function of 2B have shown that in addition to these functions 2B may also function in virus replication (van Kuppeveld *et al.*, 1996a, van Kuppeveld *et al.*, 1996b, Vankuppeveld *et al.*, 1995).

The ability of mutations in 2B to abolish RNA synthesis may relate to the function of the intermediate proteins: 2BC or 2BCP3. Expression of 2BC alone has been shown to result in cellular changes similar to those produced in virus-infected cells (Aldabe & Carrasco, 1995, Cho *et al.*, 1994, Suhy *et al.*, 2000). However it has been demonstrated that vesicles produced in this manner are unable to support viral RNA synthesis (Egger *et al.*, 2000, Jurgens & Flanagan, 2003). This will be discussed in more detail in section 1.10.1. The ability of 2BC to induce vesicle formation has been demonstrated to occur independently of the 2C NTPase activity (Cho *et al.*, 1994). Instead it has been proposed that 2BC

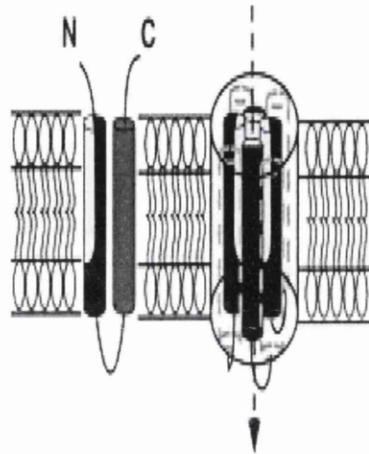


Figure 1.13 Schematic models of PV-1 2B membrane structures.

Taken from Nieva et al., 2003. In the integral hairpin, the amphipathic 32–55 helix is shown with hydrophilic and hydrophobic sides in light gray and black, respectively, while the hydrophobic 61–81 helix is dark gray.

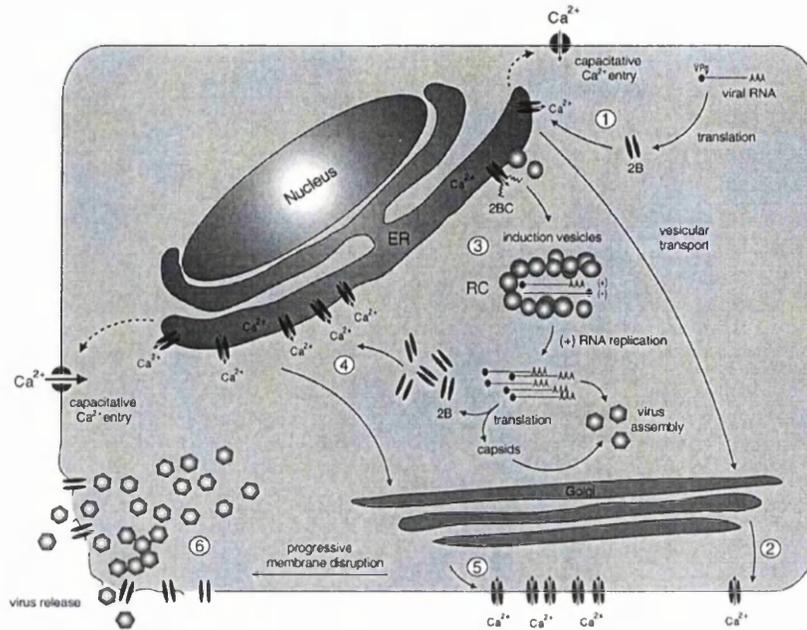


Figure 1.14. Model for the mobilization of ER-stored Ca^{2+} , the influx of extracellular Ca^{2+} and the release of virus progeny by protein 2B through the formation of membrane-embedded pores. Taken from Van Kuppeveld et al., 1997. Early events (up to 4 h p.i.) in the viral replicative cycle involve: (1) production of protein 2B and insertion in the ER membrane, leading to a release of Ca^{2+} and, as a consequence, opening of capacitative Ca^{2+} entry channels; (2) transport of 2B to the plasma membrane where it is inserted and causes influx of extracellular Ca^{2+} ; (3) induction of the membranous vesicles by protein 2BC that are required for viral plus-strand RNA replication (possibly through the release of ER-stored Ca^{2+}). Late events (from 5 h p.i.) in the viral replicative cycle involve: (4) translation of newly formed plus-strand RNAs yielding large numbers of protein 2B; insertion of these proteins in the ER membrane causes a rapid evacuation of Ca^{2+} and gives rise to an increased capacitative Ca^{2+} entry; (5) collapse of Ca^{2+} gradient maintained by the plasma membrane by massive insertion of 2B proteins; (6) progressive increase in the size of the pores formed by 2B causes disruption of the membrane and results in the membrane lesions that allow virus release.

induces vesicles by promoting the efflux of Ca^{2+} ions from the ER Lumen (van Kuppeveld et al., 1997a). The ER membrane is stabilised by Ca^{2+} binding protein- Ca^{2+} ion-phospholipid interactions. Efflux of Ca^{2+} ions from the ER leads to a reduction in membrane stability. Eventually the membrane will become significantly destabilised and collapse, resulting in the formation of vesicles. Vesicularisation, occurring in this manner, is observed in cells treated with calcium ionophores (Sambrook, 1990). However, the ability of 2BC to induce vesicle formation must be enhanced by functions contributed by the 2C domain as expression of 2B alone does not result in the same cellular rearrangements.

The 329 amino acid 2C protein is highly conserved amongst all picornaviruses but shares little homology with cellular or viral proteins. The similarities that 2C does share with known proteins are two NTPase binding motifs (motif A and B) and a series of residues that have homology to the helicases of SV40 T antigen and papillomavirus E1 (motif C) (Gorbalenya et al., 1990). Both the NTPase binding motifs and the helicase motif are located in the central region of the protein (Figure 1.15). The first NTP binding motif found in 2C contains the conserved A/GxxxxGKS/T (motif A) that has been shown to bind the phosphate groups of NTPs, while the second NTP binding motif found in 2C shares homology with NTPase binding motif B (DD/E). Motif C is found downstream of motif B and though it is found in viral helicases the function of this motif is currently unknown. Consistent with the proposed role for 2C as a viral helicase introduction of mutations that abrogated the ability of 2C to hydrolyse NTP abolished virus viability (Mirzayan & Wimmer, 1994a, Teterina et al., 1992). The requirement for NTPase activity for viral replication to occur has been confirmed by investigations into the action of guanidine hydrochloride, a reversible inhibitor of poliovirus replication, which has been demonstrated to inhibit the ATPase activity of the 2C protein (Pfister & Wimmer, 1999). However although the ability of 2C to hydrolyse ATP and GTP can be demonstrated it has yet to be proved that 2C has helicase activity (Mirzayan & Wimmer, 1994b, Rodriguez & Carrasco, 1993).

Expression of 2C, independent of the other non-structural proteins, is sufficient to cause membrane rearrangements. The ability of 2C to rearrange the cellular membranes resides in the first 88 amino acids of the protein (Teterina et al., 1997). Initial binding studies have identified an amphipathic helix present at the N-terminus of the 2C protein (amino acids 18-35) that is important for the interaction between 2C and membranes (Paul et al., 1994b).

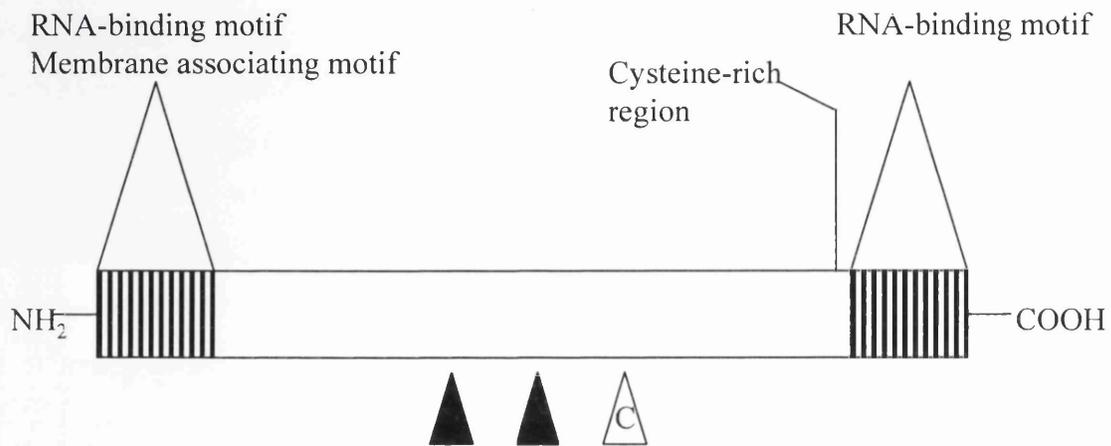


Figure 1.15 Functional domains of 2C

Amphipathic domains of the protein are illustrated with thin black vertical lines. The black triangles indicate the regions of 2C homologous to motifs A and B of the NTP binding motif. The white triangle corresponds to the region of 2C homologous to motif C of the type III helicases. Domains associated with the RNA-binding and membrane-associating characteristics of 2C have been labelled.

A subsequent study has implied that the C-terminal residues (amino acids 252-329) may also be involved in the interaction of 2C with the membrane (Teterina et al., 1997). In addition to the membrane binding domains, the 2C protein also contains two RNA-binding domains. These RNA-binding domains are located at the amino and carboxy-termini of the 2C protein (Rodriguez & Carrasco, 1995). It has since been reported that both 2C and 2BC specifically bind to stemloop b of the cloverleaf structure found at the 3' termini of the negative-sense replication intermediate in both poliovirus (Banerjee et al., 2001) and HAV (Banerjee & Dasgupta, 2001). The importance of this observation with respect to the mechanism of RNA synthesis is not yet understood. The study of temperature sensitive mutants of 2C, and studies using replication inhibitors have also suggested that 2C has a role in uncoating (Li & Baltimore, 1990) and encapsidation (Vance et al., 1997) of the viral RNA.

A variety of interactions between the proteins of the P2 region (2C, 2B and 2BC) have been detected using the yeast-two hybrid (Cuconati et al., 1998b), mammalian two-hybrid systems (de Jong et al., 2002) and GST-pull down assays (Cuconati et al., 1998b) (Figure 1.16). The contribution of these interactions to the cytopathic effects observed in picornavirus-infected cells has not yet been quantified.

1.8.5 3A, 3B and 3AB

Like the proteins of the P2 region, the non-structural proteins 3A and 3AB are found associated with cellular membranes in infected cells (and Datta & Dasgupta, 1994, Semler et al., 1982). A stretch of 22 hydrophobic amino acids within 3A mediates the interaction of both viral proteins with the cellular membrane (Towner et al., 1996). 3A, in common with 2BC and 2B, has been shown to alter the membrane permeability of the infected cell (Lama & Carrasco, 1992). Furthermore, it has been demonstrated that expression of 3A results in the down-regulation of cellular immune responses. Expression of poliovirus protein 3A has been demonstrated to reduce the amount of interleukin-6, interleukin-8 and beta-interferon (Dodd et al., 2001) secreted from the host cell, during the course of the viral infection (Doedens et al., 1997, Doedens & Kirkegaard, 1995). The latter occurs due to inhibition of the ER to Golgi transport. A further consequence of the down regulation of ER-Golgi transport is the reduction of MHC class I (Deitz et al., 2000) and the tumour necrosis factor (TNF) receptor (Neznanov et al., 2001) on the surface of infected cells. It is known that the ability of 3A to down-regulate the ER-Golgi transport resides in the

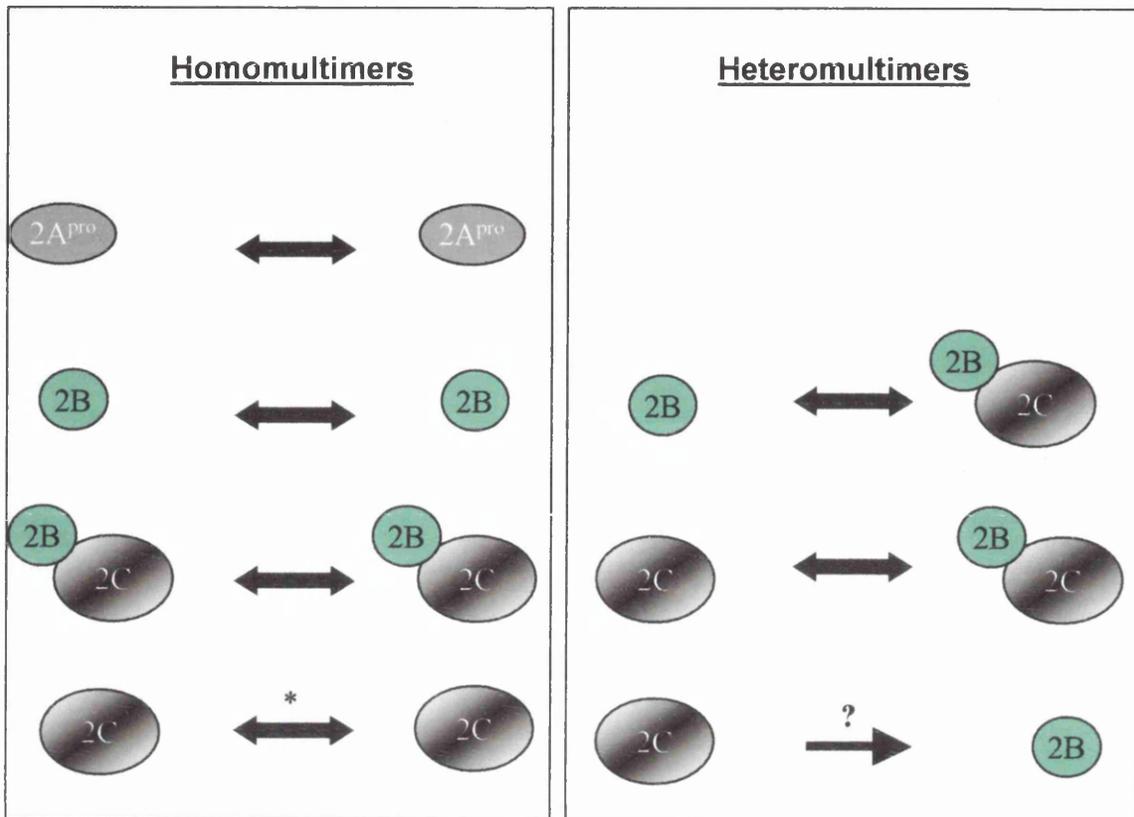


Figure 1.16. Map of P2 interactions

Double arrows indicate positive reciprocal interactions, whereas single arrows denote unidirectional interactions. Interactions detected only with a GST pull-down (Cuconati et al., 1998) and not in yeast (Cuconati et al., 1998 or mammalian two-hybrid systems (de Jong *et al.*, 2002) are marked with a *. Interactions only detectable in a yeast two-hybrid system (Cuconati et al., 1998) are marked with a ?.

N-terminal of the protein (Neznanov et al., 2001), however it remains unclear whether the increase in membrane permeability and the down-regulation of the immune system are linked or if these reflect two distinct functions of 3A.

Recently evidence has been presented that a function of 3A is required for the production of genomic sense strands, though this same function is not required for negative-sense strand synthesis. Although it is not yet understood what function of 3A is required for positive strand synthesis, the research showed that 3A was not required for the templated addition of uridylylate residues to VPg, a process known to be specific for genomic RNA synthesis (Teterina et al., 2003).

3B is VPg (Virion Protein, genome) which is found covalently attached to the 5' terminal of the genome and anti-genome. In its uridylylated form (VPg-pUpU or VPg-poly U) VPg functions as a primer for the initiation of negative and positive strand synthesis. In FMDV research has shown that viral proteins 3A and 3B have a role in determining host range (Pacheco et al., 2003).

The precursor to 3A and 3B, the intermediate protein 3AB, has a number of functions independent of those of the final cleavage products. Biochemical data has shown that 3AB, by interaction with 3CD^{pro} and 3D^{pol}, provides functional stimulation to these proteins. Interaction of 3AB to 3CD^{pro} stimulates both the binding specificity of 3CD^{pro} for the cloverleaf (in presence of 3AB interaction occurs even in the presence of 1000-fold excess of tRNA) (Harris et al., 1994) [Xiang, 1995 #9483] and the self-cleavage of 3CD^{pro} to the final cleavage products 3C^{pro} and 3D^{pol} (Molla et al., 1994, Xiang et al., 1995b). In addition, interaction of 3AB with the viral polymerase has been demonstrated to stimulate the polymerisation activity of the polymerase at low concentrations (Lama et al., 1994, Paul et al., 1994a, Plotch & Palant, 1995).

1.8.6 3C^{pro}, 3CD^{pro} and 3D^{pol}

3C^{pro} is the only protease to be conserved within all members of the picornavirus family and is responsible for the majority of the cleavages in the poliovirus polyprotein. The picornavirus 3C proteases fall into two groups with the rhino- and enteroviruses forming a closely related group containing the catalytic triad Cys-147, His-40 and Glu-71. The catalytic mechanism of the second group, which includes the aphtho-, cardio- and hepatoviruses remains unresolved.

The cleavages mediated by 3C^{pro} occur exclusively at Gln-Gly (Q-G) dipeptides. However not every Q-G residue in the viral polyprotein is cleaved by 3C^{pro} suggesting that other factors play a role in determining substrate specificity (Long et al., 1989, Nicklin et al., 1988). In addition to its role in processing the viral polyprotein 3C^{pro} has also been implicated in the cleavage of a number of cellular proteins. The cleavage, by 3C^{pro}, of the cellular transcription factors TFIIC, Oct I, TBP and CREB results in the down regulation of transcription from pol II and pol III promoters (Clark et al., 1993, Yalamanchili et al., 1997a, Yalamanchili et al., 1996, Yalamanchili et al., 1997b). The inhibition of transcription from these promoters is apparent 4 and 5 hours post-poliovirus infection respectively. In addition to direct cleavage of the transcription factors TFIIC, Oct I, TBP and CREB, evidence has also been obtained for the degradation of the transcription factor and tumour suppressor p53 by 3C^{pro} (Weidman et al., 2001). The degradation of p53 is not mediated directly by 3C^{pro} but has instead been shown to require an unidentified cellular factor, a function which is inhibited by vaccinia virus infection (Weidman et al., 2001). Weidman *et al* postulate that the ability of 3C^{pro} to mediate the degradation of p53 is advantageous to the virus as it prevents the activation of the apoptotic cellular response therefore ensuring the successful assembly and release of progeny virus from the infected cell.

As well as having a role in down-regulation of transcription, 3C^{pro} has been shown to cleave PABP, PTB and La auto-antigen. PABP is known to stimulate the translation of poly-A tailed RNA. Research suggests 3C^{pro} cleavage of La auto-antigen may result in its redistribution to cytoplasm from the nucleus where it is known to be involved in stimulating the IRES-mediated translation of the viral genome (Shiroki et al., 1999). Cleavage of PTB at multiple cleavage sites (Ala-X-X-Gln/Ala) by 3C^{pro} results in the redistribution of the cleaved fragment to the cytoplasm. A recent report has shown that the

cleaved fragments of PTB can suppress the translation of the viral genome. Back *et al* have proposed that this contributes to the “molecular switch” in the use of the viral genome from translation to replication (Back et al., 2002).

3CD^{pro} is a relatively stable intermediate protein responsible for cleavages in the viral polyprotein. Binding of 3CD^{pro} to the cloverleaf increases the affinity of PCBP from the IRES to the cloverleaf. Given that replication and translation cannot occur at the same time Gamarnik and Andino have proposed a model in which this event acts as a molecular switch moving the role of the genomic strand from translation to replication (Gamarnik & Andino, 1998). In entero- and rhinoviruses 3CD^{pro} is also responsible for the processing of the P1 precursor (Jore et al., 1988, Ypma Wong et al., 1988).

3D^{pol} is the 52 kDa RNA-dependent RNA polymerase (RdRp) encoded by the picornaviruses. 3D^{pol} is discussed in more detail in section 4.1.1 to section 4.1.1.4. Briefly, 3D^{pol} is responsible for the covalent addition of two uridylate residues to the viral protein VPg. The addition of two uridylate residues, termed uridylylation, of VPg to form VPg-pUpU enables VPg to function as a primer for the initiation of negative- and positive strand synthesis. 3D^{pol} is also responsible for the template-dependent elongation of VPg-pUpU to synthesise full-length positive and negative-sense strands.

1.9 Genomic replication

All known RNA viruses encode RNA-dependent RNA polymerases (RdRp) that mediate genomic replication. In RNA viruses with a positive-sense single-stranded RNA genome replication is initiated only after the synthesis of the viral replication proteins.

Infection of a cell by all characterised RNA viruses is associated with intracellular alterations, in particular the formation of membrane vesicles and/or membrane rearrangements. The replication of all characterised positive-stranded RNA viruses occurs on intracellular membrane structures, termed the replication complex. These replication complexes are derived from specific intracellular membranes, for example alphaviruses use endosomal and lysosomal membranes (Kaarianinen & Ahola, 2002) and nodaviruses use mitochondrial membranes (Miller et al., 2001).

The use of membranous replication complexes is believed to provide RNA viruses with three important functions. Firstly, the formation of a membrane structure localises and concentrates the proteins required for replication and assembly. Secondly, the membrane complex has been shown to protect the RNA genome within from degradation (Ahlquist, 2002). Lastly, it has been suggested that the replication complex hides the presence of double-stranded (ds) RNA and by doing so prevents or reduces the level of activation of dsRNA-induced host immune responses, like the double-stranded RNA induced protein kinase R (PKR) response and RNA silencing (Ahlquist, 2002).

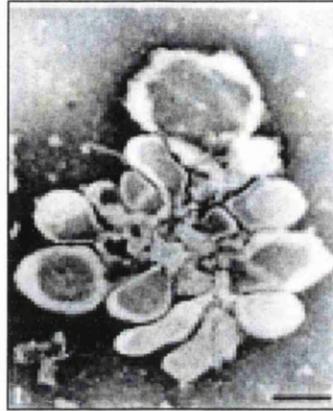


Figure 1.17 Electron micrograph of an isolated replication complex, called a rosette. Bar, 0.1 μm (Bienz et al., 1990)

1.9.2 *Cis*-acting replication elements

In vitro, the poliovirus polymerase is able to copy artificial templates. However, *in vivo* only the virus genome is replicated. It is therefore clear that the viral RNA possesses signals that allow the polymerase to differentiate the viral RNA from cellular RNAs. A number of *cis*-acting RNA elements have been identified in the genome of the picornaviruses, both internally and at the genomic termini.

1.9.2.1 Cloverleaf

The first 100 nucleotides of the 5'UTR forms a cloverleaf structure (Figure 1.18). A very similar structure has also been identified HRV 2 and HRV14 with sequence conservation of 70% (Xiang et al., 1995b). Protein binding studies have demonstrated that the 5' cloverleaf binds 3CD^{pro}, the uncleaved precursor of 3C^{pro} and the viral polymerase 3D^{pro}, to stem loop d and a cellular protein PCBP binds to stem loop b to form a ternary ribonucleoprotein complex. Experimental evidence has shown that mutations affecting the structure of the 5' cloverleaf or the binding domain of 3CD^{pro}, so disrupting the formation of the ternary ribonucleoprotein complex, result in replication deficient phenotypes (Andino et al., 1990b, Andino et al., 1990c). It is now postulated that the 5' cloverleaf provides the "switch" mechanism between translation and replication by altering the preferential binding site of PCBP 1 and 2 from domain IV to stemloop b of the cloverleaf. However, it has been estimated that the cellular concentrations of PCBP are sufficiently high to allow binding to both sites simultaneously. Although this can be explained by the localisation of picornavirus replication to membranous structures, it remains to be defined whether this is the sole mechanism of a "switch" to replication. However, some form of switch must occur as, the directional movement of the ribosomes and polymerase along the template strand is incompatible with simultaneous replication and translation (Barton et al., 1999).

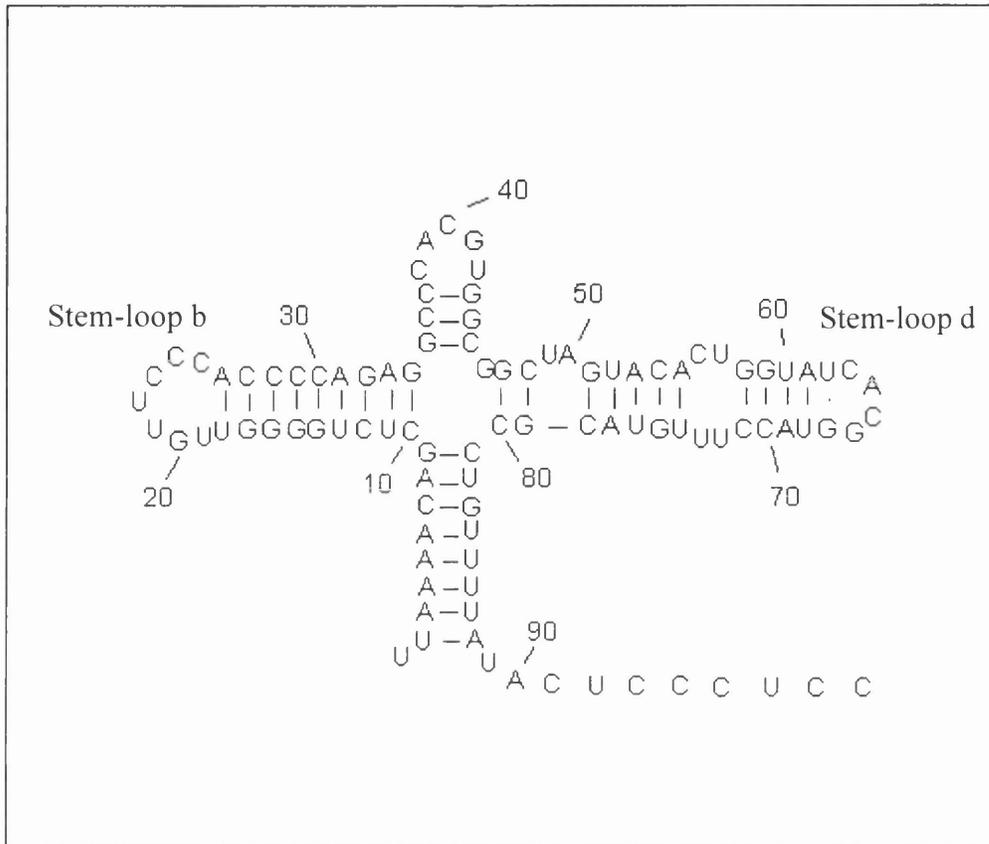


Figure 1.18 Structure of the poliovirus type 3 Cloverleaf. Adapted from <http://www.med.uni-jena.de/.../picorna/pv3sab.html>. Andino *et al* have demonstrated that a functional ribonucleoprotein complex formed around this structure, first proposed by Rivera *et al* (Rivera *et al.*, 1988), is required for the initiation of RNA synthesis (Andino *et al.*, 1990, 1993).

1.9.2.2 CRE

A *cis*-acting replication element in the ORF of the picornavirus polyprotein was initially identified in HRV14 (McKnight & Lemon, 1998). Subsequently similar *cis*-acting replication elements have been identified in cardiovirus (Lobert et al., 1999) and poliovirus (Goodfellow et al., 2000a). The CRE of poliovirus resides in the coding region of 2C. In contrast the CRE of HRV14 and cardiovirus are located in the region of the polyprotein coding for the capsid proteins. The predicted structure of these *cis*-acting replication elements is shown in figure 1.19. Goodfellow *et al* have recently shown using RNAase structural mapping that the terminal loop of the PV CRE (Figure 1.19A) is larger and so is structurally similar to that of HRV14 (shown here in Figure 1.19B) (Goodfellow et al., 2003a).

Despite the differences in positioning within the ORF, the Human rhinovirus and enterovirus *cis*-acting replication elements contain a conserved GXXXAAAXXXXXXA motif (Yang et al., 2002). Mutational analysis has shown that the first two A residues of the “AAA” sequence are critical for virus viability (Paul et al., 2000, Rieder et al., 2000). Paul *et al* have demonstrated using an *in vitro* system that these A residues template the addition of uridine residues to VPg. It has since demonstrated that this occurs by a “slideback” mechanism (Paul et al., 2003). Figure 1.20 outlines the principle by which this templated addition occurs. Viral protein 3CD^{pro} binds to the CRE and by doing so increases the affinity for the CRE by the VPg-3D^{pol} complex (step 1). A uridylylate residue (U) is covalently linked to the hydroxyl group of VPg using the first A residue as a template. The VPg-pU complex slides back and hydrogen bonds to the second A nucleotide (step 2). A second U residue, once again templated by the first A residue, is linked to the first U residue to form VPg-pUpU. Although VPg-pUpU is the primer used to initiate both positive and negative-sense RNA strands it has been demonstrated that VPg-pUpU synthesised from the CRE-templated addition of uridylylate residues to VPg is only used in the initiation of positive strand synthesis (Goodfellow et al., 2003b, Morasco et al., 2003, Murray & Barton, 2003).

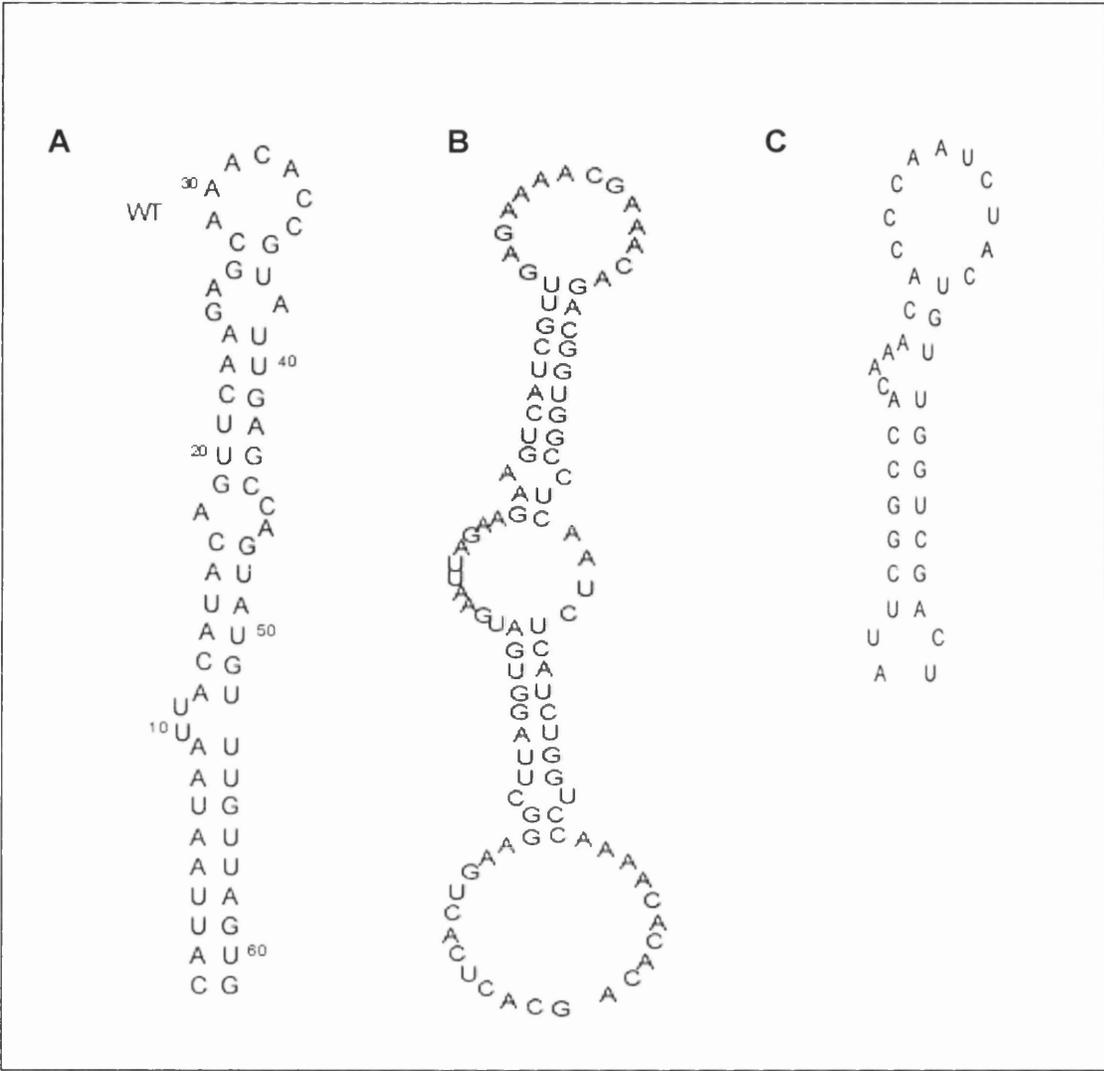


Figure 1.19 Structure of CRE in different picornavirus genera
(A) PV3 CRE found at nt 4435 - 4495 of the virus genome as described by Goodfellow *et al* (Goodfellow *et al.*, 2000) **(B)** CRE structure of HRV14 found at nt 2318-2413 as described by McKnight and Lemon (McKnight and Lemon, 1998). **(C)** Structure of the Theilers virus CRE found at nt 1568-1603 of the virus genome Lobert *et al* (Lobert *et al.*, 1999)

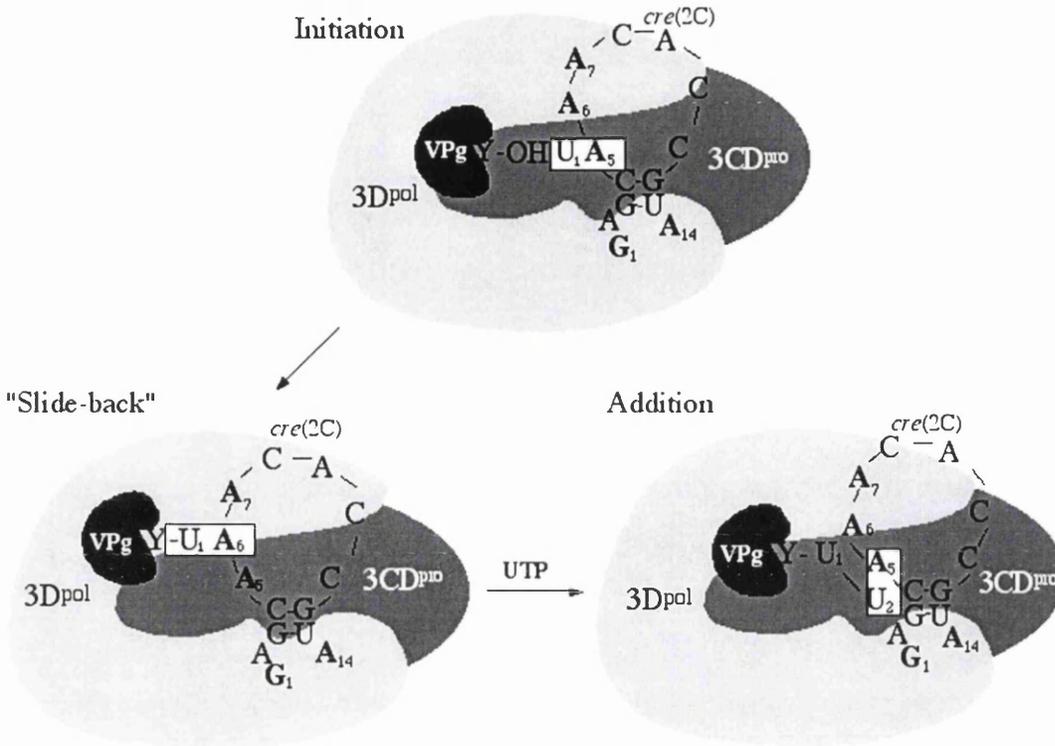


Figure 1.20 The proposed slide back mechanism for the uridylylation of VPg. Viral protein 3CD^{pro} binds to the PV *cre* and enhances the binding of the 3D^{pol}·VPg complex to the loop. During the initiation step U is linked to the hydroxyl group of tyrosine in VPg using A-5 as the template nucleotide. VPgpU then slides back to hydrogen bond with A-6 followed by the addition of a second U again on the A-5 template nucleotide. Nucleotides of the conserved motif are shown in *bold*, and nucleotides involved in the slide back are *boxed*. The figure was taken from Paul et al., 2003.

1.9.2.3 3'UTR

In comparison with the 5'UTR of poliovirus the 3'UTR has been comparatively little studied. As a consequence very little is known about the precise role of the 3'UTR in picornavirus replication. Experiments carried out using viruses that lacked the complete 3'UTR or contained a mutated 3'UTR have shown that these viruses were viable, albeit only weakly (Meredith et al., 1999, Todd et al., 1997). It would therefore appear that there is no obligatory requirement for the presence of the 3'UTR in tissue culture for the replication of picornaviruses. However, the efficiency with which replication occurs depends on its presence.

The 3'UTR of all enteroviruses appears to fold into the same core structure consisting of either two or three stem loop domains in which part of the poly A tract is included (Pilipenko et al., 1992b) (Figure 1.21). Supporting evidence for this structure has been obtained by three independent studies using both thermodynamic (Jacobson & Zuker, 1993, Melchers et al., 1997) and phylogenetic approaches (Pilipenko et al., 1992b). In contrast to the enteroviruses, the 3'UTR of rhinoviruses consists of a solitary stem-loop structure. Despite a lack of sequence and structural conservation Rohll *et al* have demonstrated that the 3'UTR of poliovirus type 3 is functionally interchangeable with that of HRV14 (Rohll et al., 1995).

Two alternative structures have been proposed for the tertiary structures of the enteroviruses. In the first model a classical pseudoknot is formed in which bases of domain Y hydrogen bond with nucleotides located in the region of the genome encoding 3D^{pol}. Alternatively, it has been proposed that the tertiary structure formed is the result of a “kissing interaction” between the nucleotides of domain X and domain Y. Although experimental evidence is available that supports the existence of a pseudoknot (Jacobson & Zuker, 1993) the formation of this structure is at odds with other evidence. Mutational analysis suggests that both site directed mutations (Melchers et al., 1997) and insertions, of up to 1000 nucleotides (Pierangeli et al., 1995, Rohll et al., 1995), used in such a way as to effectively prevent such a structure forming had no effect on the replication of the virus. Further support for a “kissing interaction” at the 3'UTR has been provided by research using poliovirus and coxsackie B virus. In these studies, viruses containing mutations designed to disrupt the “kissing interaction” were demonstrated to be replication deficient (Mirmomeni et al., 1997, Pilipenko et al., 1996).

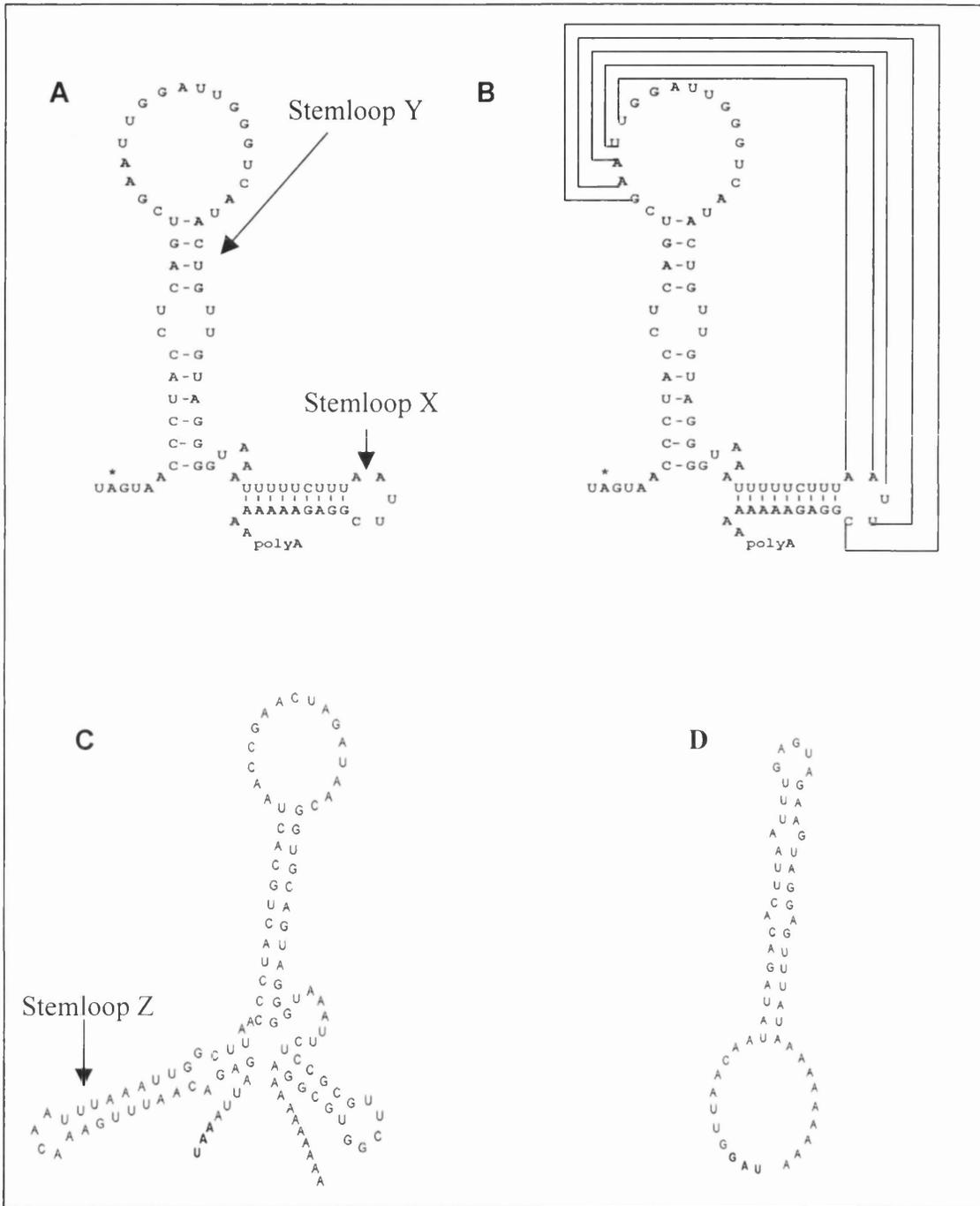


Figure 1.21. Picornavirus 3'UTR's.

(A) is the poliovirus 3'UTR (B) shows the "kissing interaction" proposed to occur between domain X and Y of the enterovirus 3'UTR (C) is the 3'UTR of coxsackievirus B4 (D) is the 3'UTR of HRV14.

Furthermore, Mirmomeni *et al* demonstrated that all isolated phenotypic pseudorevertants contained mutations that restored the “kissing interaction”. Thus, the bulk of accumulated evidence appears to favour the formation of a tertiary structure through a “kissing interaction” rather than pseudoknot formation (Mirmomeni *et al.*, 1997).

Studies using poliovirus showed that chimeric poliovirus containing the 3'UTR of either HAV or BEV replicated to significantly lower levels than those containing the 3'UTR of either HRV14 or coxsackievirus B4 (Rohll *et al.*, 1995). This suggests perhaps that modulation of the replication process may be due to interaction with cellular factors. A number of cellular factors have been found to interact with the 3'UTR. These include p105, p68 and p45 for poliovirus (Waggoner & Sarnow, 1998), p34-p36 for rhinoviruses (Roehl & Semler, 1995) and p38, p45, p47, p84 and p110 for HAV (Kusov *et al.*, 1996). Of all these proteins only p105 has been purified. Partial sequencing of this protein has indicated that it is identical to nucleolin, a protein normally found in the nucleolus of HeLa cells (Waggoner & Sarnow, 1998). The biological significance of this interaction and that of the other cellular proteins is not fully understood. However use of a chimeric poliovirus replicon, containing a mengovirus IRES which allows efficient translation in the absence of HeLa cell proteins, demonstrated a HeLa cell factor is required for RNA synthesis to proceed (Gamarnik & Andino, 1996).

1.9.2.4 Poly A tail

In addition to the 3'UTR picornaviruses encode a poly A tract at the 3'end. Although the length of poly A tail is conserved within genera, it is different between genera. Analysis of poly U tails at the 5'end of the negative-sense strand of replicative intermediates shows that they are of similar length and heterogeneity as the poly A tail of the genomic RNA. This indicates that the poly A tract is genetically encoded. Although the exact function of the poly A tail is not understood it has been demonstrated in poliovirus that deletion of the poly A tail leads to a loss of infectivity (Rohll et al., 1995, Sarnow, 1989). This has also been demonstrated in HAV (Kusov & GaussMuller, 1997) and encephalomyocarditis virus (Cui & Porter, 1995) as well as in other positive strand viruses, such as, bamboo mosaic potexvirus (Tsai et al., 1999). While in HAV the poly A tail has been demonstrated to be part of a recognition element for the viral replicase (Kusov & GaussMuller, 1997) additional functions of the poly A tail, other than binding the protein primer, have yet to be shown in poliovirus.

1.10 Models of poliovirus RNA replication

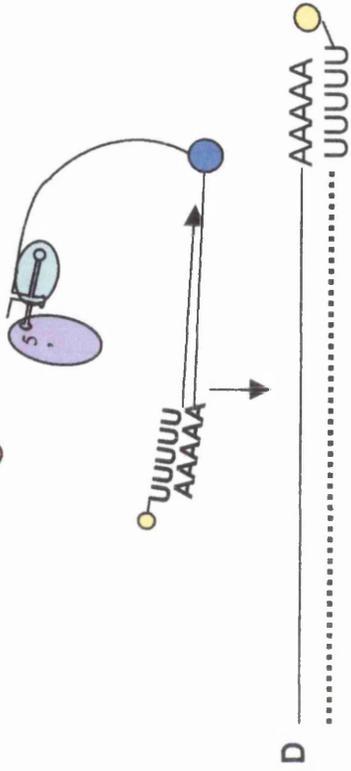
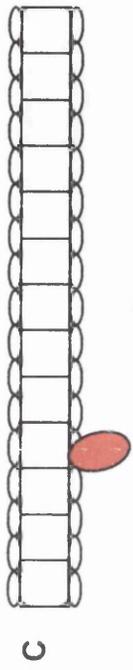
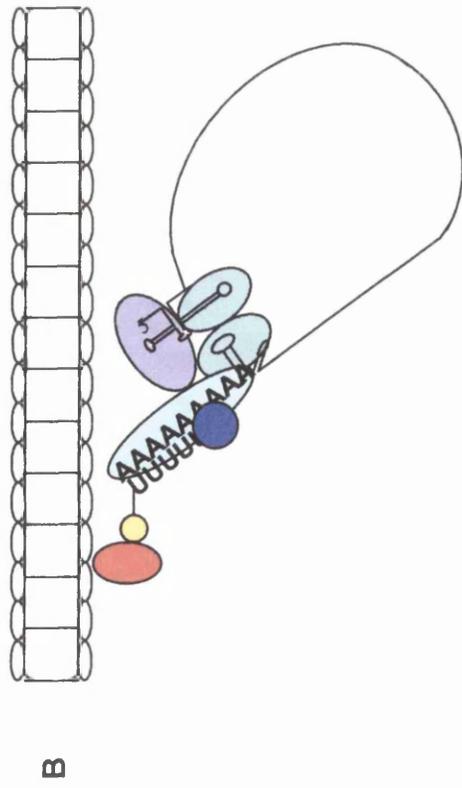
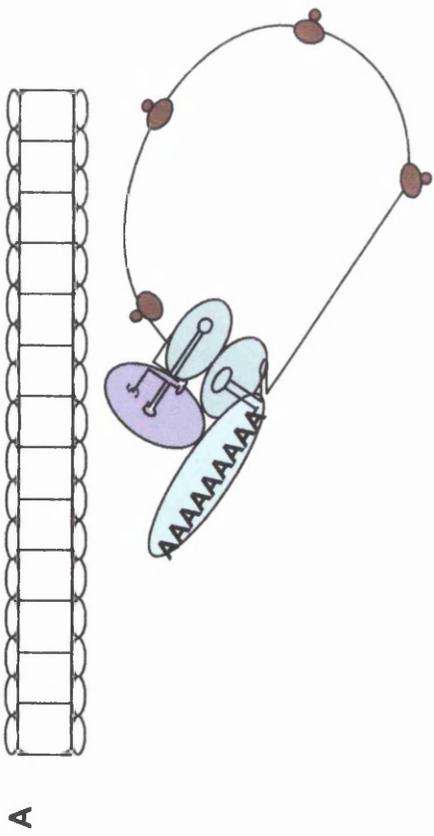
1.10.1 Model of minus strand synthesis

The first step in the replication of the poliovirus genomic RNA is the formation of a negative sense copy of the viral genome. The initiation of the negative sense copy, like the initiation of positive-strand synthesis, requires the presence within the cell of the uridylylated version of VPg. Recent literature has suggested that the VPg-pUpU that is synthesised from the CRE is not used for the initiation of negative-sense RNA strands (Goodfellow *et al.*, 2003b, Morasco *et al.*, 2003, Murray & Barton, 2003). Although no definitive proof exists as yet, it has been suggested by more than one research group that the VPg-poly (U) could be used as a primer for initiation of negative strands.

The proposed model for the formation of the negative strand replication intermediate is illustrated in figure 1.22. The viral protein 3AB is shown inserted into the endoplasmic reticulum (ER). In accordance with the data obtained by Andino *et al* (Herold & Andino, 2001) the synthesis of the negative-sense strand is shown to occur following circularisation of the viral genome. Circularisation of the viral genome is believed to occur via a protein-protein bridge formed between two molecules of 3CD^{pro}. Following circularisation of the genomic RNA the poly (A) tail is in close proximity to VPg (3B). In the current model for the synthesis of the negative-sense RNA strand VPg-poly (U) is required for initiation. This is templated by the poly (A) tail in a reaction catalysed by the RdRp 3D^{pol}. VPg-poly (U) is then elongated by the viral polymerase until a full-length negative-sense strand has been formed. The negative-sense strand is only found in virus-infected cells in the form of a partially double-stranded or fully double stranded molecule. The fully double-stranded molecule is often described as the replication intermediate or RI genomic form.

Figure 1.22 Model of negative strand synthesis (Not to scale)

3CD^{pro} (green oval) and the cellular protein PCBP (purple oval) interact with each other and with the 5' cloverleaf structure to form a circular RNP complex. The binding of the proteins and the circularisation of the genome leads to the inhibition of translation. The inhibition of translation initiation enables the virus to switch to replication. Translation continues only as long as ribosomes actively engaged in protein synthesis are maintained on the genome (A). Once the template RNA is cleared initiation of RNA synthesis can occur. In the model 3AB, VPg (yellow sphere), 3A (red oval), is shown embedded in the ER membrane. The addition of uridylate residues to VPg is templated for by the poly A tail in a reaction catalysed by the viral polymerase 3D^{pol} (Blue sphere) (B). Negative strand synthesis initiates through the elongation of the VPg-poly U primer by 3D^{pol}. The synthesis of negative-sense RNA disrupts the circular structure of the RNP (C).



Key:

3D ^{pol}	3CD ^{pro}	3D ^{pro}	3CD ^{pro}
3AB	PCBP	3AB	PCBP
3A	Membrane	3A	Membrane
-pUpU	Negative-Sense Strand	VPg-pUpU	Negative-Sense Strand
	Positive-Sense strand	PABP	Positive-Sense strand

1.10.2 Model of plus-strand synthesis

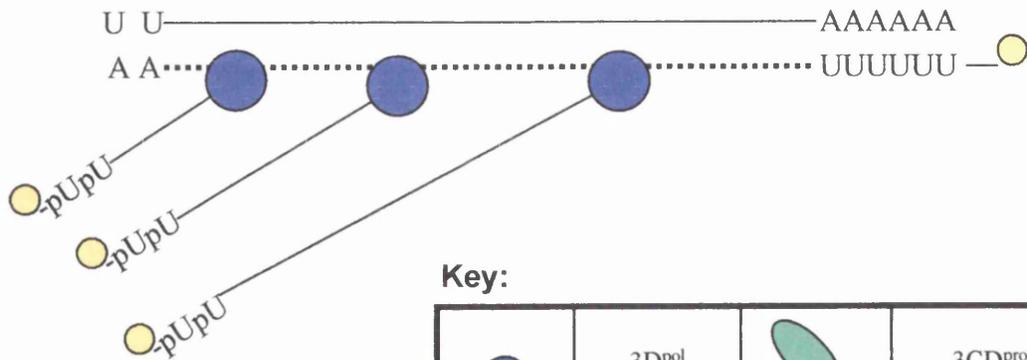
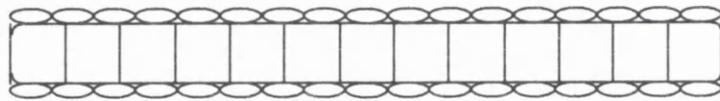
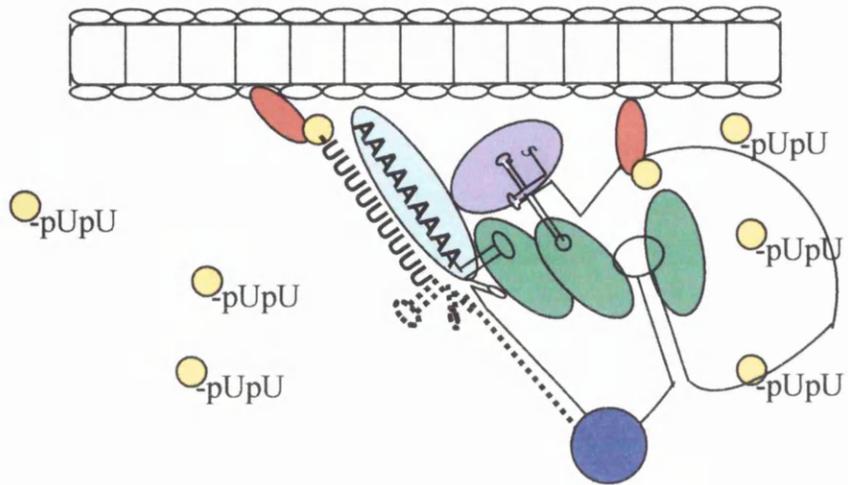
The current model of positive-sense strand synthesis in poliovirus is shown in figure 1.23. In the model proposed by Goodfellow *et al* and others, the CRE, throughout the duration of negative-sense strand synthesis, templates the formation of a pool of VPg-pUpU (Goodfellow *et al.*, 2003b). Following synthesis of VPg-pUpU on the CRE the protein is released from the template. This release is believed to occur due to a structural motif in the polymerase that prevents the addition of further U residues to the VPg (Paul *et al.*, 2003).

Following termination of the negative-sense strand the viral RdRp can use the VPg-pUpU complexed with the 3' termini of the negative-sense strand as a primer to initiate the synthesis of positive-sense strand. The mechanism by which the 3D^{pol}-VPg-pUpU complex is translocated to the 3' end of the negative-sense strand is unknown. The synthesis of genomic-sense strands within a virus-infected cell occurs with greater efficiency than that of the negative-sense strands. Two features of the current model of positive-strand synthesis can be used to explain this bias. Firstly, in the replication complex the negative-sense strand is only ever found in a fully double-stranded (termed RF) or partially double-stranded (known as RI) complex. As a result of this, the only residues available for the CRE-templated VPg-pUpU primer to interact with are found at the 3' termini of the negative-sense strand. Secondly, *in vitro* it has been demonstrated that the formation of VPg-pUpU by 3D^{pol} occurs relatively inefficiently when uridylylation is templated by poly A (Paul *et al.*, 2003). In contrast, CRE-templated addition of U residues by 3D^{pol} *in vitro* has been shown to be a relatively efficient process (Paul *et al.*, 2003, Paul *et al.*, 2000).

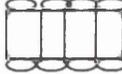
Following initiation of positive-strand synthesis, unwinding of the template by 3D^{pol} ensures that the residues at the 3' termini of the RF template are free to interact with further polymerase-primer complexes. A large pool of preformed primer would enable 3D^{pol} to rapidly initiate multiple rounds of positive-strand synthesis on any one RF template. Thus, asymmetric replication, which results in a 50:1 bias towards genomic-strand synthesis (Jarvis & Kirkegaard, 1992), occurs as a consequence of differential rates of initiation between positive- and negative-strands.

Figure 1.23 Model for the synthesis of positive-sense RNA strand

During negative-sense strand synthesis (dotted line) the CRE acts as a template for the production of VPg-pUpU. The VPg-pUpU is retained within the replication complex (A). Following completion of negative-sense synthesis the VPg-pUpU can associate with the terminal AA residues at the 3' end of the negative sense strand (B). The viral polymerase (blue sphere) can elongate the VPg-pUpU thus initiating positive-strand synthesis.



Key:

	3D ^{pol}		3CD ^{pro}
	3AB		PCBP
	3A		Membrane
	VPg-pUpU	Negative-Sense Strand
	PABP	—	Positive-Sense strand

1.11 Project aims

The information encoded by all RNA viruses is two-fold. At the primary level the sequence acts as a message carrying all the information required for synthesis of essential structural and non-structural virus protein to occur. However in RNA viruses the control of protein expression, genomic replication and even packaging of the progeny genomes resides in RNA secondary structure and RNA motifs.

The picornavirus family provides the largest single grouping of RNA viruses. Poliovirus, the prototype picornavirus, is arguably the best characterised of all mammalian viruses. The available information regarding the mechanism of picornavirus replication has not changed significantly in ten years. However, the identification of the CRE in HRV14 in the early 90's, has opened up the possibility that picornavirus genomes may contain further *cis*-acting replication elements that have yet to be identified. The overall aim of this project was to further the current understanding of picornavirus replication. Three approaches were used in this study of picornavirus replication:

- a) To identify the binding sites of the non-structural proteins 2C, 3AB, 3CD^{Pro} and 3D^{Pol} within the genome and anti-genome.
- b) To investigate whether highly conserved regions of the virus genome, which had been identified through the use of suppression of synonymous site variation (SSSV) analysis by a collaborator of the laboratory, were of functional significance in terms of picornavirus replication.
- c) Previous research in the laboratory had identified a coding change (N18Y) in the viral polymerase that compensated for deleterious changes in the 3'UTR (Meredith et al., 1999). It was hoped that characterisation of the effects of N18Y on the biochemical attributes of the virus polymerase would provide further insight into the importance of the 3'UTR in terms of the efficiency of picornavirus replication.

The “primordial soup” hypothesis states that all life originated from self-replicating RNAs. This evolutionary view is supported by the available evidence which suggests that, despite occupying a diverse range of niches, a basic replication strategy exists that is common to all RNA viruses. Thus, while the overall aim of this project was to obtain valuable information regarding the mechanics of picornavirus RNA synthesis, the information obtained, due to the common evolutionary origins of all RNA viruses, provides a valuable insight into replication strategies that may be employed by less easily studied viruses like the human caliciviruses and HCV.

2 Identification of protein–RNA interactions using the Yeast three-Hybrid system

2.1 Introduction

RNA-protein interactions are fundamental to a number of cellular processes including post-transcriptional regulation, translation and RNA stability. In addition, the targeted expression of eukaryotic genes has been demonstrated to be important for the correct development of *Drosophila* embryos, and morphogenesis of the algae *Acetabularia acetabulum* occurs as a result of interactions between RNA transcripts and specific cellular proteins (Edwards et al., 2001, also reviewed in Okita & Choi, 2002). Amongst RNA viruses the specific interactions that occur between proteins, of cellular and viral origin, and the viral RNA (including in certain viruses subgenomic mRNAs) are critical to the efficiency with which the virus can replicate. Indeed the ability of the viral RNA to efficiently interact with cellular proteins has been shown to be a determinant of cell tropism and virulence. Interruption of protein-RNA interactions important for virus genomic replication provides one area for therapeutic drug development to explore.

Research over the past 20 years has elucidated the significance that RNA-protein interactions have for the successful replication of the picornaviruses. It has been demonstrated that RNA-protein interactions modulate the use of the viral genome between its functions as a template for translation or genomic amplification. It is also likely, though no encapsidation signal has yet been found, that an interaction between a viral protein and the positive-sense genome is required for the specific encapsidation of the genomic RNA into the virus particle. While the binding sites for a number of the cellular proteins involved in the translation of the genome have been elucidated (see table 1.5) the interactions of the individual viral proteins with the viral RNA has been less comprehensively documented.

Several techniques have been developed to facilitate the study of a number of RNA-protein interactions. These include phage display (Laird-Offringa & Belasco, 1995) and methods of *in vitro* selection like SELEX (systemic evolution of ligands by exponential enrichment) (Ellington & Szostak, 1990, Tuerk & Gold, 1990). SenGupta and colleagues have

developed an *in vivo* yeast-based system to identify proteins that interact with RNA (SenGupta et al., 1996). This method is based on the observation that a number of cellular transcription regulators are composed of two domains, a DNA binding domain and an activation domain. Brent and Ptashne (Brent & Ptashne, 1985) showed that the activation domain of the yeast transcription factor GAL4 could be fused to the DNA binding domain of the *E. coli* LexA transcription factor to create a functional transcriptional activator. Subsequent experiments by a number of groups, (Ma & Ptashne, 1988, Triezenberg et al., 1988) demonstrated that the activation domain did not have to be part of the same polypeptide as the DNA binding domain to reconstitute a functional transcriptional activator. Fields and Song subsequently demonstrated that this feature could be used to study the interactions between proteins, which are not transcription factors, if the potential interactors were expressed as chimeras with the DNA binding domain or the activation domain (Fields & Song, 1989). The method developed by Fields and Song, termed the yeast two-hybrid system, utilised yeast genetics to allow the ability of the yeast to synthesise amino-acids to function as a reporter gene alongside the more conventional reporter gene β -galactosidase.

The yeast three-hybrid screening method, developed by SenGupta *et al* is an extension of the yeast two-hybrid system to allow the study of protein-RNA interactions. It consists of three (fusion) constructs that interact to reconstitute the transcriptional activator. The first of these fusion proteins is constitutively expressed in *Saccharomyces cerevisiae* (*S.cerevisiae*) strain R40coat. In the system that was utilised in this study the fusion protein consists of a fusion between a GAL4 binding domain (GAL4BD) and the bacteriophage MS2 coat protein. The second hybrid construct in the system is often referred to as the “bait” hybrid as it contains a fusion of the 19nt RNA hairpin that has been shown to interact directly with the bacteriophage MS2 coat protein, and the RNA of interest.

The third hybrid construct, termed the “prey” allows for the fusion of the activation domain of GAL4 in-frame with a known or unknown protein of interest that is believed to interact with the expressed RNA. It is only when the RNA expressed as a fusion with MS2 interacts with the “prey” protein that the GAL4 activation domain (GAL4 AD) can interact with the GAL4 BD thereby releasing the repression of reporter genes that are under the control of a GAL4 promoter (Figure 2.1A). The regulatory regions for the reporter gene *lacZ* and the auxotrophic marker histidine contain a number of copies of the GAL4 binding site that act as upstream activation sequences (UAS) in yeast.

Traditionally the yeast three-hybrid system is used to “fish” for proteins that interact with an RNA sequence of known importance. This approach has been used successfully to identify cellular proteins that interact specifically with sequences and structures identified in mRNAs and viral RNAs. However, the yeast three-hybrid can also be used to “fish” for RNA that interacts with a protein of interest. Although not used as frequently the use of the yeast three-hybrid system as a technique to identify RNA has been successfully employed by groups researching in the fields of virology (Kang et al., 1999, Sokolowski et al., 2003) and cellular biology (Edwards et al., 2001, Sengupta et al., 1999). The work documented in this chapter explores this approach as a method to identify putative *cis*-acting replication elements within the genome of poliovirus type 3.

Previously it had been demonstrated that the P1 region of the poliovirus genome could be deleted without affecting the ability of the virus to replicate (Percy et al., 1992) so the investigation concentrated on examining the P2 and P3 regions of the genome. To do this a cDNA library was created that would be representative of the entire P2 and P3 regions of the viral genome. The library consisted of 50-250 bases of cDNA fused with a cDNA of the 19nt MS2 hairpin. The 50-250 base cDNA fragments were generated by restriction enzyme digestion. The viral proteins 2C, 3AB, 3CD^{pro} and the viral polymerase (3D^{pol}) that have already been characterised as having important roles in viral replication were used to “fish” the cDNA library to identify RNA sequences that specifically bound the viral proteins. Subsequent mutational analysis would be used to confirm a role for any RNA sequences retrieved from the cDNA library as *cis*-acting replication elements. Any subsequent analysis would also hopefully shed light on the exact function the retrieved RNA sequence influences and consequently improve the current understanding of the mechanics of poliovirus replication.

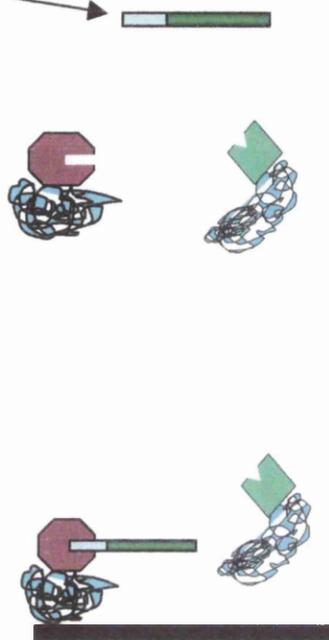
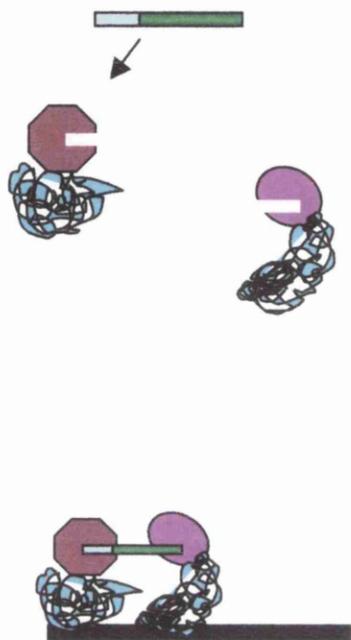
Figure 2.1 Illustration of the principle behind the yeast three-hybrid system.

The yeast- three hybrid system consists of three constitutive parts a bait protein containing a transcriptional binding domain (blue domain) fused to the MS2 coat protein (mauve). An RNA expressed as a fusion with the MS2 packaging signal, and lastly a prey protein containing a fusion of a protein of interest (purple) with a transcriptional activator (blue domain). If the RNA and protein interact (2.1A) transcriptional repression is relieved and the reporter gene is expressed. Transcriptional repression is not relieved if the protein and RNA fail to interact (2.1B).

2.1A

2.1B

Hybrid RNA



Interactions between the two proteins results in the release of transcriptional repression of the histidine and β -galactosidase genes

No interaction between the two proteins results in the maintenance of transcriptional repression

Key:

Component	Definition	Component	Definition
	Bacteriophage MS2 coat protein recognition sites		Prey protein A
	RNA "bait"		Transcription factor activation domain
	Bacteriophage MS2 coat protein		Prey protein B
	Transcription factor binding domain		

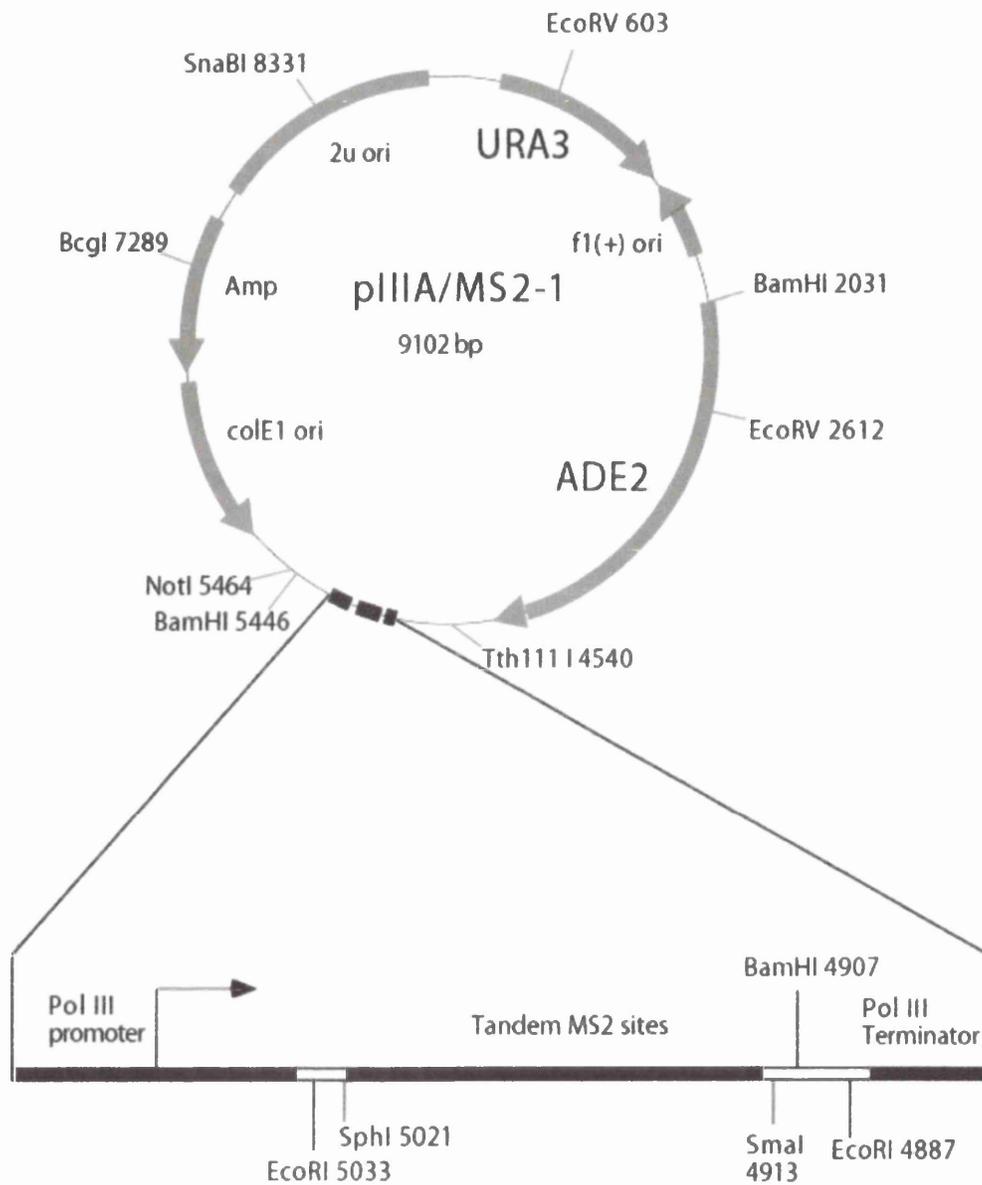
2.2 Creating a cDNA library

2.2.1 pIIIA/MS2-1

The hybrid RNA functions within the yeast three-hybrid system to lock the RNA to be screened onto the promoter binding protein. The interaction between the bacteriophage MS2 hairpin and coat protein is the most commonly used RNA-protein interaction to accomplish this. This is primarily because the bacteriophage MS2 coat protein and its binding partner interact with high affinity (Lowary & Uhlenbeck, 1987).

The plasmid pIIIA/MS2-1 (Figure 2.2) is an expression vector containing a cDNA used to produce the hybrid RNA following transformation into *S. cerevisiae*. The RNA is transcribed in large quantities (up to 1300 copies per cell) from the RNAase P promoter present in the plasmid by RNA polymerase III (reviewed by Bernstein et al., 2002). The RNA produced contains 5' and 3' sequences derived from RNAase P and two MS2 coat protein binding sites (Figure 2.3A). The introduction of two MS2 coat protein binding sites into the hybrid RNA utilises the known ability of the MS2 coat protein to bind cooperatively to adjacent MS2 binding sites to strengthen the interaction between the hybrid RNA and the promoter binding hybrid protein. RNA sequences to be screened using the yeast three-hybrid system can be cloned into the cDNA using either the *Sma I* or the *Sph I* site. This orientates the RNA of interest downstream or upstream, respectively, of the twin MS2 coat protein binding sites.

The plasmid pIIIA/MS2-1 also contains the genes *URA3* and *ADE2* the gene products of which enable selection of yeast maintaining the plasmid on uracil deficient and adenine deficient medium respectively. The presence of the *ADE2* gene on the cDNA is of particular importance as selection on adenine deficient media provides a useful colorimetric method to identify false positives. In the absence of adenine, *S. cerevisiae* R40coat strain attempts to synthesise adenine in a *de novo* manner, which results in the accumulation of a red metabolite. However, in cells capable of synthesising the *ADE2* gene product no accumulation of the red metabolite occurs and the cells remain white in colour.



SmaI and SphI sites are unique and can be used to insert fragment.

Figure 2.2 vector map of pIIIA/MS2-1 used in the construction of the cDNA library.

Taken from <http://www.biochem.wisc.edu/wickens/3H/>

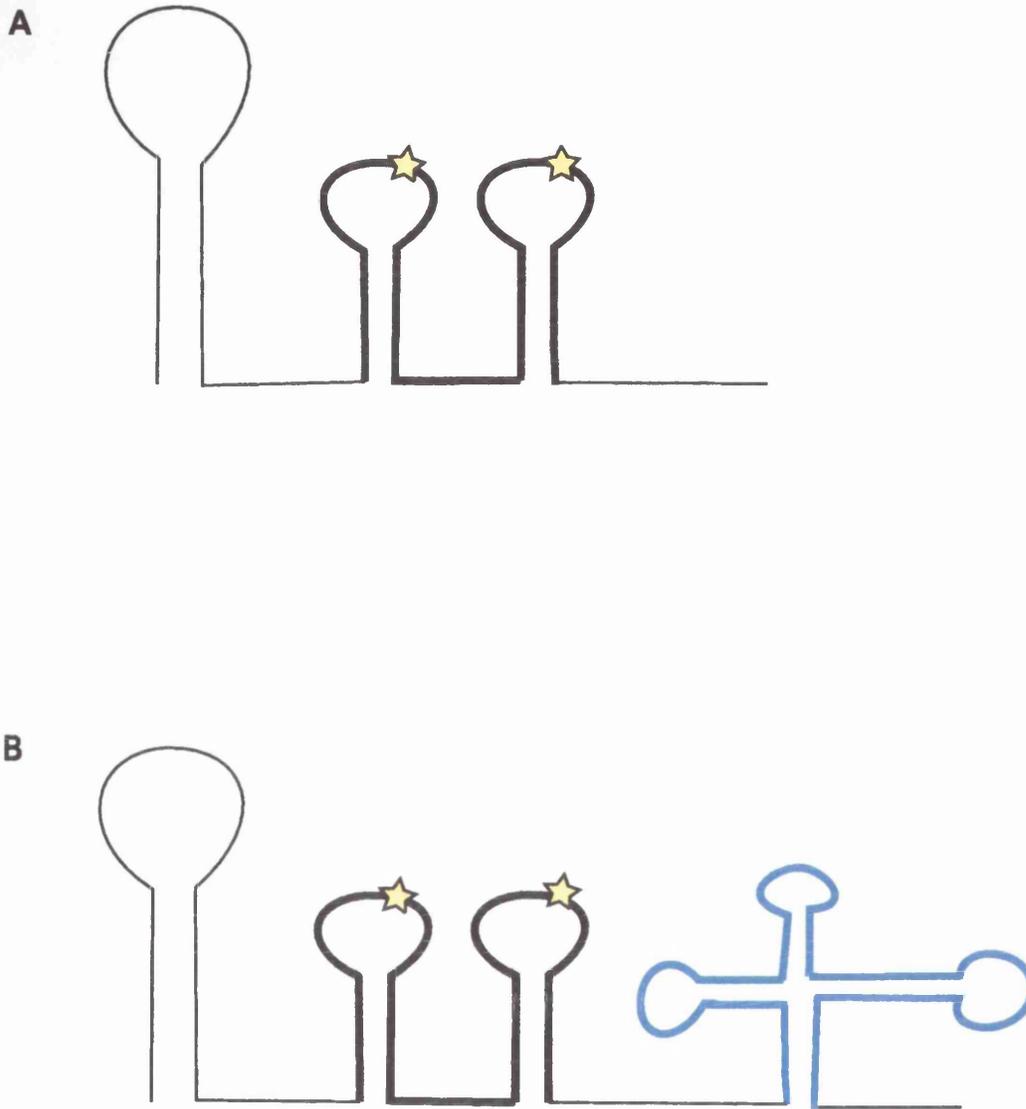


Figure 2.3 Schematic structure of hybrid RNAs expressed from the cDNA p IIIA/MS2-1 upon transformation into *S.cerevisiae* strain R40coat. The hybrid RNAs are transcribed by RNA polymerase III from an RNAase P promoter. The hybrid RNA (A) maintains the 5' stemloop structure and the 3' end of the RNAase P RNA (thin lines). The internal region of the RNAase P RNA has been removed and two MS2 coat protein binding sites have been introduced (bold black line). Both of the MS2 coat protein binding sites contain a mutation (star) known to enhance the RNA-protein interaction (Lowary and Uhlenbeck, 1987). An example of a hybrid RNA is shown in schematic B. This shows the RNA of interest, in this case the cloverleaf (bold blue line), fused downstream of the MS2 sites.

A cell that can activate the HIS3 gene independently of the RNA plasmid will “lose” pIII/MS2-1 during the course of replication. In colonies that have lost pIII/MS2-1 a red colour will appear as the adenine in the media becomes limiting.

2.2.2 Generation of cDNA fragments

To identify novel viral protein-RNA interactions involved in viral replication, using the yeast three-hybrid system, a cDNA library was created. As it had been well documented that the sequences encoding the capsid proteins could be removed without affecting poliovirus replication these sequences were excluded from the cDNA library.

The decision to exclude the sequences of the structural proteins from the cDNA library to be screened meant that the sequences of the non-structural proteins could be derived from cDNA's of subgenomic replicons available within the laboratory. pT7FLC/Rep is a subgenomic replicon that was constructed from the poliovirus type 3 Leon strain. In pT7FLC/Rep VP4, VP2 and the amino terminus of VP3 have been deleted and replaced by the CAT gene, in-frame with the rest of the viral genome. Monitoring of the production of CAT by enzymatic assay following transfection of RNA, transcribed from pT7FLC/Rep, into mammalian cells showed that the RNA was amplified following transfection and could be packaged *in trans* by helper virus (Percy et al., 1992). The viral sequences used in the creation of the cDNA library were obtained from a subgenomic replicon derived from pT7FLC/Rep called pT7FLC/Rep3. The only difference between pT7FLC/Rep3 and pT7FLC/Rep is that pT7FLC/Rep3 contains a larger deletion of the P1 region of the genome (Barclay et al., 1998).

Analysis of the sequence of pT7FLC/Rep3 using the computer package Vector NTI identified three restriction enzymes that recognise a number of sites distributed throughout the cDNA of the viral genome. All three restriction enzymes, *Rsa I*, *Alu I* and *Hae III*, recognise four base palindromic sequences and leave a blunt-end following restriction of the sequence. Figure 2.4 shows the distribution of the palindromic recognition sites of each enzyme within the plasmid pT7FLC/Rep3.

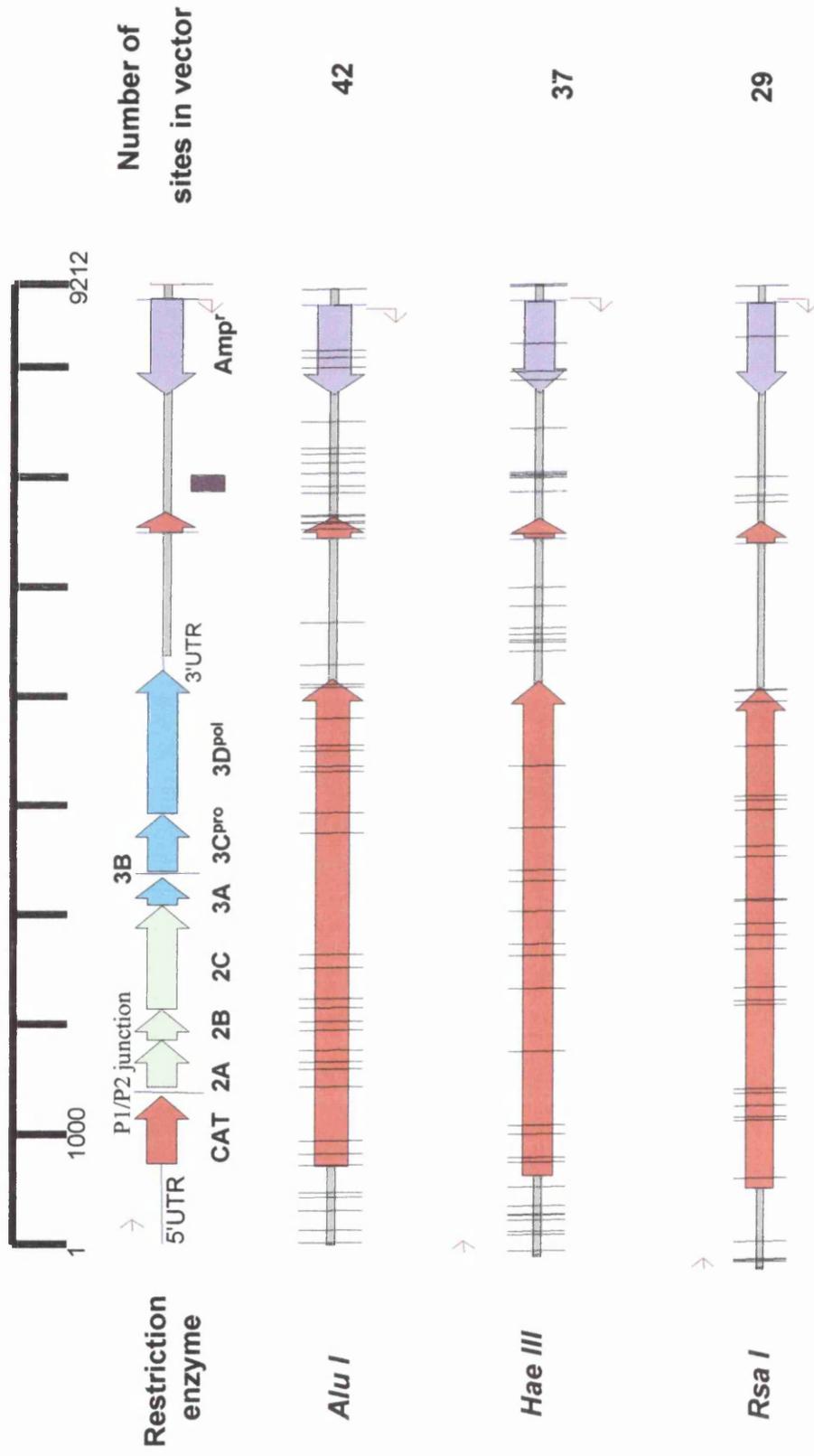


Figure 2.4 Distribution of restriction enzyme sites *Alu I*, *Hae III* and *Rsa I* within the cDNA pT7FLC/Rep3

Figure 2.4 shows that though the palindromic sequences recognised by each restriction enzyme are widely distributed throughout the genome relatively large fragments exist where, independently, no recognition sites are present. For example, the Vector NTI analysis of the location of the four base palindromic sequence recognised by *Alu I* (AGCT) identified a 1.2 kb fragment in the region of the cDNA encoding the viral proteins 3A, 3B and 3C^{pro} which was resistant to digestion (Figure 2.4). However, the Vector NTI analysis of the location of the four base palindromic sequences recognised by *Hae III* (GGCC) and *Rsa I* (GTAC) showed the presence of five and six recognition sequences within this fragment respectively. The overlapping coverage between the three enzymes is such that complete restriction enzyme digestion of the entire plasmid pT7FLC/Rep3, using an enzyme mix containing *Rsa I*, *Alu I* and *Hae III*, identifies only four regions of the genome where fragments larger than 200 bases in size are obtained. Two of these lie within the cDNA encoding the viral proteins and generate fragments of 222 and 247 bases in size respectively and two are located in the vector sequence (502 and 375 bases in size respectively). Overall, the average size of fragment produced following complete restriction enzyme digestion of pT7FLC/Rep3 is 87 nucleotides in length, which is reduced to 77 bases if only the sequences corresponding to the viral cDNA are included for analysis.

It was important that the cDNA library represented every sequence and possible structure in the viral genome. The decision to create the cDNA library using a method of partial restriction enzyme digestion was taken on the basis of the information on fragment size provided by the computational analysis. Analysis of the products of complete restriction enzyme digestion using Vector NTI of fragments derived from the cDNA of the viral genome showed that a number of fragments which would be generated were smaller than 50 nucleotides in size. These sequences comprised 36% of the total number of cDNA fragments obtained from the viral genome and ranged from 6-49 bases in size with an average size of 21 nucleotides. It was believed that a number of these small fragments were too small to form secondary structures or be successfully purified from agarose following complete digestion of pT7FLC/Rep3. Adding further support to the decision to use a method of partial restriction digestion to create the cDNA library was the observation that a number of the small fragments were derived from the poliovirus 5'UTR which contains two well-characterised and essential *cis*-acting secondary structures: the IRES and the cloverleaf.

The cloverleaf structure constitutes the first 100 nucleotides of the viral genome. Vector NTI analysis showed that the first 127 nucleotides of the cDNA sequence contained two *Alu I*, one *Hae III* and three *Rsa I* recognition sites and that digesting the plasmid pT7FLC/Rep3 to completion using these enzymes would result in the production of fragments of 10, 26, 17, 16, 8 and 50 bases in size. These fragments are too small to generate the independent stem-loops that form the cloverleaf structure. This is important as an interaction between the viral protein 3CD^{pro} and the cloverleaf structure has been documented to be essential to the replication of the virus (Andino et al., 1990a). The presence of the cloverleaf in the cDNA library, either as a complete structural motif or as independent stem-loops structures, provides an important internal control for the yeast three-hybrid cDNA library screen.

The optimal fragment size for screening using the yeast three-hybrid is between 100-150 bases in size. To ensure that every sequence was represented in the cDNA library the decision was made to optimise the generation of cDNA fragments in the range of 50-250 bases. This decision was made so that the cDNA library would include the two fragments of 222 and 247 bases that are not restricted by the three enzymes used to generate the cDNA library. It was decided to set the lower range of fragments of cDNA at 50 bases to ensure that every fragment of cDNA in the library was capable of forming secondary structure. Optimisation of the partial restriction enzyme digestion method to generate fragments ranging in size between 50-250 bases identified the optimal conditions as being 30 minutes at 37° C using 0.1 U of each enzyme in a final volume of 50µl. An additional benefit of using a partial restriction enzyme digestion method to create the cDNA library should be that each fragment generated is random in nature as it can not be pre-determined which sites recognised by the restriction enzymes will be cleaved during the digestion period.

2.2.3 Construction of the cDNA library

The cDNA fragments were purified from agarose and ligated into the plasmid pIII/MS2-1 at the *Sma I* (Figure 2.2) site thus fusing the cDNA fragment downstream of the MS2 sites. This is illustrated schematically in figure 2.3 B. Prior to ligation with the purified cDNA fragments the linearised pIII/MS2-1 was treated with 1 U shrimp alkaline phosphatase for 1 hr at 37 °C to minimise the risk of recovering religated pIII/MS2-1 following transformation. The ligations were redigested with *Sma I* to further minimise the possibility of recovering religated vector, purified and transformed into *E. coli* DH5 α using standard electroporation conditions (see section 6.4.10.2). Transformed colonies were selected by growth on media containing ampicillin.

The genomic coverage that was required to ensure that the cDNA library was truly representative was calculated using the equation (Sambrook et al., 1989):

$$N = \ln(1-P) / \ln(1-F)$$

$$F = \text{median size of fragment} / \text{total genome size}$$

$$P = \text{probability required}$$

$$N = \text{Number of independent clones required}$$

Using the above equation it was calculated that the number of independent clones in the cDNA library that were required to guarantee 99.99 % probability that the entire P2/P3 region was represented within the cDNA library from the partial restriction digest of pT7/Rep3 was 635. The number of independent colonies within the cDNA library created by partial restriction digest was 2167.

Colony polymerase chain reaction (PCR) was used to identify the presence of insert in a number of independent clones in the cDNA library (Figure 2.5). A number of the amplified PCR products appear to be greater in size than 100-200 nt in length (marked by * in figure 2.5). This can be explained in two ways. Firstly, more than one fragment may be ligated end-end in the cDNA or secondly it is DNA greater than 100-200 nt in size that has been carried through the purification procedure. Automatic sequencing of the purified PCR product of the seventeen clones, which generated a positive signal, using the forward

oligonucleotide primer O-AB-6I (see Table 6.2 materials and methods) was carried out at the in-house facility.

The results of the sequence analysis of the seventeen cDNAs are shown in table 2.1. Table 2.1 shows that the sequences obtained from eleven of the seventeen were derived from the poliovirus cDNA. The sequences obtained although few in number do offer some insight into the created cDNA library. The sequences of the cDNA correspond to both the genomic-sense poliovirus RNA and the negative-sense replication intermediate. For instance, the sequence of cDNA obtained from clones 2 and 4 shows that both these clones express a hybrid RNA corresponding to nucleotides 3648-3749 of the PV3 genome. However cDNA 2 expresses the sequence in its genomic form whilst the hybrid RNA expressed by cDNA 4 is the anti-sense form of the same sequence. This provides evidence that the cDNA library provides "bait" for interactions that the viral proteins may have with the negative-sense replication intermediate as well as interactions that may occur between the proteins and the genomic RNA. Secondly, the sequences corresponding to the poliovirus genome that were obtained cover almost the entire length of the poliovirus genome. This level of coverage suggests that the cDNA library should be representative of the entire viral genome. Lastly, of the seventeen cDNA sequenced two, cDNAs 1 and 18, express RNA of the same region though not the same size. The RNA produced by cDNA 1 covers nucleotides 4407-4510 of the poliovirus genome, whilst the RNA expressed by cDNA 18 is much larger in size (nt 4158-4532). This sequence data informs that some fragments larger than 250 bases in size have come through the purification step in creating the library. More importantly, the presence of cDNAs fragments within the library that overlap smaller cDNA fragments of the same region, should enable the identification of the minimum protein-binding site of the viral protein through the identification of the smallest cDNA fragment that can be successfully retrieved by the viral protein.

Although sequencing of 0.1% of the cDNA is not statistically enough to determine that the library was random and nonbiased the sequencing of the cDNA strongly suggests that the partial restriction digestion method has created a cDNA library that is random in nature and that the restriction enzymes have not biased the library towards particular regions of the genome or the size of cDNA fragments obtained.

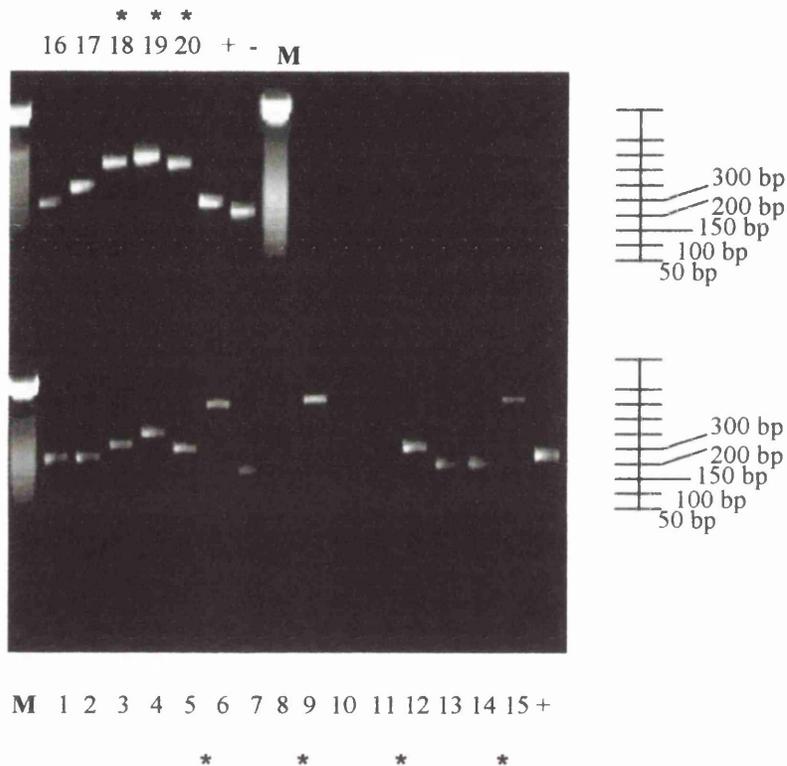


Figure 2.5 Colony Polymerase Chain Reaction of independent clones from the cDNA library.

PCR analysis of the cDNA library was used as a preliminary analysis of independent clones to confirm the presence of insert. PCR analysis was carried out using oligonucleotide primers MS2 leader and MS2 terminator that bind to the regions flanking the insert. The presence of insert was determined by comparing the PCR products of independent cDNA clones with a known positive control, pIIIaMS2-1^{2Cre}(+), and negative control pIIIaMS2-1 (-). Independent clones that show the presence of insert are greater than 100-200 bases in size are marked with a *.

cDNA number	PV3 nt	Orientation	cDNA number	PV3 nt	Orientation
1	4407-4510	Sense	13	Cloning vector	-
2	3749-3648	Antisense	14	Cloning vector	-
3	Cloning vector	N/A	15	2431-2642 + Additional sequences	Antisense N/A
4	3561-3749	Sense	16	1120-1105	Antisense
5	632-492	Antisense	17	318-372	Sense
6	-	-	18	4158-4532	Sense
7	6049-6008	Antisense	19	4295-4441	Sense
9	3940-3763	Antisense	20	6553-6407	Antisense
12	Cloning vector	-			

Table 2.1. Sequence analysis of PCR products amplified from cDNA library. N/A, not applicable; -, no insert

2.2.4 Identification of known RNA cis-acting replication elements in cDNA library

As a number of *cis*-acting replication elements have been identified in the poliovirus genome it was important that these should be represented in the cDNA library (Lib_{RE}). PCR was carried out using primers designed to amplify the cloverleaf, CRE and the 3'UTR (Materials and methods Table 6.2). To confirm that the primers did not bind in a non-specific manner to the plasmid vector, pIII/MS2-1 was used as negative control for the PCR reaction. As a positive control for the PCR the replication signals were amplified from the subgenomic replicon pT7FLC/Rep3. A second library (Lib_{DNAase}) that was constructed by Dr. I Goodfellow in the laboratory, by partial DNAase treatment was also tested by PCR for the presence of the three known replication signals.

Confirmation of the presence of the cloverleaf within the cDNA library provides an internal control for the detection of RNA elements by the viral protein 3CD^{Pro}. The cloverleaf was amplified using the forward primer O-AB-2I (for primer details see table 6.2 materials and methods) that binds to nucleotides 1-18 of the poliovirus cloverleaf. Primer ID1 was used in conjunction with the reverse primer O-AB-3I that binds to nucleotides 89-106 of the poliovirus cloverleaf. Figure 2.6A shows that the cloverleaf could be detected in both cDNA libraries using PCR.

The second *cis*-acting replication element that was analysed for its presence within the cDNA libraries was the CRE. The CRE was amplified by the primers 2C_{ss} and 2C_{as} that have been documented previously (Goodfellow et al., 2000b). Figure 2.6B shows that using these primers the CRE could be shown to be present in both cDNA libraries.

Lastly it was decided to see if the poliovirus 3'UTR could be detected in either of the cDNA libraries. The 3'UTR was amplified using the forward primer IG10 that binds to nucleotides 7208-7308 in conjunction with the reverse primer IG11 that binds to nucleotides 7392-7308. Figure 2.6C shows that the 3'UTR could be detected in both cDNA libraries.

Figure 2.6 therefore confirms that all three of the known replication signals are represented within both cDNA libraries.

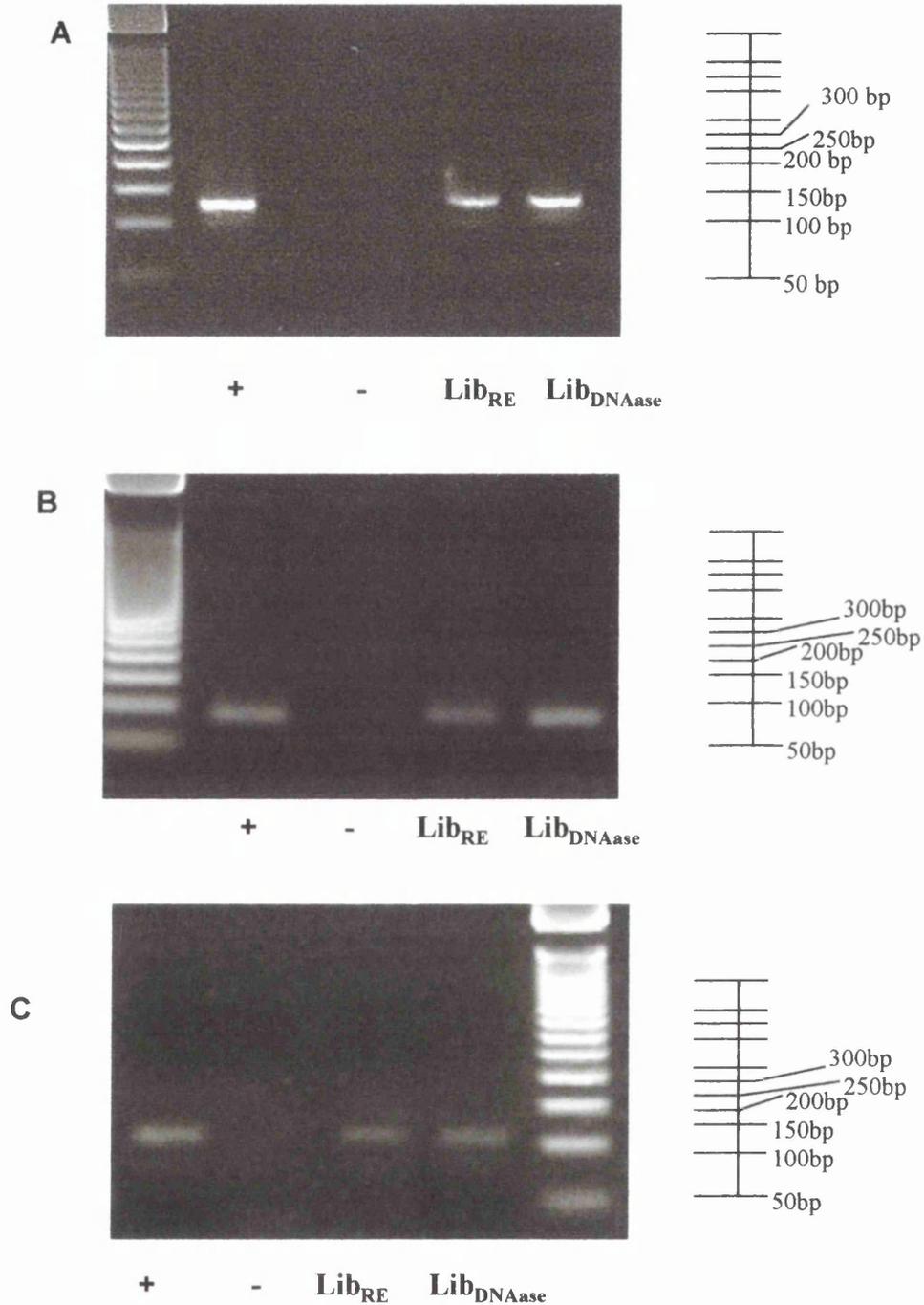


Figure 2.6. PCR screening for the presence of known CRE within the cDNA library.

PCR analysis was used to confirm the presence of CRE within the cDNA library. The presence of CRE was determined by comparing the PCR products of the cDNA library with a known positive control, p T7FLC/REP3 (Percy *et al.*, 1992) (+), and negative control pIIIaMS2-1 (-). The PCR analysis showed the presence of the cloverleaf (2.6A), CRE (2.6.B) and 3'NTR (2.6C) within cDNA libraries made by partial restriction digest (libRE) and DNAase treatment (libDNAase).

2.3 cDNA expressing viral non-structural proteins

The non-structural proteins that were used in this investigation were 2C, 3AB, 3CD^{pro} and 3D^{pol}. To express the viral non-structural proteins as fusions with the GAL4 AD the coding sequences were fused in-frame with the coding sequence of the GAL4AD in the plasmid vector pACTII (Figure 2.7).

The non-structural proteins 2C, 3AB and 3D^{pol} had been cloned into pACTII by Dr.I.Goodfellow. Previously, a coding change of a histidine to an alanine residue at residue 40 of 3C^{pro} had been engineered into a cDNA of 3CD^{pro} that was available in the lab. This coding change renders 3C^{pro} proteolytically inactive (Hammerle et al., 1991, Paul et al., 2000). This plasmid, pET3CD^{H40A} was used to obtain the sequences required to construct a GAL4AD/3CD^{H40A} fusion.

3CD^{H40A} was preferred in this screen to allow identification of RNA that was bound by 3CD^{pro} and prevent identification of RNA bound by 3C^{pro} alone. In addition the use 3CD^{H40A} would possibly result in a reduction in any toxicity that may be associated with expressing a protein with proteinase activity within *S.cerevisiae*.

The sequence of 3CD^{H40A} was amplified using the primers IG25 and IG26 (Materials and Methods Table 6.2). Following amplification the purified PCR product was digested with the restriction enzymes *Not I* and *Bam HI*. The PCR product and pACTII, which had been digested in a similar way, were purified from agarose prior to being ligated together. The ligations were purified prior to transformation into *E.coli* DH5 α . The presence of insert was detected by digesting the cDNA recovered from the transformants with *Not I* and *Bam HI*. Plasmids positive for insert showed the presence of an additional band of approximately 2 kb in size in comparison to pACTII, which had been similarly digested with *Not I* and *Bam HI*, when visualised by agarose gel electrophoresis.

The constructs used and the predicted molecular weights of the fusion protein expressed from the plasmids are summarized in table 2.2.

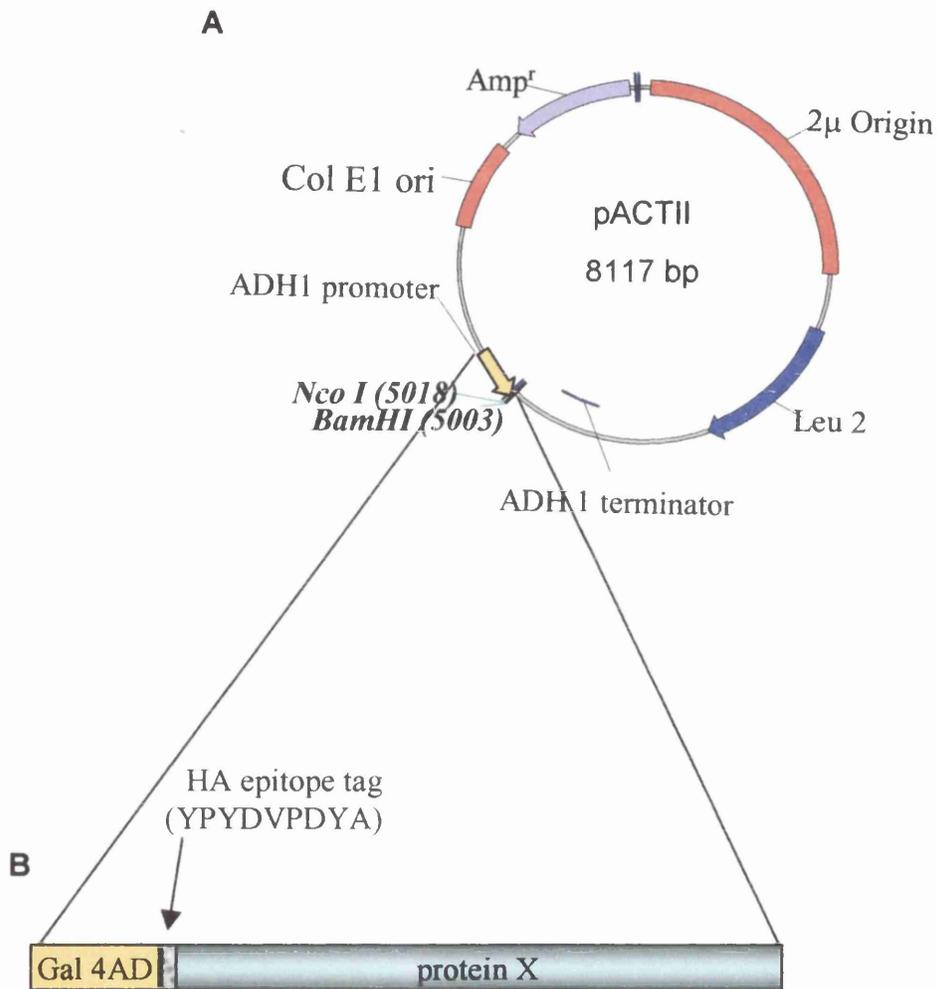


Figure 2.7 Yeast three hybrid vector map.

All constructs are cloned into the expression plasmid pACTII using the *Nco* I/ *Bam*HI restriction sites (A). Expression of the protein in *S.cerevisiae* generates an N-terminal fusion of the cloned protein with the Gal 4 activation domain (B).

cDNA (Laboratory identification code)	Viral protein expressed as fusion with GAL4 activation domain	Functions associated with viral protein	Predicted Molecular weight of protein + Gal4AD tag (19.1kDa)
pACTII2C	2C	<ul style="list-style-type: none"> •ATPase •Encapsidation •Role in RNA synthesis •Membrane rearrangements 	56.6 kDa (37.5 kDa- Gal4AD tag)
pACTII3AB	3AB	<ul style="list-style-type: none"> •Membrane association •Stimulation of 3D and 3CD •Essential for RNA replication 	31.2 kDa (12.1kDa – Gal4AD tag)
pACTII3D	Viral polymerase (3D)	<ul style="list-style-type: none"> •RNA synthesis •Polymerisation-dependent unwinding of ds RNA •Uridylylation of VPg •Interaction with Sam 68 	71.4kDa (52.4 kDa- Gal4AD tag)
pACTII3CD ^{H40A}	3CD ^{H40A} (Proteolytically inactive form of the protein)	<ul style="list-style-type: none"> •Viral protein processing, •Formation of RNP complex 	91.2 kDa (72.1kDa-Gal4AD tag)

Table 2.2 cDNAs expressing viral protein- Gal4AD fusion used during yeast three-hybrid screen

2.3.1 Expression of the viral proteins in *S. cerevisiae* R40

Expression of foreign proteins within *S. cerevisiae* can result in morphological changes such as a small colony phenotype when the yeast cells are cultured on agar and slow growth in liquid culture. Any changes in the morphology are a consequence of a feature of the expressed protein that is toxic to the yeast. Previously it has been shown that the poliovirus non-structural proteins 2A^{pro} and 2BC were toxic when expressed in yeast (*S. cerevisiae*) (Barco & Carrasco, 1995, Barco et al., 1997). Barco and Carrasco have demonstrated that the ability of yeast expressing 2A^{pro} to grow corresponds to the acquisition of mutations which decrease the proteolytic activity and transactivation functions of the expressed protein (Barco et al., 1997).

Transformation of *S. cerevisiae* with the fusion protein constructs pACTII3AB, pACTII3CD^{H40A} or pACTII3D^{pol} did not result in any obvious morphological changes that would correlate with the expression of the proteins being toxic to the *S. cerevisiae* R40coat cells. Transformation of *S. cerevisiae* R40coat with the fusion protein construct pACTII2C resulted in the slow growth of the cells using both solid and liquid culture mediums. The slow growth of the cells was observed phenotypically as small colony size when cultured on solid media. This is a classic phenotype of yeast cells expressing a toxic product. The observation that 2C was toxic when expressed in yeast was unsurprising as it had been previously observed that 2BC was toxic to yeast cells (Barco & Carrasco, 1995). The toxicity of both 2BC and 2C is likely to relate to their ability to modify cellular membranes (Cho et al., 1994).

To confirm that the poliovirus fusion proteins were expressed in *S. cerevisiae* R40 protein expression was determined by western blot analysis. Though the viral proteins are fused to the (19.1kD) GAL4 activation domain, the pACTII vector also contained an influenza haemagglutinin epitope, YPYDVPDYAG, (Figure 2.7), which was expressed as part of the fusion protein. To detect expression the fusion proteins that has been immobilised onto a nitrocellulose membrane were probed with a mouse monoclonal antibody that had been raised to the haemagglutinin epitope, diluted 1:5000 in 4% milk powder/PBS. The use of an antibody raised to the haemagglutinin epitope to detect the immobilised protein enabled the same antibody to be used to detect the expression of the different viral proteins. Following washing with PBS/0.1% Tween 20, the blot was incubated with HRPO

conjugated anti-mouse IgG (Pierce). The immunoreactive proteins were visualised after further washing using the Supersignal West chemiluminescent detection kit (Pierce).

Figure 2.8 shows that expression of all the fusion proteins could be detected by western blotting. The inherent toxicity of 2C raised the concern that expression of the protein would result in the selection of variants of 2C that had been grossly truncated. Despite the obvious toxicity to the yeast cells expressing viral protein 2C, western blot analysis showed that the 2C was not apparently truncated by the yeast cells (Figure 2.8B). No further analysis was carried out to identify whether the 2C expressed by the yeast carried mutations that would abrogate any of the protein functions.

The molecular weight of the fusion protein of 3AB and the GAL4AD was predicted at 31.2 kDa. The molecular weight of the fusion protein was calculated as the combined total of the molecular weight of the GAL4 AD tag, as provided by Clontech, and 3AB, as calculated by Pallansch *et al* (Pallansch *et al.*, 1984). Unexpectedly the majority of the antibody could be consistently shown to react with a protein that migrated at approximately 50kDa. A faint band marked by a black arrow on figure 2.9C is localised to the predicted molecular weight. As no bands of this size were obtained when probing the extracts of *S.cerevisiae* transformed with cDNA expressing the other viral proteins (Figure 2.8 A, C, D and E) this cannot be explained by antibody cross-reaction. One possible explanation for the observed increase of the molecular weight of 3AB could be due to its hydrophobic nature. Within virus-infected cells, 3AB is inserted into the membrane. It is possible that the higher molecular weight observed for this protein is the result of an aggregate of membrane or protein that was not disrupted by the boiling of the sample prior to loading on the protein gel.

The molecular weights of the fusions between the GAL4 AD and the viral proteins 3D^{pol} and 3CD^{H40A} were predicted to be 71.5 and 91.2 kDa respectively. Western blot analysis (Figure 2.8D and E respectively) confirmed that both 3D^{pol} and 3CD^{H40A} were expressed by *S.cerevisiae*. An additional band was detected when probing for 3CD^{H40A} (Figure 2.8E). Further western blot analysis showed that this band was not always present suggesting that it is a product of proteolytic degradation or mechanical shearing of the protein that has occurred as a result of the process used to obtain the yeast extracts.

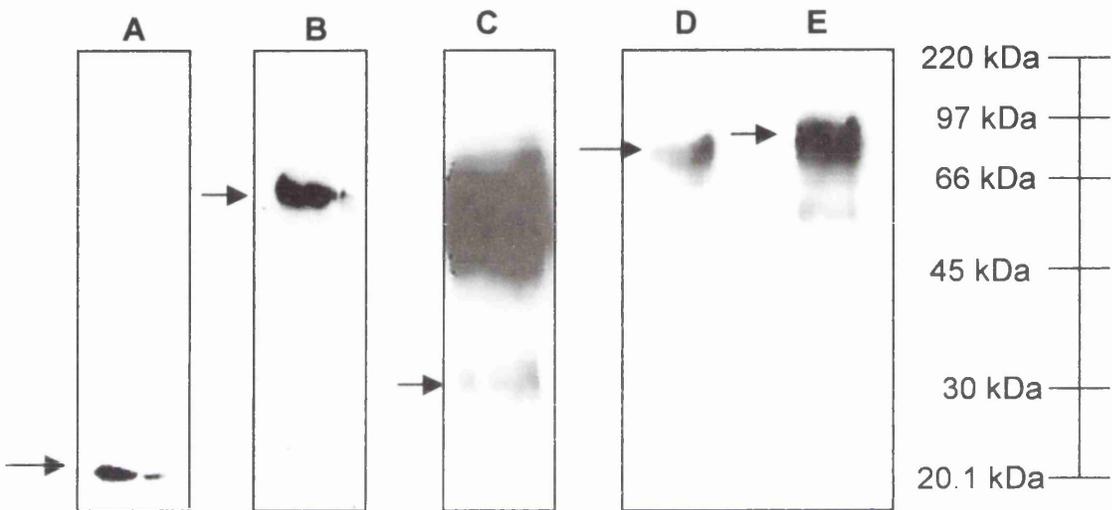


Figure 2.8 *Western blot analysis of yeast extracts*

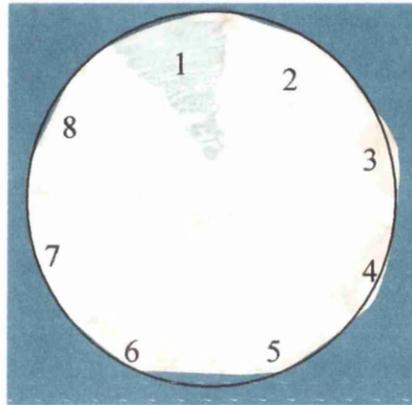
Expression of the Gal4 AD alone (**A**). Expression of fusion of Gal4AD and viral protein 2C (**B**). Expression of fusion of Gal4AD and viral protein 3AB (**C**). Expression of Gal4AD fused with viral protein 3D^{pol} (**D**) and 3CD^{H40A} (**E**). Expressed protein with the expected molecular weight for each fusion protein have been marked with a black arrow.

2.4 Identification of a characterised viral protein-RNA interaction

Although expression of the full-length fusion proteins in *S.cerevisiae* could be detected immunologically it was not known whether the yeast three-hybrid would be able to detect interactions with the viral RNA when it was expressed as a fusion.

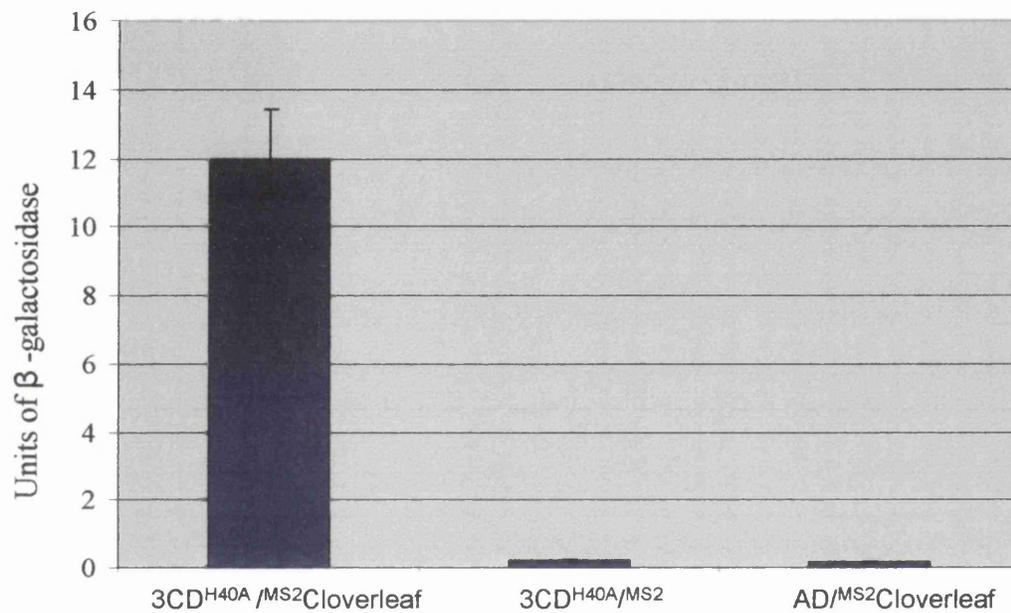
One of the best characterised protein/RNA interactions in poliovirus occurs between the 5' terminal cloverleaf RNA structure and viral protein 3CD^{pro}. Using the yeast three-hybrid screening system interactions between protein and RNA can be identified in two ways both qualitatively by β -galactosidase filter-lift assay and quantitatively by liquid β -galactosidase assays. To confirm that the yeast three-hybrid system could be used successfully to distinguish positive viral protein/RNA interactions a β -galactosidase filter lift assay was carried out using *S. cerevisiae* transformed with the fusion proteins and with a plasmid encoding only the MS2 hairpin RNA or a plasmid containing a fusion of the MS2 hairpin RNA with the cloverleaf or CRE.

The β -galactosidase filter lift assay, shown in figure 2.9, demonstrates that the expression of the 19nt MS2 hairpin was not enough to release the repression of transcription from HIS3/ β -galactosidase promoter when it was expressed in conjunction with the viral proteins. The β -galactosidase filter lift assay did however detect a positive interaction between 3CD^{H40A} and the ^{MS2}cloverleaf. Figure 2.9 shows that no interaction was observed between any of the other proteins and the cloverleaf.



Number	Protein	RNA	Result
1	3CD ^{H40A}	MS2Cloverleaf	+
2	3CD ^{H40A}	MS2	-
3	3AB	MS2Cloverleaf	-
4	2C	MS2Cloverleaf	-
5	3D	MS2Cloverleaf	-
6	AD	MS2Cloverleaf	-
7	AD	MS2	-
8	3CD ^{H40A}	MS2	-

Figure 2.9 . Qualitative β -Galactosidase filter assay. The transformed *S.cerevisiae* were selected for by growth on media lacking leucine and uracil.



Viral protein/RNA interaction

Figure 2.10. Interactions of 3CD^{H40A} as determined by liquid β -galactosidase assay.

Each interaction was tested for β -galactosidase production in triplicate. Units of β -galactosidase were calculated as a mean of the three independent reactions. Error bars were calculated to show the range of β -galactosidase units produced between the independent samples.

To investigate the strength of the interaction between 3CD^{H40A} and the ^{MS2}cloverleaf a liquid β -galactosidase assay was carried out using ONPG as the substrate. Because of the quantitative nature of liquid assays they can be used to compare the relative strength of the interactions, though Estojak *et al* demonstrated that in the yeast two-hybrid system there was no direct correlation between β -galactosidase and the k_d of an interaction (Estojak *et al.*, 1995). All liquid assays were carried out using *O*-nitrophenolgalactosidase (ONPG) as the substrate. The strength of the interactions was calculated using the Miller equation (Miller, 1972).

To confirm that the yeast three-hybrid could discriminate clearly between the known interaction and additional RNAs, the liquid ONPG assay was carried out using *S.cerevisiae* that had been transformed with pACTII3CD^{H40A} and cDNA expressing one of the ^{MS2}cloverleaf or ^{MS2}CRE. The results of the liquid β -galactosidase are displayed graphically in figure 2.10. These results show that the positive reaction is 12 times stronger than the negative controls.

2.5 Background activation of the β -galactosidase promoter

3-aminotriazole (3-AT) is a competitive inhibitor of the HIS3 gene product. To diminish the number of false positives obtained during the library screen, concentrations of 3-AT were titrated to reduce the background whilst permitting “real” positives to grow. Although typically 3-AT is used at a concentration of 2-5mM it has previously been demonstrated that strong interactions can be detected on 25mM 3-AT (Bernstein *et al.*, 2002, SenGupta *et al.*, 1996).

To investigate the level of background activation of the reporter genes that occurred, a titration was carried out to detect the ability of yeast transformed with the fusion protein expression vector and the “empty” RNA vector pIII/MS2-1 to grow on media containing 3-AT. The growth of yeast over a range of 0mM to 30mM 3-AT was assessed. The results are displayed in table 2.3.

The ability of the yeast expressing the viral proteins to activate the HIS3 reporter gene non-specifically correlates with the resistance of *S.cerevisiae* to 3-AT. As a negative control the GAL4 activation domain (AD) and MS2 RNA were co-expressed in *S.cerevisiae*. As documented in table 2.3 no growth of *S.cerevisiae* transformed with the GAL4 AD and the MS2 RNA hairpin was observed at a concentration of 3 mM 3-AT or higher.

Table 2.3 shows whilst yeast expressing a known positive, 3CD^{H40A} and MS2Cloverleaf, were able to grow on agar containing a concentration of 3-AT up to 30 mM, no other tested interaction was able to sustain growth successfully above a concentration of 10 mM 3-AT. In addition the resistance to the presence of 3-AT in the growth medium, as displayed in table 2.3, correlates well with the known lack of RNA-binding specificity of the viral proteins 3AB and 3D^{pol} (Pata et al., 1995, Xiang et al., 1995a).

On the basis of the results of the background activation test a concentration of 5 mM 3-AT was used for cDNA library screens involving 3CD^{H40A} and 3D^{pol} and a concentration of 7mM 3-AT was used in cDNA library screens using 3AB. In the absence of any selection, expression of 2C in *S.cerevisiae* is toxic and as a consequence the ability of *S.cerevisiae* to grow over a range of 3-AT concentrations was not investigated. When undertaking cDNA library screens using 2C a concentration of 3 mM 3-AT was used to select for positive interactions.

Protein-RNA interaction	Concentration of 3-AT(mM)				
	0	3	10	20	30
3CD ^{H40A} /MS2 cloverleaf	+++	++	++	+	+
GAL4AD/MS2	+++	-	-	-	-
3CD ^{H40A} /MS2	+++	+	-	-	-
3AB/MS2	+++	+	+	+/-	-
3D ^{p01} /MS2	+++	++	-	-	-

Table 2.3 Sensitivity of “false positives” to the inhibitor 3-AT. Transformed *S.cerevisiae* were selected for by growth on leucine and histidine deficient media. Growth was marked according to density (on a scale of + to +++). No growth was marked with a dash.

2.6 cDNA library screen

Initially problems were encountered in obtaining a transformation efficiency high enough (10^5 - 10^6 cfu/ml) to generate a truly representative cDNA library screen. However altering the method of transformation of the expression plasmids into *S.cerevisiae* from a method of co-transformation to one of sequential transformation did result in a significant increase in the transformation efficiency of the cDNA library from 10^3 to $10^5/10^6$ cfu/ml. Viability of the transformants as tested by growth on leucine-deficient media was calculated as being approximately 1×10^8 cfu/ml for cells expressing 3CD^{H40A}, 3D^{pol} and 3AB and 1×10^7 cfu/ml for cells expressing 2C.

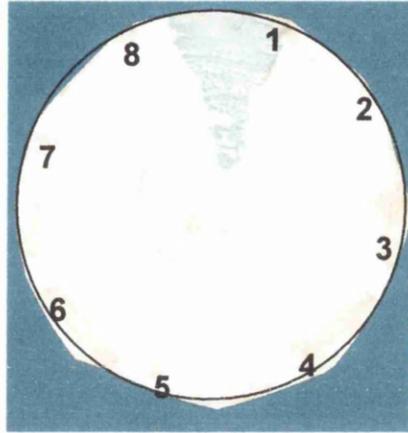
Table 2.4 gives a list of the transformation efficiencies obtained for the cDNA screens as determined by growth on uracil-deficient media. The transformation efficiencies given in table 2.4 are approximations based on colony counting. The potential for colonies to merge during the incubation period does attach a margin of error to the calculated transformation efficiencies. Table 2.4 shows that the transformation efficiency for cDNA library screens using the proteins 3CD^{H40A}, 3D^{pol} and 3AB ranged from 4×10^4 to 2×10^6 cfu/ml. In contrast, the transformation efficiency for cells expressing 2C was low, ranging between 1 and 3×10^4 cfu/ml, presumably the low transformation efficiency was due to the toxic effects that 2C expression has on *S. cerevisiae*.

Following transformation of *S.cerevisiae* expressing the AD fusions, with the cDNA library, the *S.cerevisiae* were plated out on a media (Leu- His-) containing 3-AT, at pre-determined levels, that would select for protein-RNA interactions. The plates were incubated at 28 °C for seven days after which white colonies were picked and plated onto fresh selective media containing 3-AT. The colonies were allocated a number and the growth and colour of the colonies after 3-4 days was assessed. At this stage the colonies could be split into two groups: the putative positives and RNA-independent colonies.

RNA-independent colonies are colonies of yeast that can grow well in the absence of histidine, the selection for interactions, but which have lost the cDNA that expresses the hybrid RNA. RNA-independent colonies, as has already been described, can be identified by a change in the colour phenotype of the cells. Yeast cells that have maintained the RNA plasmid are white in colour while RNA-independent colonies are wholly pink/red in colour or contain pink segments.

A common observation of the cDNA screens conducted was that a significant drop of the number of white colonies would be seen following replating. The most plausible explanation for this given that many of these were small white colonies is that in areas where the 3-AT has been degraded, or is lower, the *S.cerevisiae* can begin to grow through.

The remaining white colonies that had come through screening were streaked onto leucine and uracil (Leu⁻Ura⁻) deficient media. A few more colonies did not survive the plating onto Leu⁻ Ura⁻ media suggesting that the colour-selection for RNA-independent activation of the reporter gene is not absolute. Table 2.4 shows that β -galactosidase activity was not detectable by quantitative analysis for the colonies that came through the screening procedure for the library versus the viral proteins 2C, 3AB and 3D^{pol}. Figure 2.11 shows the result of one of the qualitative β -galactosidase filter-lift assays carried out over the course of the cDNA library screen.



Number	Protein	RNA	Result
1	3CD ^{H40A}	MS2Cloverleaf	+
2	3CD ^{H40A}	MS2	-
3	3CD ^{H40A}	Unknown 1	-
4	3CD ^{H40A}	Unknown 2	-
5	3CD ^{H40A}	Unknown 3	-
6	3CD ^{H40A}	Unknown 4	-
7	3CD ^{H40A}	Unknown 5	-
8	3CD ^{H40A}	Unknown 6	-

Figure 2.11 Results of a qualitative β -galactosidase filter assay using *S.cerevisiae* recovered from a cDNA library screen. The transformed *S.cerevisiae* were selected for by growth on media lacking leucine and uracil.

2.7 Discussion of results

The yeast three-hybrid system (figure 2.1) has become an important tool in the study of RNA-protein interactions. Much of the work carried out using the yeast three-hybrid system has concentrated on the discovery or conformation of proteins that bind a known RNA sequence. The work described in this chapter approaches the subject from an alternate angle. In this chapter the poliovirus replication proteins were used, in the context of the yeast three-hybrid system, to try and identify RNA partners for poliovirus proteins.

As has previously been described, no RNA partners for the proteins were identified. This is intriguing as preparatory work showed that the cloverleaf *cis*-acting replication element was present within both cDNA libraries. It is unlikely that this is due to the transformation efficiency reported in this study (Table 2.4) being too low to detect a positive interaction. Table 2.5 shows data derived from the published literature. Table 2.5 shows that the transformation efficiencies reported in this study at least for the viral proteins 3D^{pol} and 3CD^{H40A} are in accordance with those published.

It is possible that the reason that no positive interactions were identified was because the fusion protein was misfolded. Although no investigation was carried out to confirm that the fusion proteins, 2C, 3AB, 3D^{pol} and 3CD^{H40A} were correctly folded, results that were obtained suggested that the viral proteins were not grossly altered in their structural form. Firstly, preparatory work showed that the construction of the fusion protein and the hybrid RNA did not alter the structures of 3CD^{H40A} and ^{MS2}cloverleaf significantly as the known protein-RNA interaction could be detected. Secondly, upon transformation of the cDNA from which the viral protein 2C was expressed the transformed *S. cerevisiae* showed a distinct phenotypic change that suggested that 2C was toxic to the yeast cells. As 2C is known to cause rearrangement of intracellular membranes it is likely that the toxicity observed in yeast cells transformed with the 2C expression plasmid is a consequence of 2C being functionally active in the context of the fusion protein.

Protein	Library	Transformants	Number of screened transformants (colony colour white after 7 days)	2 nd screen (White) Leu ⁻ His ⁺ 3-AT (growth ++/+++)	RNA – independent positives after 2 nd screen	β-galactosidase positive
3D	LibRE	1 x 10 ⁵	148	19	11	0
3D	LibDNAase	5.2 x 10 ⁴	217	38	16	0
3CD ^{H40A}	LibRE	3.2 x 10 ⁵	15	1	3	0
3CD ^{H40A}	LibDNAase	2 x 10 ⁶	65	5	19	0
3AB	LibRE	5.4 x 10 ⁴	108	1	25	0
3AB	LibDNAase	4 x 10 ⁴	64	1	13	0
2C	LibRE	2.6 x 10 ⁴	11	0	5	n/a
2C	LibDNAase	1.6 x 10 ⁴	83	0	63	n/a

Table 2.4 Screens for cDNA that interacts with poliovirus viral proteins

RNA	Transformants	White	β -galactosidase positive	RNA dependent	Positive interactors	Relevant In vivo
Ash1 3'UTR	2×10^6	66	66	11	4	2
Histone SL	3×10^5	4	4	1	1	1
HCV 3'X	3×10^5	6000	-	-	4	-
Influenza NP	5×10^5	-	120	10	2	2
Ash E3	5×10^5	49	23	9	1	1
<i>T. cruzi</i> SL	2×10^6	400	100	-	1	-

Table 2.5. Library screens for cDNAs that interact with RNAs. Table 2.5 was adapted from Bernstein, D. S., N. Buter, et al, 2002 and references therein. Abbreviations HCV: hepatitis C virus, NP: Nucleoprotein, UTR: Untranslated region

No interaction between 3AB or 3D^{pol} and the viral RNA has been as well characterised as the interaction that occurs between 3CD^{H40A} and the cloverleaf. In addition the expression of 3AB and 3D^{pol} by *S.cerevisiae* was not accompanied with phenotypic changes, consequently determining that these proteins are functional in the context of the fusion protein is difficult. Direct evidence for the correct functioning of the 3D^{pol}-fusion could be tested, by assessing the ability of the purified fusion protein to catalyse the uridylylation of VPg and the polymerisation ability of the polymerase using the documented *in vitro* assays. However, this requires the expression and purification of large quantities of the protein in yeast. Direct evidence for the correct folding of the multifunctional protein 3AB within the context of the fusion protein would be easier to obtain. Function could be determined by confirming its known association with cellular membranes and interactions with 3CD^{pro} and 3D^{pol}. Additionally, the ability of 3AB to alter membrane permeability could be exploited to determine the functionality of the 3AB-fusion. None of these approaches were followed as the results of the background activation test using 3-AT provided evidence that the 3D^{pol} and 3AB-fusion proteins retained RNA-binding activity. Non-specific RNA-binding activity is a known characteristic of both 3D^{pol} and 3AB and so this provides indirect evidence that both these proteins are folded into a structurally functional conformation. The western blot analysis carried out to confirm expression of the fusion proteins in *S. cerevisiae* provided further indirect evidence that 3AB was folded in the correct conformation as it clearly suggested the 3AB-fusion was capable of forming a higher order complex, a characteristic associated with functional 3AB.

The identification of RNA binding activity for 3AB, 3D^{pol} and 3CD^{H40A} and the toxic effect of 2C suggests that the reason that no RNA binding partners were identified was not due to the abolition of function of the proteins, due to a gross structural change. One consequence of expressing the viral proteins as a fusion is that the fusion GAL4 activation domain may be positioned structurally to block the RNA binding domain of the protein partially or completely. Figures 2.9 and 2.10 show that the RNA binding domain of 3CD^{H40A} is not blocked as the interaction between 3CD^{H40A} and the ^{MS2}cloverleaf can be detected indirectly through assaying for β -galactosidase activity. No positive test could be carried out for the other proteins though the titration of 3-AT showed that 3AB and 3D^{pol} showed non-specific RNA binding activity.

One possible limitation of the yeast three-hybrid system arises from the observation that optimal reporter gene activity is obtained when the hybrid RNA contains inserts of

between 50 and 200 nucleotides in length. Table 2.6 displays the length of RNA sequences known to interact with proteins. Table 2.6 shows that the majority of identified RNA binding sites fall within the optimal range of the yeast three-hybrid system. The size limitations on the insert making up the hybrid RNA does preclude the identification of protein binding sites that are formed by the tertiary structure of the RNA molecule of interest. For example, in this study the construction of the cDNA library used a method of partial restriction enzyme digestion to obtain overlapping fragments of between 50 and 200 nucleotides. One drawback to this approach is that using short overlapping fragments requires the RNA sequence that makes up the protein-binding site to be linear in nature. An additional problem that may be encountered is that the removal of flanking sequences may alter the most thermodynamically favoured secondary RNA structure for the sequence of interest.

It is therefore possible that the “positive” cognate RNA binding partners of the viral protein are misfolded in the context of the hybrid RNA. Misfolding of this sort, illustrated in figure 2.12 has been described in other studies (Cassiday & Maher III, 2001, Edwards *et al.*, 2001) and can prevent interaction of the RNA with both the bacteriophage MS2 coat protein and/or the protein of interest. Edwards *et al* in 2001 overcame this problem by inserting a G-C clamp into the hybrid RNA (Figure 2.12).

Although the mis-folding of RNA provides a rational explanation for why no positive partners were identified in the cDNA screen, engineering a G-C clamp into pIII/MS2-1 may not offer a solution to the problem. In the situation where a known RNA structure is being tested for an interaction with an unknown protein partner(s), the RNA can be manipulated by the insertion of a G-C clamp to form a secondary structure that has been described by other techniques. In contrast, where the RNA provides the unknown variable, insertion of a G-C clamp may create structures that are not present in the viral RNA naturally, or may destroy secondary structures that are found. While positive interactions may be obtained from library-screens carried out using RNA that has been structurally “fixed” by the introduction of a G-C clamp, the presence of the G-C clamp must be accounted for when describing and elucidating the functional relevance of any identified RNA sequences.

One other factor that may play a part in explaining the results obtained in this study is the orientation of the RNA relative to the MS2 coat protein binding sites. The orientation of

RNA	Length of RNA (nt) excluding MS2	Cellular or Viral Origin	Protein	Reference
ash1 E3	127	Cellular	She3p	1
NRE	57	Cellular	Pumilio	2
Histone RNA hairpin	28	Cellular	HBP(SLBP)	3
IRE	51	Cellular	IRP-1	4
bI4	1600	Cellular	Leu RS bI4 maturase	5
TAR	58	Viral	Tat	4
HIVy	139	Viral	HIV Gag	6
RSVMY	320	Viral	RSV Gag	7
HCV SL3.2	48	Viral		8
HCV 3'X	98	Viral	L22,L3,S3 mL3	9
Poliovirus cloverleaf	100	Viral	3CD PCBP 2	10, 11
Poliovirus CRE	60	Viral		12

Table 2.6. Known RNA-protein interactions in both cellular and viral systems.

References: 1(Long et al., 2000), 2 (Sonoda & Wharton, 1999), 3 (Martin et al., 2000), 4 (SenGupta et al., 1996) , 5 (Rho & Martinis, 2000) 6, (Bacharach & Goff, 1998), 7 (Lee et al., 1999), 8 (You et al., 2004), 9 (Wood et al., 2001), 10 (Andino et al., 1990a), 11 (Gamarnik & Andino, 1998) and 12 (Goodfellow et al., 2000a).

the RNA sequence and the MS2 sites can affect the strength of the reporter gene signal. Insertion of the iron response element (IRE) upstream of the MS2 sites has been shown to increase transcription of the hybrid RNA two to three-fold above the level of transcription observed when the IRE is positioned upstream of the MS2 sites (Bernstein et al., 2002). However Bernstein *et al* have reported that systematic studies carried out to test this feature of the yeast three-hybrid system have reported that both orientations yield specific activation of the reporter gene by the RNA-protein interaction. This is supported by the observation that published literature on the yeast three-hybrid system has used both the downstream orientation used in this study and the upstream orientation as used by Bernstein *et al*. Therefore it is unlikely that the sole reason no positive RNA-protein interactions were detected was due to the orientation of the RNA to the MS2.

It is however still possible that the orientation of the RNA sequence of interest with respect to the MS2 sites remains a factor. One hypothesis that cannot be ruled out is that the binding of the MS2 coat protein to the MS2 binding sites results in the covering of the viral protein binding site on the RNA of interest (illustrated in figure 2.13). Construction of a second cDNA library using the upstream positioning of the RNA with respect to the MS2 coat protein binding sites and repeating the yeast three-hybrid screen would possibly confirm whether this is indeed the case.

One further possibility to be explored is the effect of reducing the level of the His3p competitive inhibitor 3-AT used to suppress “false positives” within the screen. Although the levels of 3-AT used within the cDNA library screen fall within the ranges described in other studies (Bernstein et al., 2002), it may be that the levels were too stringent for the detection of weak or transient specific interactions between the viral proteins and RNA. Repetition of the library screen in the presence of a lower level of 3-AT, in the range of 1-2.5 mM 3-AT might generate positive results, despite a relative increase in the background of “false” positives.

Even if the hybrid RNA and proteins are correctly folded the straightforward yeast three-hybrid system detailed in figure 2.1 does not account for the modulation of protein function by protein-protein and multiple protein-RNA interactions. Four types of multi-protein interactions have been identified. These are independent interactions, bridged interactions, coupled interactions and complex interactions. These are illustrated in figure 2.14.

Independent interactions, as illustrated in figure 2.14A, occur when two proteins independently bind to two sites on the same RNA. The binding sites used by the two proteins are separate and the binding of the proteins is unaffected by the binding of the other protein. An example of an independent interaction used to mediate the function of the RNA is the interaction of the bI4 maturase and the leucyl-t RNA synthetase with the *cob* bI4 intron (Rho & Martinis, 2000).

Bridged interactions (Figure 2.14B) occur when two proteins are required to mediate binding to a single RNA site. Long *et al* showed in 2000 that the apparent interaction of She3p with the ASH1 RNA only occurred in the presence of She2p (Long et al., 2000).

Coupled interactions, as shown diagrammatically in figure 2.14C, occur when two proteins interact with each other and with the RNA. Examples of coupled interaction include the interaction that occurs between the HIV-1 Tat protein, the cellular protein CyclinT1 and HIV TAR RNA (Bieniasz et al., 1999, Bieniasz et al., 1998). The formation of a coupled interaction may occur in two different ways. Firstly formation of the coupled interaction may require an initial interaction between one of the proteins involved in the complex interaction and the RNA to facilitate the formation of a protein-protein interaction with the second protein in the setup. On the other hand prior formation of the protein-protein interaction may be required before interaction with the RNA by the proteins can occur. Bieniasz *et al* have shown that the interaction which occurs between HIV-1 Tat, TAR and Cyclin T1 occurs according to the latter procedure (Bieniasz et al., 1999).

Complex interactions i.e. where more than two proteins are involved in a series of RNA-protein and protein-protein interactions (Figure 2.14D) are also known to exist (Sonoda & Wharton, 2001).

In each of the examples above the multiple protein interaction are known to be functionally important. In addition to this, the interactions have all been described through modification of the basic yeast three-hybrid system. The condensed nature of the genomes of RNA viruses, in particular the picornaviruses, requires that the information carrying capacity of the genome is maximised. One method that the picornaviruses have used to achieve this is that the number of proteins coded for by the viral ORF is increased through the use of a proteolytic cascade (Figure 1.14), which generates functional intermediate proteins. Taking the P3 region of the poliovirus ORF as an example the viral intermediate proteins 3AB

Figure 2.12 *A schematic diagram of the secondary structure of hybrid RNA molecules with and without the use of a G-C clamp.*

The RNA sequence is a short stem-loop X indicated by the thick blue line. The natural secondary structure context of stem-loop X is shown in the inset. Thin black lines correspond to the RNAase P leader and trailer sequences. The thick black lines correspond to the MS2 coat protein binding sites. **A** shows the predicted secondary structure of the hybrid RNA expressed in the absence of stem-loop X. **B** The secondary structure of the hybrid RNA formed following the insertion of stem-loop x – a structure in which neither the MS2 coat protein binding sites or stem-loop x is folded in the functional conformation. **C** The introduction of a G-C clamp (boxes plus black lines) restores the conformation of both stem-loop x and the MS2 coat protein binding sites. Diagram is adapted from Bernstein et al., 2002

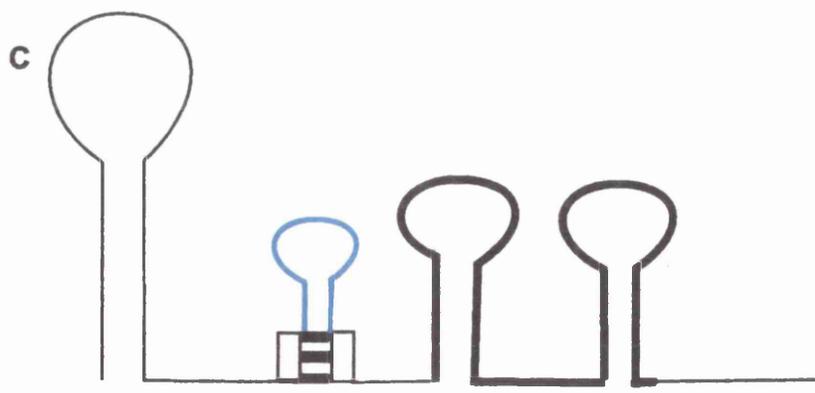
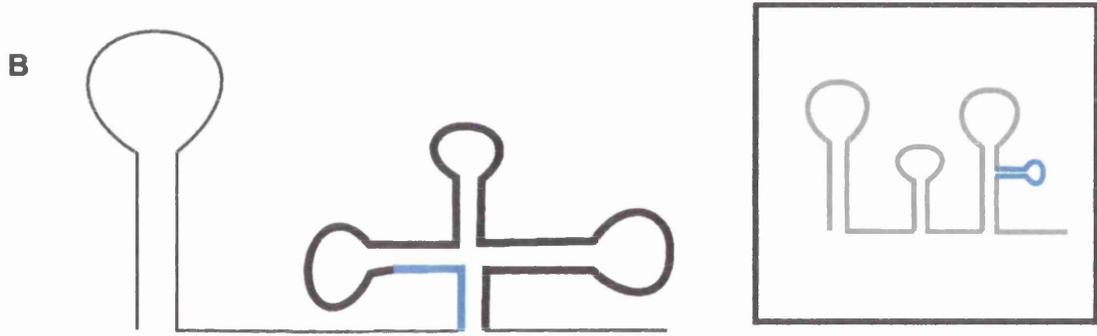
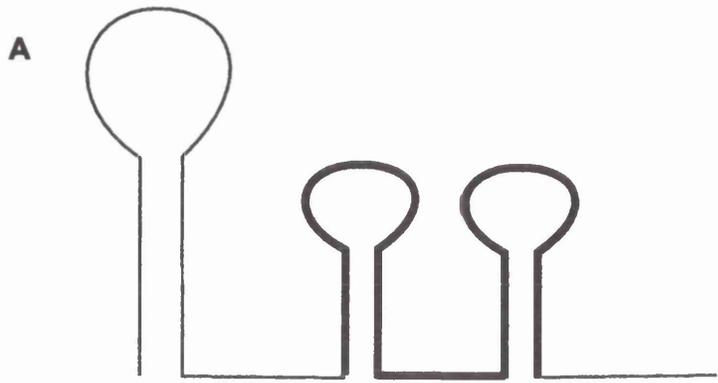
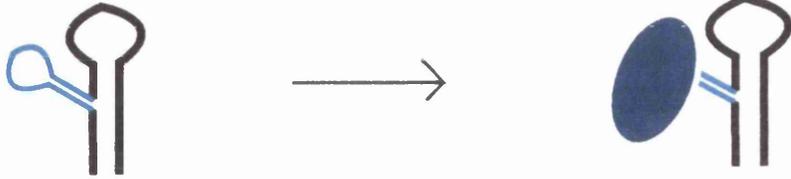


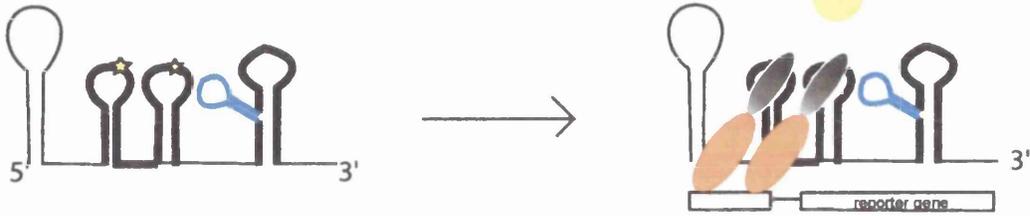
Figure 2.13 Schematic diagram illustrating how altering the orientation of the RNA to the MS2 binding sites can alter the interaction of the protein with the RNA transcript

(A) The protein binding site (bold blue line) is shown in its natural position and with its cognate binding protein (blue oval) bound. When the RNA of interest is positioned downstream of the MS2 coat protein recognition sites (Bold black lines) binding of the MS2 coat protein (Grey oval) prevents the protein of interest from recognising the presence of its binding site **(B)**. As a consequence no interaction occurs between the transcriptional DNA binding domain (orange oval) and the transcriptional activation domain (yellow sphere). When the MS2 coat recognition site are positioned upstream of the RNA sequence of interest **(C)** the binding of the MS2 coat protein does not interfere with the binding of the protein to its cognate RNA binding site. As a result the DNA binding domain and the transcriptional activation domain can interact thus activating transcription of the reporter gene.

A



B



C



KEY :

Component	Definition	Component	Definition
	Mottf A cognate binding protein		Bacteriophage MS2 coat protein
	Transcription factor Activation domain		RNA motif A
	Transcription factor DNA binding domain		

(RNA-binding protein involved in polymerisation and precursor to VPg) and 3CD^{pro} (sequence-specific RNA binding activity and proteinase activity) are known to have functions that differ from those of the final cleavage products that include a proteinase (3C^{pro}), evasion of the immune system (3A), the protein primer required for the initiation of RNA synthesis (3B=VPg) and the viral polymerase (3D^{pol}).

A further way of achieving maximum functional capacity from a small genome is for the encoded proteins to modulate the function of each other and cellular proteins by the use of protein-protein interactions. This is certainly the case for poliovirus. The documented interaction between 3AB and 3CD^{pro} results in the stimulation of the autocatalytic cleavage of 3CD^{pro} that is a fairly stable intermediate protein (Molla et al., 1994, Xiang et al., 1995a). The poliovirus intermediate protein 3AB has also been shown to positively modulate the polymerisation activity of the 3D^{pol} (Lama et al., 1994, Paul et al., 1994a, Plotch & Palant, 1995). Figure 2.15 is a summary of the interactions and known functions of the proteins encoded by the P3 region of poliovirus (Xiang et al., 1998). The picture is further complicated intracellularly by the presence of cellular proteins that may modulate the viral proteins. An interaction between 3D^{pol} and Sam68 has been previously described in the literature (McBride et al., 1996) although as yet no function has been ascribed to this interaction. Other interactions may exist between the viral proteins and cellular proteins that have not yet been identified.

It is therefore possible that the reason for the failure to detect specific RNA binding sites for the proteins screened in this study resides in the fact that a coupled interaction occurs following an interaction between two proteins. Studies on known RNA-protein interactions have overcome this problem by modifying the basic yeast three-hybrid system through the introduction of one or more additional cDNAs into the *S.cerevisiae* R40coat strain (Bieniasz et al., 1999, Sonoda & Wharton, 2001). Extension of this method to the poliovirus screen described in this study would for example express 3AB from a multicopy plasmid as a non-fusion protein in the presence of 3D^{pol} or 3CD^{pro}. This type of multi-protein cDNA library screen could be extended to search for cellular protein/viral protein complexes through the introduction of a cDNA library expressing cellular proteins as fusions and the viral proteins expressed as non-fusions.

One further modification of the yeast three-hybrid system has been described that enables screening for proteins that are involved in complex interactions. Sonoda *et al* described in

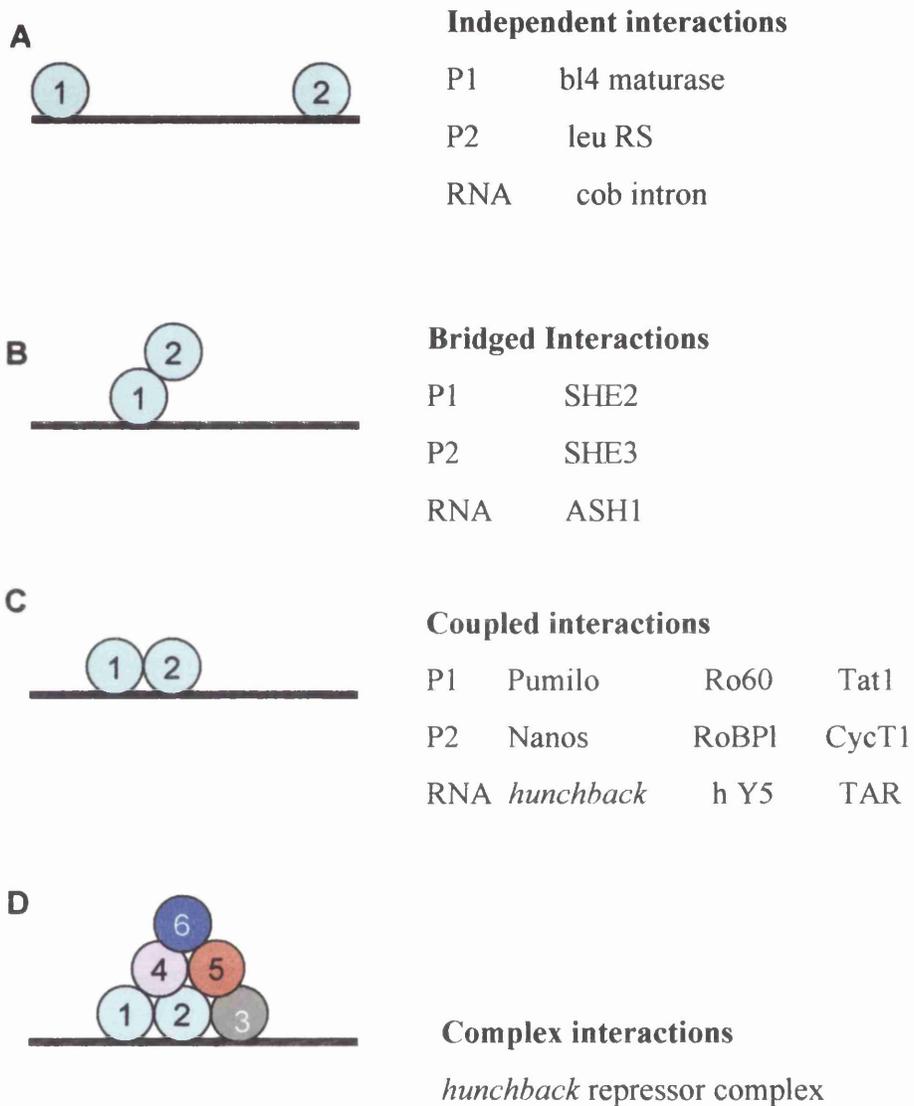


Figure 2.14 *Analysing multicomponent interactions*

A diagram of four permutations of complexes involving one RNA and two or more proteins. Each type of permutation has been analysed in the yeast three-hybrid system. Adapted from Bernstein et al Methods 2002. Black bold line corresponds to the RNA 1 and 2 correspond to proteins 1 and 2. Examples of each type of multicomponent interaction are shown adjacent to the relevant diagram. Relevant citations are **A** Rho and Martinas., 2000 **B** Long *et al.*, 2000 **C** Sonoda *et al.*, 1999; Bouffard *et al.*, 2000; Bieniasz *et al.*, 1998, 1999 and **D** Sonoda *et al.*, 2001

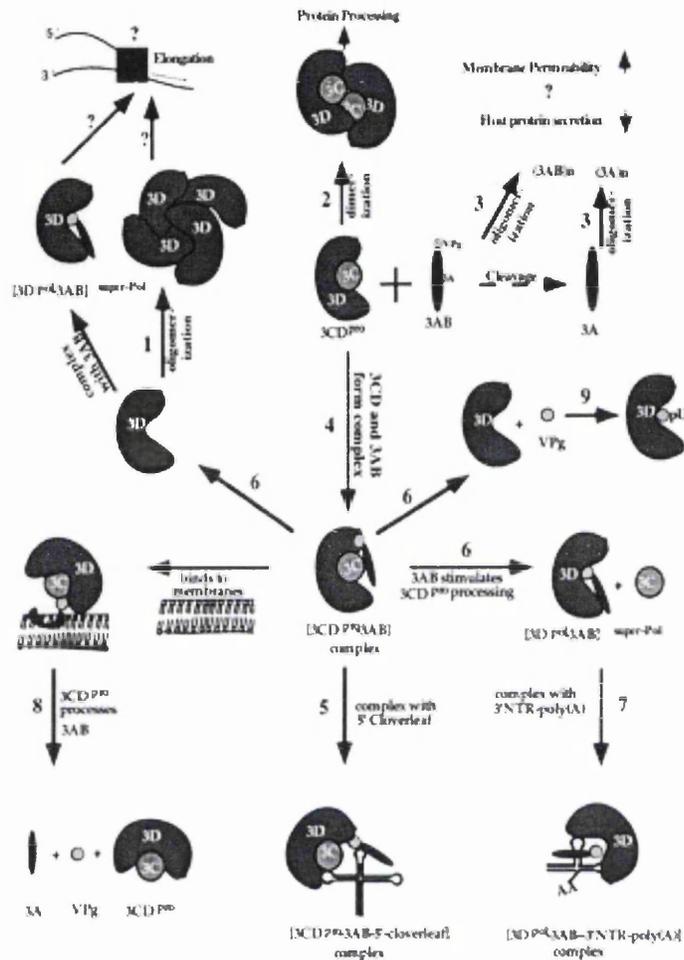


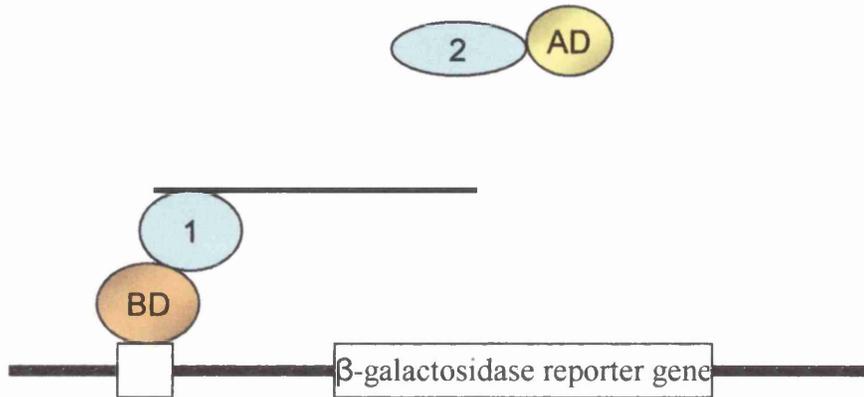
Figure 2.15 Working model for interactions and diverse functions of the P3 polypeptides (1) 3D^{pol} is depicted as a tetramer although its precise status in solution is not known. 3D^{pol} oligomers are more active than monomers in RNA chain elongation, and the same may be true for the 3AB/3D^{pol} complex. (2) The function of 3CD^{pro}/3CD^{pro} dimers is not known, but they may be required for autocleavage and/or viral protein processing. (3) 3AB and its cleavage product 3A alter host membrane permeability and inhibit protein secretion. Whether these proteins must assemble to dimers or multimers (through interactions of their hydrophobic domain) to produce these effects is unknown. (4) 3AB and 3CD^{pro} can form a complex in solution that (5) can bind with high specificity to the 5' cloverleaf, a process required for the initiation of positive-strand RNA synthesis. (6) When 3AB and 3CD^{pro} are bound to each other, 3AB can stimulate the autocleavage of 3CD^{pro} into 3C and 3D^{pol}. 3D^{pol} and 3AB may remain in a tight complex with increased RNA polymerization activity. (7) A 3AB/3D^{pol} complex is likely to recognize the 3' NTR of the genome for the initiation of negative-strand RNA synthesis. (8) The half-life of 3AB is regulated by 3CD^{pro}, since membrane-bound 3AB can be cleaved by 3CD^{pro} into 3A and 3B (VPg), two cleavage products whose functions are distinct from those of the 3AB precursor. (9) VPg, in turn, is uridylylated by 3D^{pol} to VPgpU(pU), the primer for 3D^{pol}. Taken from Xiang et al, 1998

1999 the curiosity of the interaction between *hunchback* RNA and the cellular protein Nanos. Sonoda *et al* demonstrated that while Nanos would interact with the *hunchback* RNA when it was tethered by Pumilio, no interaction was observed when the hybrid RNA was tethered by the MS2coat protein/MS2 hairpin linkage (Sonoda & Wharton, 1999). A similar phenotype was observed by Bouffard *et al* in 2000 when investigation protein interactions with hY5 RNA (Bouffard et al., 2000). The strategy employed to identify protein-RNA interactions like these is summarized in figure 2.16A.

The principle of the modified yeast three-hybrid system outlined in figure 2.16A is that the linkage provided by the interaction between the MS2 hairpin and the MS2 coat protein/DNA binding domain fusion can be functionally replaced by any known RNA/protein interaction. Bouffard *et al* chose to fuse the protein Ro60 with the DNA binding domain. The known interaction between Ro60 and hY5 RNA was then used as bait in a library screen of cellular proteins fused to the activation domain. Sonoda *et al* expanded the technique further by the introduction of a third non-fusion protein into the equation. This is illustrated in figure 2.16B.

Given what is already known about the interactions that occur between the proteins encoded by the P3 region of the viral polyprotein the use of complex interactions such as those described above is likely. It is already known that in poliovirus the RNA specificity of the intermediate protein 3CD^{pro} resides in 3C^{pro} rather in the viral polymerase 3D^{pol} component of the protein (Andino et al., 1993, Andino et al., 1990c). Therefore other protein-protein interactions that occur during the course of the viral replication cycle may be responsible for mediating the interaction between proteins and genome. Modification of the yeast three-hybrid though not attempted during the course of this study does provide an interesting avenue of future work.

A



B

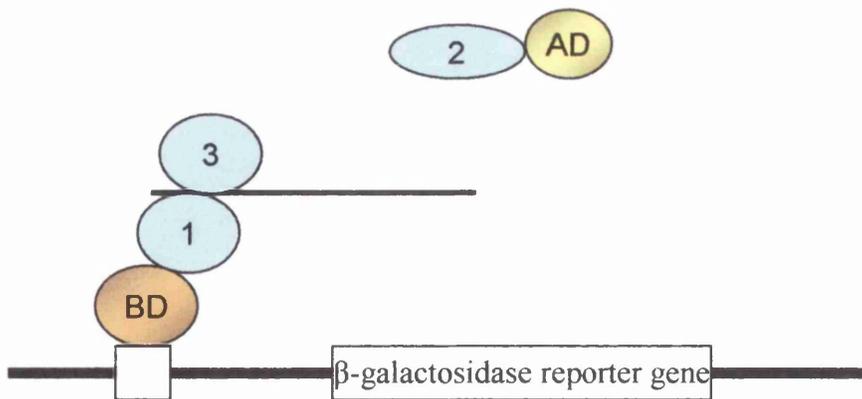


Figure 2.16. Screens to identify proteins from multicomponent complexes

(A) a screen for protein-RNA complexes. The DNA binding domain is fused to an RNA-binding domain of choice (1) which interacts with the RNA. The RNA/RNA binding domain (BD) complex is then used as bait for an unknown protein (2) fused to the activation domain (AD). The basic principle was modified further by Sonoda et al in 2001 (Sonoda *et al*, 2001) by the introduction of non-hybrid component (3) to form a ternary RNA complex that is used as bait to screen unknown AD fusions.

2.8 Future work

One of the major concerns identified during the course of this study has been the inability of the yeast three-hybrid system to identify an interaction between 3CD^{H40A} and the ^{MS2}cloverleaf during the cDNA library screen. The use of PCR did however show that the cloverleaf element was present, in its entirety, within the cDNA library (Figure 2.6A).

One experiment that should be carried out is to identify whether the cloverleaf present within the cDNA library retains 3CD^{H40A} binding activity. To determine this, cDNAs expressing the ^{MS2}cloverleaf RNA should be retrieved from the library using colony hybridization (Sambrook et al., 1989). In this technique, an imprint of bacterial colonies transformed with the cDNA library is made from an agar plate onto nitrocellulose membrane. Lysing of the bacterial cells, imprinted on the nitrocellulose membrane, leaves the DNA extracted from each cell immobilised on the membrane at the precise location that the cell was imprinted. The immobilised DNA can be subsequently probed for using radio-labelled oligonucleotides or RNA that are anti-sense with respect to the target-sequence. The presence of the target-sequence can be visualised by exposing the nitrocellulose membrane to X-ray (autoradiography) film. By orienting the nitrocellulose membrane correctly to the agar plate bacterial colonies corresponding to signals positive for the target-sequence on the nitrocellulose membrane can then be picked and amplified.

Isolation of the plasmids expressing ^{MS2}cloverleaf RNA would enable a number of questions to be answered. The ability of the individually expressed ^{MS2}cloverleaf RNAs to interact with 3CD^{H40A} independently of the other cDNAs could be monitored by using the quantitative and qualitative β -galactosidase assays described within this chapter.

Sequencing of the retrieved cloverleaf cDNAs will identify if any mutations have been accumulated that might block or reduce the binding of 3CD^{H40A}. Lastly, the retrieval of cDNA clones expressing the ^{MS2}cloverleaf RNA using the southern blotting technique will enable determination of the number of copies of the cloverleaf present within the cDNA libraries. Sequencing of the retrieved cDNAs would also enable determination of whether the majority of the ^{MS2}cloverleaf RNA expressed in *S. cerevisiae* is functional (sense-strand) or non-functional (anti-sense strand) with respect to 3CD^{H40A} binding. If the majority of the ^{MS2}cloverleaf RNA is expressed as sense-strand than the observation that it could not be retrieved using 3CD^{H40A} in the yeast three-hybrid cDNA library screens raises

questions regarding the sensitivity of the yeast three-hybrid screen in comparison with PCR. The question of the sensitivity of the yeast three-hybrid system can be answered using a sensitivity assay. Basically, a defined copy number of cDNA will be used per transformation. Each yeast three-hybrid screen will be carried out using a dilution of pIII^A/MS² cloverleaf as part of the total copy number used. The remainder of the cDNA used to achieve the total copy number used per transformation will be either pIII^A/MS²-1 or pIII^A/MS²CRE. PCR will be carried out on the samples at the concentrations used in the yeast three-hybrid screen both in the presence and absence of cDNA encoding different sequences. Questions that should be answered by the results of this sensitivity assay include:

- How many copies are required to be present in the library before the cloverleaf can be successfully picked out of the cDNA yeast three-hybrid library screen?
- Using PCR what is the minimum copy number of the cDNA pIII^A/MS² cloverleaf that results in a positive identification?
- Does the presence of additional cDNAs encoding different RNA structures decrease the sensitivity of both PCR and the yeast three-hybrid system?

The results accumulated in the sensitivity assay may generate an explanation for the failure of the yeast three-hybrid screen outlined in this chapter to identify novel RNA-protein interactions. Modifying the conditions of the yeast three-hybrid to take account of the results of the sensitivity assay results may lead to the identification of other RNA-protein interactions. If not then modification of the yeast three-hybrid system to account for multi-component interactions may provide a solution to the problems encountered in this study.

3 Analysis of the functional significance of a region of suppression of RNA sequence variation in poliovirus

3.1 Introduction

The functional importance of RNA secondary structure in the replication of RNA viruses has long been established. RNA elements have been described for both mammalian and prokaryotic viruses that signal functions as diverse as translation, genome replication and encapsidation. The recent use of computational methods as a way of identifying potential RNA secondary structures of functional importance has been documented for a number of RNA viruses, including hepatitis C virus (Tuplin et al., 2002) and Ebola virus (Crary et al., 2003). Computational analysis was also used to provide supplementary evidence for the existence of the CRE in poliovirus (Goodfellow et al., 2000b).

Numerous algorithms have been developed independently to aid the identification of RNA secondary structures that may be of biological significance. Traditionally algorithms have been based on thermodynamics. Thermodynamic algorithms, for example MFOLD analysis (Mathews et al., 1999, Zuker, 1989, Zuker, 2003), rely on data experimentally obtained on the free energy of stems and loop structures (Antao et al., 1991, Antao & Tinoco, 1992, Varani et al., 1991, Xia et al., 1997). The structures predicted to form using thermodynamic calculations have been demonstrated to be in good accord with data provided from experimental methods of analysis such as nuclear magnetic resonance (NMR) (Tinoco & Bustamante, 1999), chemical (Mathews et al., 1999) and enzymatic probing (Goodfellow et al., 2003a).

In contrast to thermodynamic methods of analysis, that predict the formation of structures on the basis of a single sequence, new algorithms have been developed that take advantage of comparative sequence data. Suppression of synonymous site variation (SSSV), one such algorithm, works by analysing genomic alignments for sequence constraints that are present but which can be demonstrated to be independent of the protein coding sequence. The maintenance of such sequences within a population may result from constraints placed on the genome by the presence of structures or sequences that have an essential or important biological function. Analysis of the enterovirus genus using SSSV analysis has

identified a number of regions of suppression. This chapter describes the application of molecular techniques to one of these regions to investigate whether its functional significance can be determined.

3.2 Analysis of SSSV amongst the Enteroviruses

Suppression of the sequence diversity of a region of a viral genome may result from constraints placed on the genome by the presence of structures that have a direct bearing on virus replication (Simmonds & Smith, 1999).

The level of SSSV within a genome is calculated using parsimony analysis. This method calculates the proportion of similarity between aligned sequences at each codon. In order to identify regions of suppression; variability is normalised to account for codon degeneracy. Variability is then plotted over a sliding window. In this case a sliding window of 21 codons was used. Areas of codon suppression below the baseline level are predicted to consist of secondary structures or sequences of functional importance, as the underlying RNA sequence must be constrained by factors unrelated to protein function.

Parsimony analysis carried out on human enteroviruses by Prof. P. Simmonds (University of Edinburgh) identified regions of SSSV (Figure 3.1) centred on nucleotides 3841, 5339 and 6467. The nucleotide numbering refers to the position of the nucleotide within the open reading frame (ORF) of the polyprotein i.e nucleotides 1-3 are the initiation codon AUG. No SSSV was identified in the region of the genome encoding the capsid proteins. The areas of SSSV centred on nucleotide 5339 (blue arrow) and nucleotide 6467 (purple arrow) are localised within the region encoding the viral polymerase. The marked area of SSSV at position 6467 is located within nucleotides 6768-7148 of the full-length poliovirus genome. The marked area of SSSV centred on nucleotide 3841 occurs in the region of the genome where the CRE (black arrow) is localised. The CRE was initially identified using MFOLD analysis. The identification of the CRE is important for two reasons. Firstly, it provides supporting evidence that SSSV analysis can identify structures of functional significance by its ability to identify a known element. Secondly, it informs that elements of functional importance should be predicted by more than one method of computational analysis.

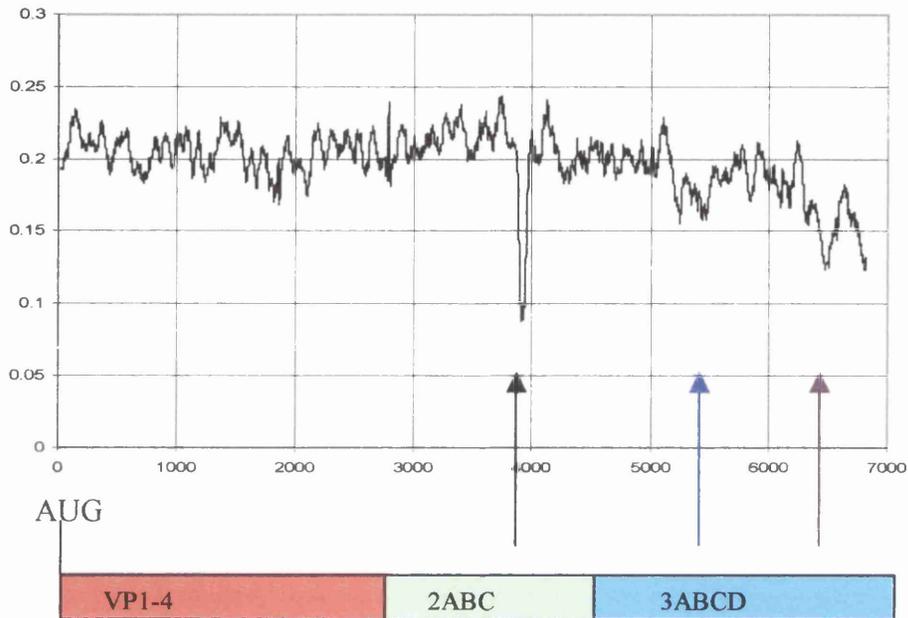


Figure 3.1. Suppression of synonymous site variation analysis of the enterovirus genus. Synonymous codon variability analysis was used to analyse an alignment of enterovirus genomes. The alignment generated by CLUSTAL used all enterovirus sequences available from GENBANK, which differed from each other by >1%. Marked areas of suppression were identified at positions 3841, 5539 and 6467. The CRE (black arrow) localises to the area of suppression at position 3841. The marked area of suppression at positions 5539 (blue arrow) and 6467 (purple arrow) localise to the region of the genome encoding the viral proteins 3C and 3D. The nucleotide numbering refers to the position of the nucleotide in the ORF i.e. nucleotides 1-3 are the initiation codon AUG.

3.3 MFOLD analysis of poliovirus type 3 Leon nt 6768-7148

One of the most commonly used computational algorithms for predicting RNA secondary structure is the MFOLD programme (Zuker, 1989, Zuker, 2003), that calculates the free folding energy of a given RNA structure. The algorithm works by calculating the pairing-number (P-num) value that assesses the likelihood of any one base pairing with any other. A high P-num value indicates that a base is promiscuous in its interaction with other bases during folding for bases with low P-num values the converse is true. The P-num value calculated for any one sequence can be presented either as a structure format plot or a P-num plot.

MFOLD analysis carried out of the region of codon suppression in the genome of poliovirus type 3/Leon, using the MFOLD web server, <http://www.bioinfo.rpi.edu/applications/mfold>, identified a highly structured region with a predicted free energy of -93.7 (Figure 3.2). The P-num values for the region are low (Figure 3.3) consistent with a region in which secondary structure would be predicted to form.

The presentation of the data as a structure format plot predicts the formation of two stem-loop structures. The first of these stem-loop structures (structure I) is 87 nucleotides in length. The predicted structure contains five bulge loops and a terminal loop of eleven nucleotides. The second of the stem-loop structures (structure II) shows a less classical stem loop structure. Structure II shows a trident like- structure with three stem-loops that protrude from two bulge loops present with the structure. The stem-loops that crown the structure each contain a terminal loop of four or five residues. MFOLD analysis of the equivalent regions within the genomes of poliovirus type 1 (Mahoney) and type 2 (Lansing) predicts the formation of similar structures (Figure 3.4).

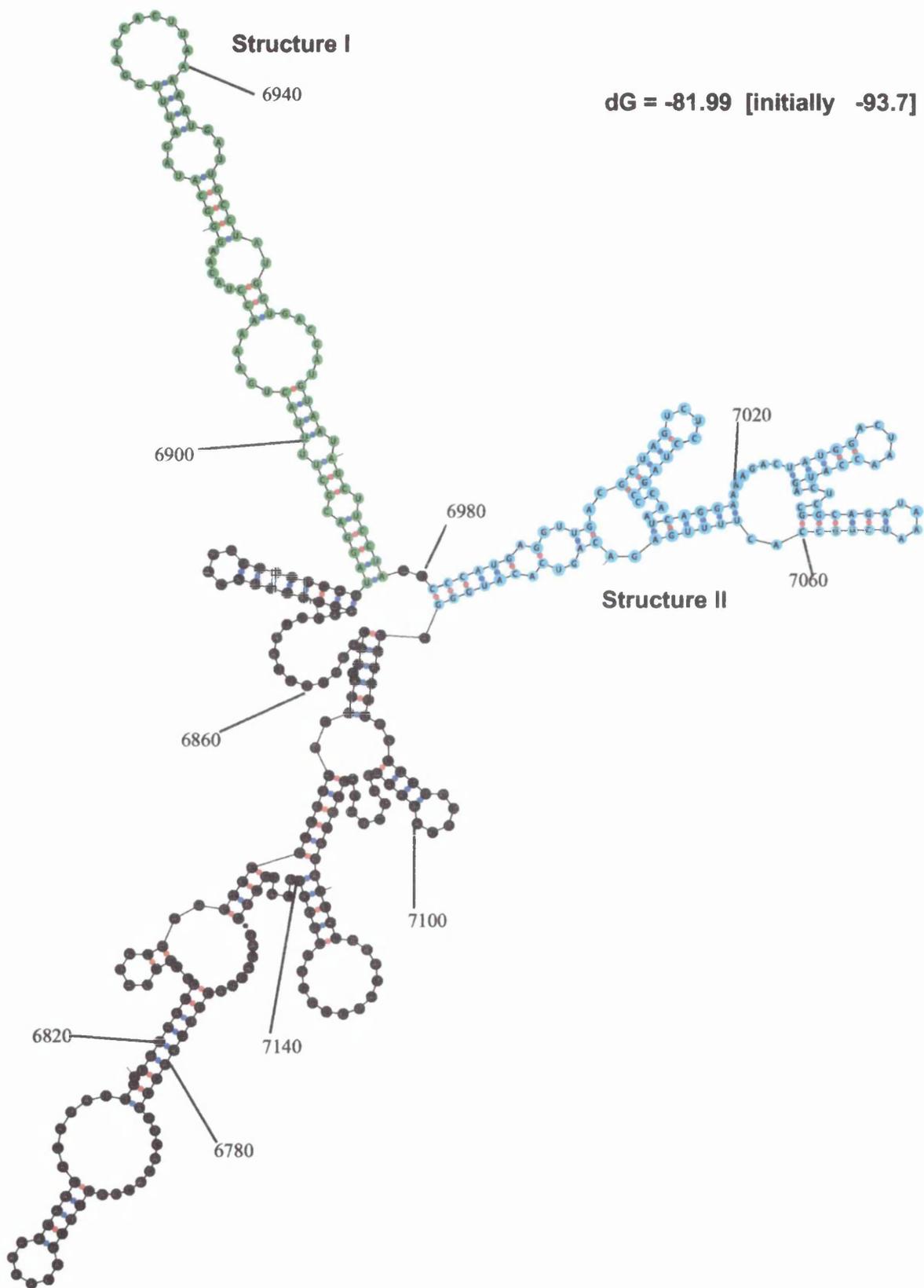


Figure 3.2 Structure format plot of nt 6768-7148 of the genome of poliovirus type 3 (Leon). MFOLD analysis was carried out using the MFOLD server. The MFOLD analysis predicted the formation of two stem-loop structures: Structure I and Structure II, which are highlighted in green and blue respectively.

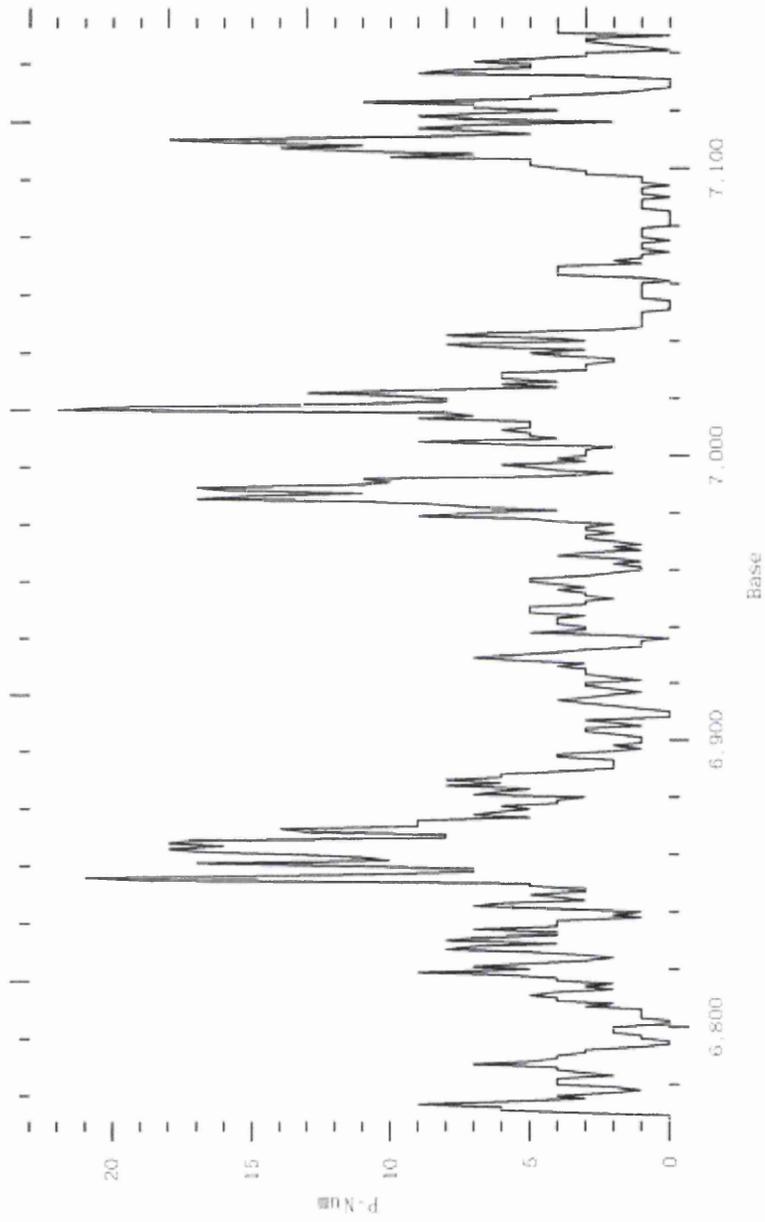


Figure 3.3 *P-Num plot for the region of the poliovirus type 3 (Leon) genome nt 6768-7148*

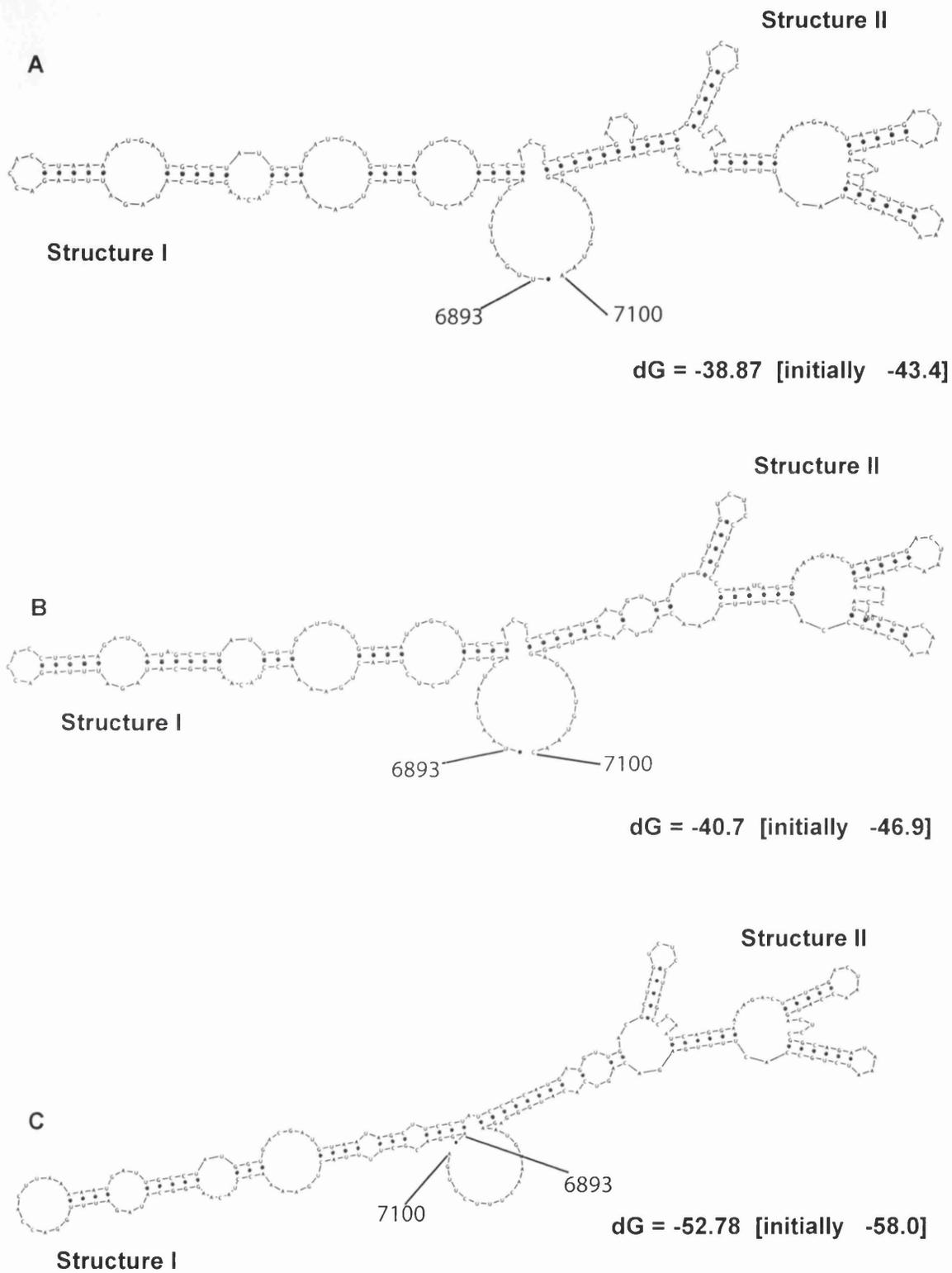


Figure 3.4 MFOLD analysis of nt 6893-7100 in poliovirus type 1-3
(A) MFOLD analysis of nt 6893-7100 of poliovirus type 1 (Mahoney),
(B) MFOLD analysis of nt 6983-7100 of poliovirus type 2 (Lansing) and
(C) MFOLD analysis of nte 6983-7100 of poliovirus type 3(Leon)
 MFOLD analysis of the RNA sequence was carried out using the MFOLD server.

3.4 Disruption of the stem loop structures I and II

If the secondary structures identified by MFOLD analysis are important or critical for virus replication then the introduction of mutations that disrupt these structures, but which maintained the correct protein, would be expected to create a phenotype. Similarly, if the underlying sequence, independent of the presence of structure, is essential then mutations that alter the sequence, but not the structure of the region, would create a phenotype. Synonymous sequence changes to be into structures I and II that would cause their collapse were designed initially by eye though computer analysis was used latterly (see appendix 3 for linear information on all described sequence changes). Unless otherwise stated, numbering of the nucleotides refers to its position in the full-length genome.

3.4.1 Disruption of structure I

Initially the design of changes that were to be introduced to the sequence was determined by manual analysis. The effect that the introduction of nucleotide alterations would have on the RNA structures was predicted using MFOLD analysis. Residues were identified that could be altered which would cause a targeted collapse of structure I or structure II. Analysis of structure I by eye identified three possible regions to mutate. For purposes of analysis, structure I was subdivided into three sections. The first of these sections was the terminal loop of structure I, the second possible region was a conserved GCCU stem situated in the middle of the structure and lastly the base of structure I.

Oligonucleotide primers were designed to allow mutagenesis of structure I using an overlapping PCR strategy described later on in this chapter. In order to aid the identification of cDNAs carrying mutagenic sequences restriction sites were engineered into the primers. In order to fully disrupt the base stem region of structure I two primers were designed, O-BSmut3Dpol and O-BSmut3Dpolrev (see materials and methods, table 6.2), that introduced five point mutations. In addition the primer pair had been designed so that two of the point mutations that were to be introduced would create a unique *Nhe I* restriction site in the viral cDNA sequence.

MFOLD analysis of the predicted structure of the region following introduction of the non-coding mutations U₆₉₅₈G, A₆₉₆₇U, A₆₉₇₀U, U₆₉₇₄A and C₆₉₇₅G showed a slight increase in the free energy state compared to that of the parental sequence (Figure 3.5). MFOLD

analysis predicts that the base stem region of structure I up to the second bulge loop will collapse. However the terminal loop and the fifth bulge loop are predicted to remain intact. The introduction of these changes to the RNA sequence is predicted by MFOLD analysis to collapse structure II into four independent stem-loop structures.

Computational programs can be used to identify the least energetically stable structure from a dataset of synonymous variants. As the number of possible synonymous variations on a 99 nucleotide sequence is in the region of 4^{33} per base limitations must be placed within the algorithm to enable calculation of the results within a suitable time-frame.

Analysis of the folded region 6768-7148 showed that a minimum of four non-coding nucleotide changes introduced into structure I would result in its minimum free energy state being formed. Using computational analysis two independent series of four mutations were identified that resulted in the collapse of structure I. Computational analysis predicted that the introduction of the non-coding changes C₆₉₁₃A, C₆₉₂₂A, U₆₉₂₉C and C₆₉₅₂U would result in the collapse of structure I. The MFOLD structure format plot, shown in figure 3.6 predicts that the introduction of these mutations would result in the formation of a free energy state of -86.8 in comparison with the wild-type poliovirus sequence that has a free energy state of -93.7. The MFOLD analysis shows that structure I has been completely disrupted at the top of the structure though the base of the structure remains unaffected. In addition the MFOLD analysis clearly shows that structure II and the flanking regions are unaffected by the introduction of these mutations when compared to the MFOLD analysis of the wild-type poliovirus type 3 sequence. An oligonucleotide primer, O-AD-1E, was designed that allowed the introduction of these mutations into structure I using an asymmetric PCR strategy described later in this chapter. As a consequence of the mutations being introduced into the parental sequence a *Bgl II* restriction enzyme site was created that could be utilised for screening purposes.

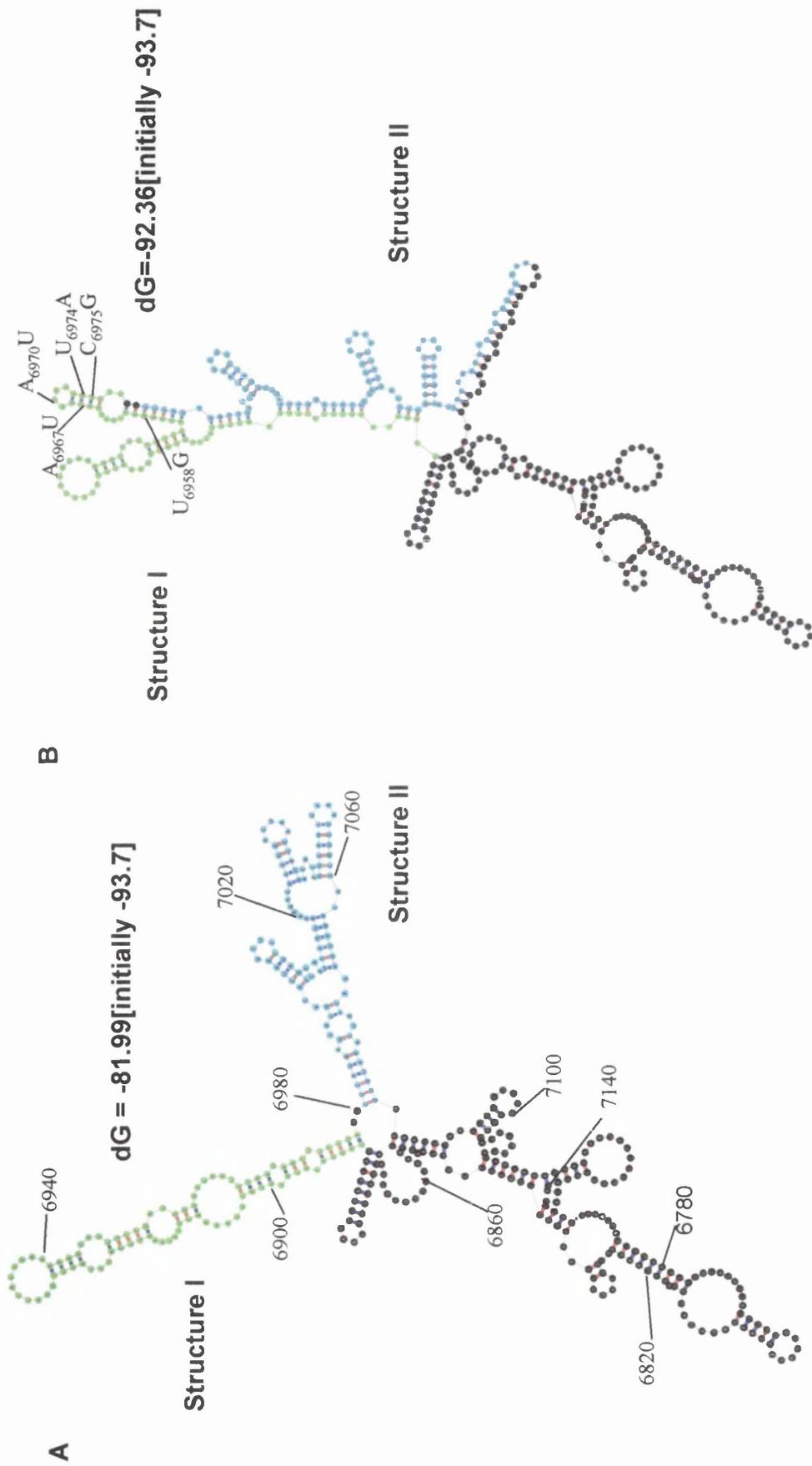


Figure 3.5 Structure format plot of nucleotides 6773-7148. (A) shows the structure format plot of nt 6773-7148 of poliovirus P3/leon/37 (type3). (B) shows the predicted effect of introducing the mutations U₆₉₅₈G, A₆₉₆₇U, A₆₉₇₀U, U₆₉₇₄A and C₆₉₇₅G into structure I. The introduction of these mutations into the poliovirus sequence creates a *Nhe* I restriction enzyme site. In the structure format plots shown structure I and structure II are highlighted in green and blue respectively.

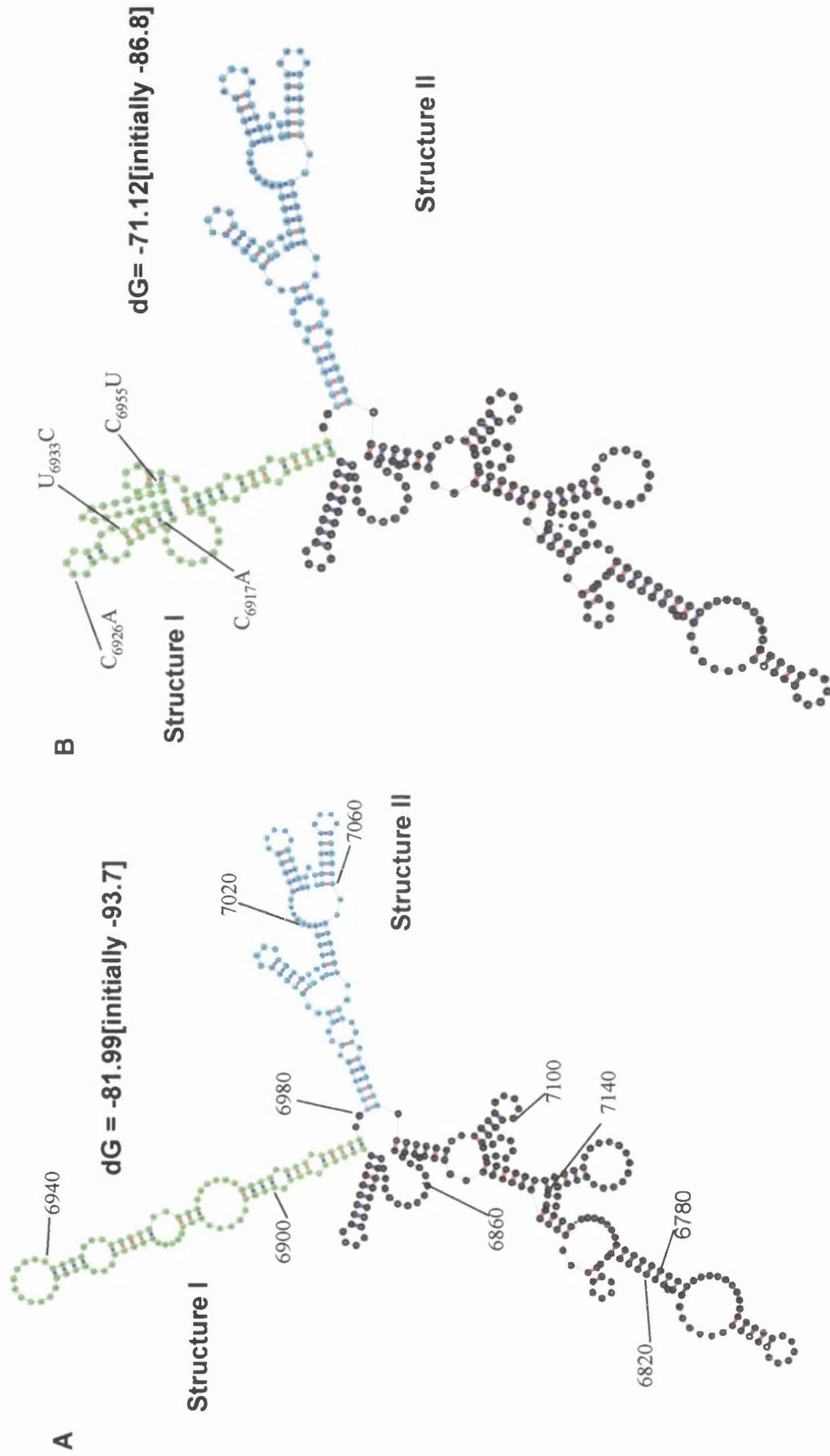


Figure 3.6 Structure format plots of nucleotides 6773-7148. (A) shows the structure format plot of nt 6773-7148 of poliovirus P3/leon/37 (type3). (B) shows the predicted effect of introducing the mutations C₆₉₁₇A, C₆₉₂₆A, U₆₉₃₃C and C₆₉₅₅U into structure I. The introduction of these mutations into the poliovirus sequence creates a *Bgl* II restriction enzyme site. In the structure format plots shown structure I and structure II are highlighted in green and blue respectively.

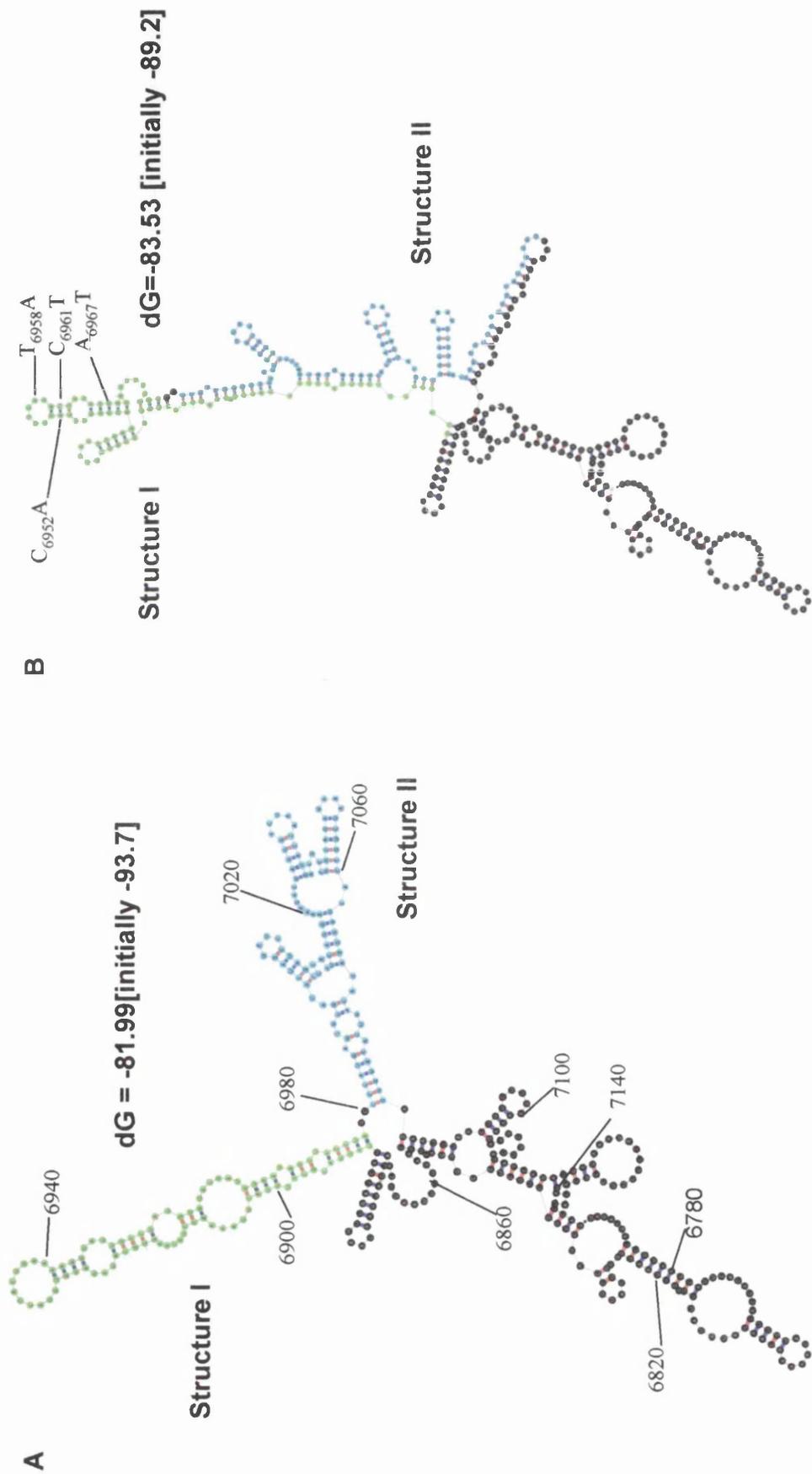


Figure 3.7 Structure format plot of nucleotides 6773-7148. (A) shows the structure format plot of nt 6773-7148 of poliovirus P3/leon/37 (type3). (B) shows the predicted effect of introducing the mutations C₆₉₅₂A, T₆₉₅₈A, C₆₉₆₁T and A₆₉₆₇T into structure I. The introduction of these mutations into the poliovirus sequence creates a *Nde* I restriction enzyme site. In the structure format plots shown structure I and structure II are highlighted in green and blue respectively.

The computational analysis also identified that the introduction of the non-coding changes C₆₉₅₂A, T₆₉₅₈A, C₆₉₆₁T and A₆₉₆₇T was capable of collapsing structure I. The MFOLD structure format plot predicts that the introduction of the sequence changes C₆₉₅₂A, T₆₉₅₈A, C₆₉₆₁T and T₆₉₆₇A would result in a structure that was significantly less stable than the parental sequence that has a free energy state of -93.7 (Figure 3.7). The MFOLD analysis predicts that the introduction of these nucleotide changes would result in the disruption of both structure I and structure II (Figure 3.7). The introduction of these mutations to the viral cDNA sequence, by the oligonucleotide primer O-AD-5F created an *Nde I* site that could be utilised for screening purpose.

3.4.2 Disruption of structure II

Analysis of the folded region 6768-7148, using computer analysis, identified two sets of mutations that would result in the collapse of structure II. An early version of the computational algorithm used to identify the least energetically favourable structure from a dataset of synonymous variants identified that the introduction of seven nucleotide changes into the sequence of structure II would result in the collapse of the structure. The MFOLD structure format plot predicts that the introduction of the sequence changes A₇₀₃₇U, C₇₀₄₃A, C₇₀₄₄A, G₇₀₄₆A, G₇₀₄₉A, C₇₀₅₆U and G₇₀₅₈A would result in a structure that was significantly less stable (free energy state of -81.3) than the parental sequence that has a free energy state of -93.7. The structure format plot (Figure 3.8) predicts that structure II will be completely disrupted, by the introduction of the aforementioned mutations, except for the 6 base pair stem found at the base of the structure. MFOLD analysis predicts that structure I and the flanking regions are unaffected by the introduction of these mutations when compared to the MFOLD analysis of the wild-type poliovirus type 3 sequence. An oligonucleotide primer, O-AD-3E, was designed that allowed introduction of these mutations into structure II using an asymmetric PCR strategy described later in this chapter. As a consequence of the mutations being introduced into the parental sequence an additional *Apo I* restriction enzyme site was created that was used for screening purposes.

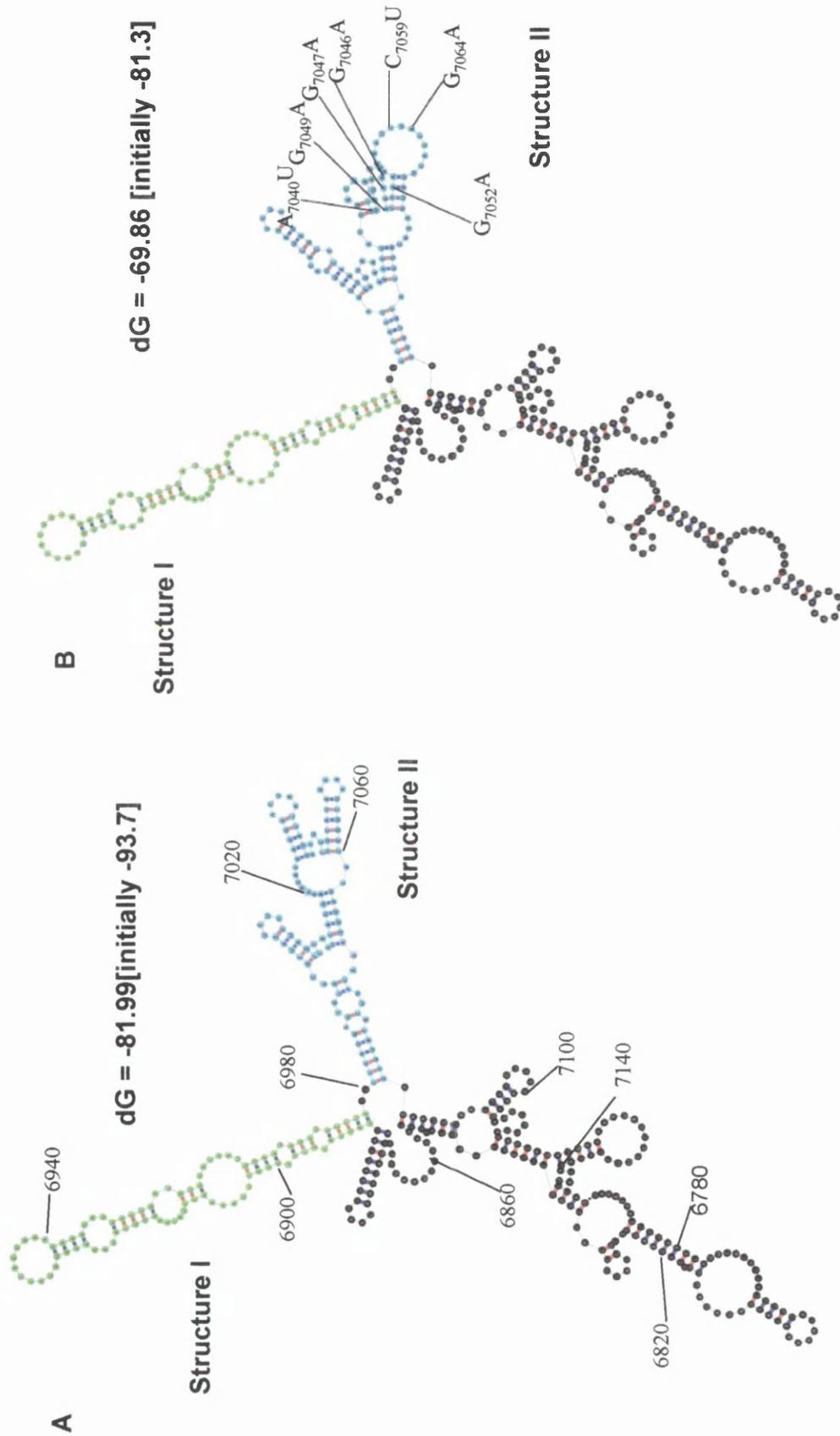


Figure 3.8 Structure format plot of nucleotides 6773-7148. (A) shows the structure format plot of nt 6773-7148 of poliovirus P3/leon/37 (type3). (B) shows the predicted effect of introducing the mutations A₇₀₄₀U, G₇₀₄₅A, G₇₀₄₇A, G₇₀₄₆A, G₇₀₄₉A, G₇₀₅₂A, C₇₀₅₉U, and G₇₀₆₄A into structure I. The introduction of these mutations into the poliovirus sequence creates a *Apo I* restriction enzyme site. In the structure format plots shown structure I and structure II are highlighted in green and blue respectively.

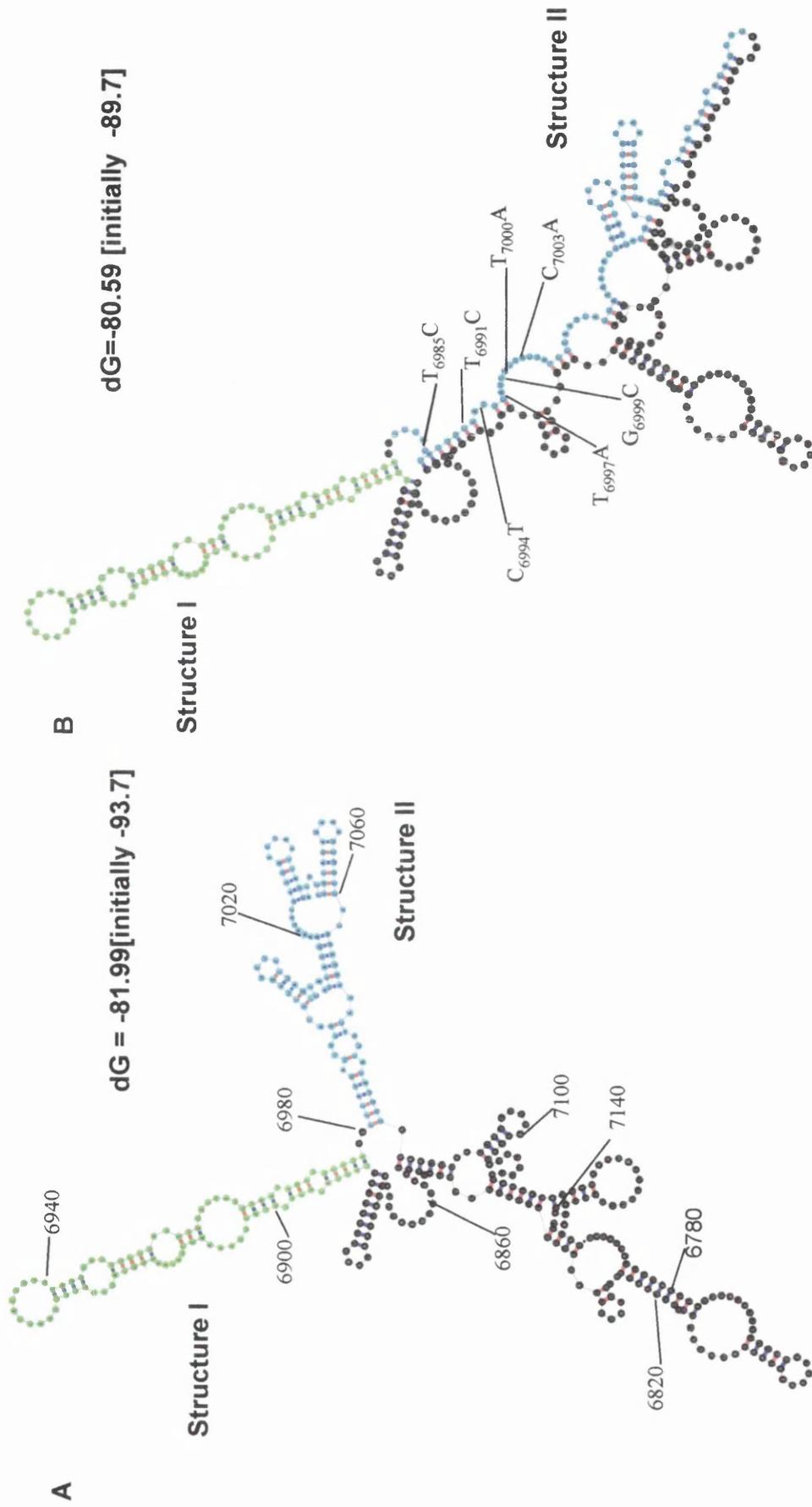


Figure 3.9 Structure format plot of nucleotides 6768-7148. (A) shows the structure format plot of nt 6773-7148 of poliovirus P3/leon/37 (type 3). (B) shows the predicted effect of introducing the mutations $T_{6985}C$, $T_{6991}C$, $C_{6994}T$, $T_{6997}A$, $A_{6998}T$, $G_{6999}C$, $T_{7000}A$ and $C_{7003}A$ into structure II. The introduction of these mutations into structure II results in the introduction of an *Nsi*I restriction site. In the structure format plots shown structure I and structure II are highlighted in green and blue respectively.

A later version of the computer algorithm detected that the introduction of eight nucleotide changes would collapse structure II. These were T₆₉₈₅C, T₆₉₉₁C, C₆₉₉₄T, T₆₉₉₇A, A₆₉₉₈T, G₆₉₉₉C, T₇₀₀₀A and C₇₀₀₃A. The MFOLD structure format plot predicts that the introduction of the sequence changes T₆₉₈₅C, T₆₉₉₁C, C₆₉₉₄T, T₆₉₉₇A, A₆₉₉₈T, G₆₉₉₉C, T₇₀₀₀A and C₇₀₀₃A would result in a structure that was significantly less stable than the parental sequence that has a free energy state of -93.7 (Figure 3.9). The MFOLD analysis predicts that the introduction of these nucleotide changes would result in the collapse of structure II. In contrast, the MFOLD analysis predicts that structure I will be unaffected, by the introduction of these nucleotide changes. An oligonucleotide primer, O-AD-4F, was designed that allowed introduction of these mutations into structure II using an asymmetric PCR strategy described later in this chapter. As a consequence of the mutations being introduced into the parental sequence an additional *Nsi I* restriction enzyme site was created that was used for screening purposes.

The calculated free energies for each of the mutated RNA structure and the structure format plots predicts that the synonymous changes introduced should successfully disrupt the targeted structure (either structure I or structure II).

cDNA construct	Luciferase Replicon/ infectious clone	Mutations introduced (nt refer to position in full- length viral genome)	Unique identifying restriction enzyme site (nt location)	Predicted effects of substitution on structure	Laboratory identification code
pT7luc3D ^{Nhe I} +R	Luciferase Replicon	U ₆₉₅₈ G, A ₆₉₆₇ U, A ₆₉₇₀ U, U ₆₉₇₄ A, C ₆₉₇₅ G	<i>Nhe I</i>	Partial disruption of structure 1 up to the second bulge loop. Complete structural disruption of Structure 2	P-AE-6A
pT7FLC3D ^{Nhe I}	Infectious cDNA	U ₆₉₅₈ G, A ₆₉₆₇ U, A ₆₉₇₀ U, U ₆₉₇₄ A, C ₆₉₇₅ G	<i>Nhe I</i>	See above	P-AE-5A
pT7luc3D ^{Bgl II} +R	Luciferase Replicon	C ₆₉₁₃ A, C ₆₉₂₂ A, U ₆₉₂₉ C, C ₆₉₅₂ U	<i>Bgl II</i>	Disruption of structure 1 above the first bulge loop and including the terminal loop. No destabilisation of structure 2	P-AE-6B
pT7luc3D ^{Apo I} +R	Luciferase Replicon	A ₇₀₃₇ U, G ₇₀₄₃ A, G ₇₀₄₄ A, G ₇₀₄₆ A, G ₇₀₄₉ A, C ₇₀₅₆ U, G ₇₀₅₈ A	<i>Apo I</i>	Disruption of structure 2. No destabilisation of structure 1	N/A
pT7FLC ^{Apo I}	Infectious cDNA	A ₇₀₃₇ U, G ₇₀₄₃ A, G ₇₀₄₄ A, G ₇₀₄₆ A, G ₇₀₄₉ A, C ₇₀₅₆ U, G ₇₀₅₈ A	<i>Apo I</i>	See above	N/A
pT7luc3D ^{Nde I} +R	Luciferase Replicon	C ₆₉₅₂ A, T ₆₉₅₈ A, C ₆₉₆₁ T and A ₆₉₆₇ T	<i>Nde I</i>	Disruption of structure 1	P-AE-7B
pT7luc3D ^{Nsi I} +R	Luciferase Replicon	T ₆₉₈₅ C, T ₆₉₉₁ C, C ₆₉₉₄ T, T ₆₉₉₇ A, A ₆₉₉₈ T, G ₆₉₉₉ C, T ₇₀₀₀ A and C ₇₀₀₃ A	<i>Nsi I</i>	Disruption of structure 2	P-AE-8A

Table 3.1 Subgenomic replicon and Infectious cDNAs constructs

3.5 Site directed mutagenesis of the region nt 6768-7148

In order to investigate whether disruption of structures I and II would result in a replication phenotype the mutations were built into a replicon containing the firefly (*Photinus pyralis*) luciferase gene, constructed by Dr. I Goodfellow, based on the well characterised replicon pT7Rep3 (Percy et al., 1992).

The cDNA fragments, bearing the desired changes in structures I and structure II, which were built into the replicon were created using two different PCR mutagenesis strategies. The cDNA constructs engineered in the course of this study are described in table 3.1. All mutagenic PCR reactions, as described below, were carried out using *pfu*^{platinum} polymerase supplied by Invitrogen.

3.5.1 PCR mutagenesis using an overlap strategy

The overlap PCR mutagenesis strategy requires 4 primers (Figure 3.10). The first primer pair requires a forward primer (primer 1) that anneals to the 5' end of the region that is to be amplified and a reverse primer (primer 3) that anneals in a downstream position. The reverse primer was designed to contain the point mutations that were to be introduced into the sequence.

The second primer pair was designed so that the Forward primer (primer 2) had a region of complementarity to the 3' region of the PCR product produced from the first primer pair. Where the point mutations would allow for the introduction of a unique restriction site for screening purposes complementarity of the primers was designed around this site. The reverse primer (primer 4) in the second primer pair like the forward primer of the first primer pair introduces a restriction enzyme site. As the first round products share a region of complementarity the PCR products can anneal together and act as template in the second-round PCR reaction that uses the primer pair (1+4) to generate the full-length product. The final PCR product was cloned into the plasmid vector of choice using the restriction enzyme sites X and Y. PCR mutagenesis using an overlap strategy was used to construct pT7luc3D^{NheI}+R.

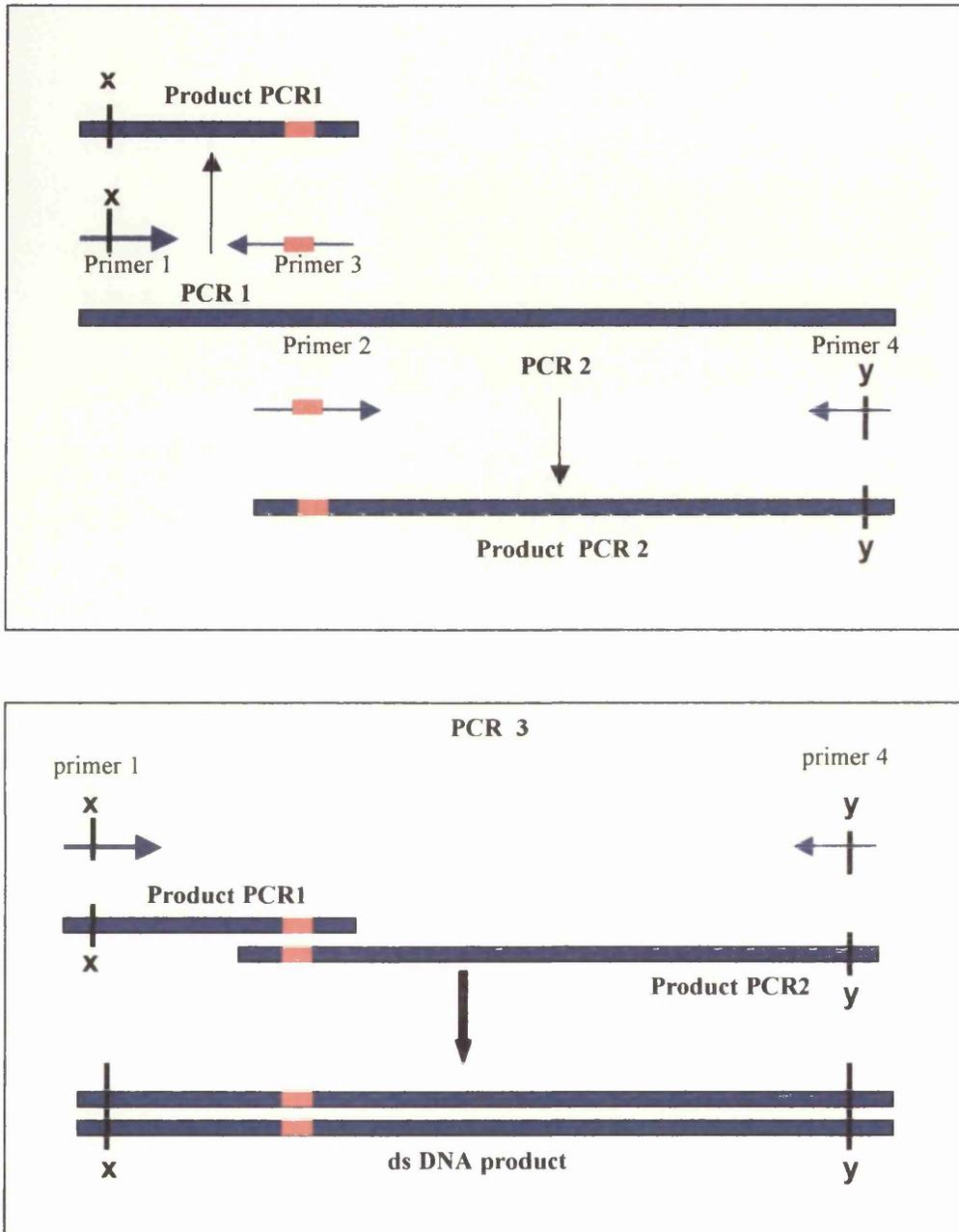


Figure 3.10 Mutagenesis using an overlapping PCR Strategy.

Figure 3.10 outlines the requirements for each stage of the overlap PCR method to create a full-length PCR product. Basically, the first round products share a region of complementarity and can anneal together. They can therefore be used as templates in a PCR reaction which uses the external primer pair (primers 1+4) to generate the full-length product. The PCR product can then be cloned into the vector of choice using the restriction enzyme sites X and Y. In Figure 3.10 the parental DNA sequence is shown in blue while the area of mutagenesis is highlighted in red.

3.5.2 Mutagenesis using an asymmetric PCR strategy

In addition to the overlap PCR strategy as a way of introducing mutations into cDNA, an asymmetric PCR strategy was also employed (Figure 3.11). This technique requires the use of three primers. The first primer pair requires a forward primer (primer B) that anneals to the 5' end of the region that is to be amplified and a reverse primer (primer C) that anneals in a downstream position. The forward primer (primer B) is designed with approximately 10 nucleotides of wildtype sequence flanking on either side the region into which the point mutations are to be introduced. The reverse primer (primer C) in the first primer pair introduces a restriction enzyme site (X) that was used for cloning into the plasmid vector of choice.

The presence of 10 nucleotides of wildtype sequence at the 5' end of the PCR product produced from the first primer pair enables the polymerase to use the anti-sense strand of the PCR product as a reverse primer in the second round of PCR. The Forward primer (primer A) in the second round PCR anneals to the 5' end of the region to be amplified. The forward primer like the reverse primer used in the first round PCR introduces a restriction enzyme site (Y) that is used for cloning the PCR product into the plasmid vector.

Thus the second round of PCR uses the purified PCR product from the first round PCR as a "megaprimer" to generate a full length PCR product. Mutagenesis using the asymmetric PCR method was used in the construction of replicons pT7luc3D^{Apol}+R, pT7luc3D^{Nde I}+R, pT7luc3D^{Nsi I}+R and pT7luc3D^{BglII}+R. In constructing pT7luc3D^{BglII}+R an annealing and extension step, of six cycles prior to the second round of PCR, was shown to increase the yield of full length product.

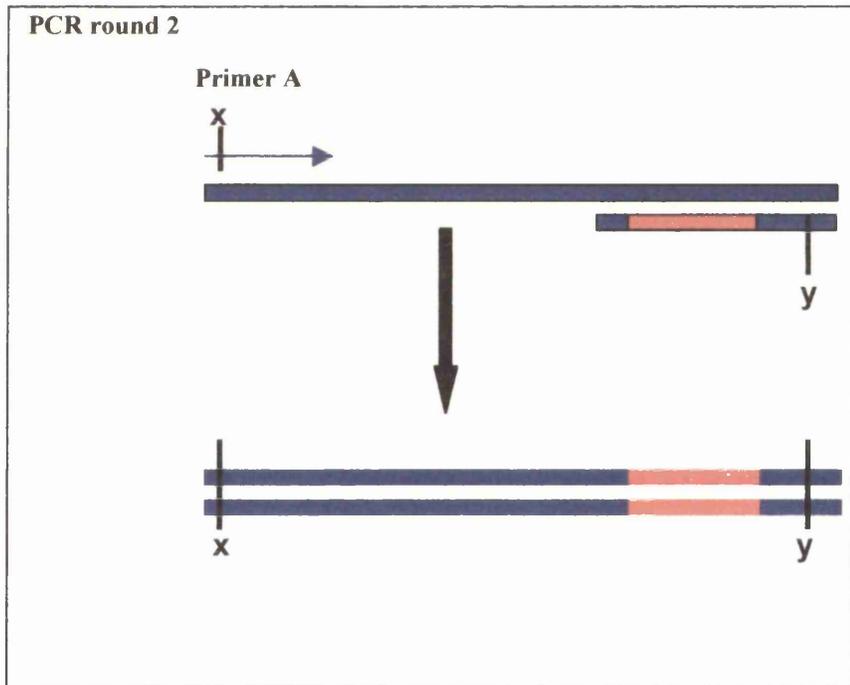
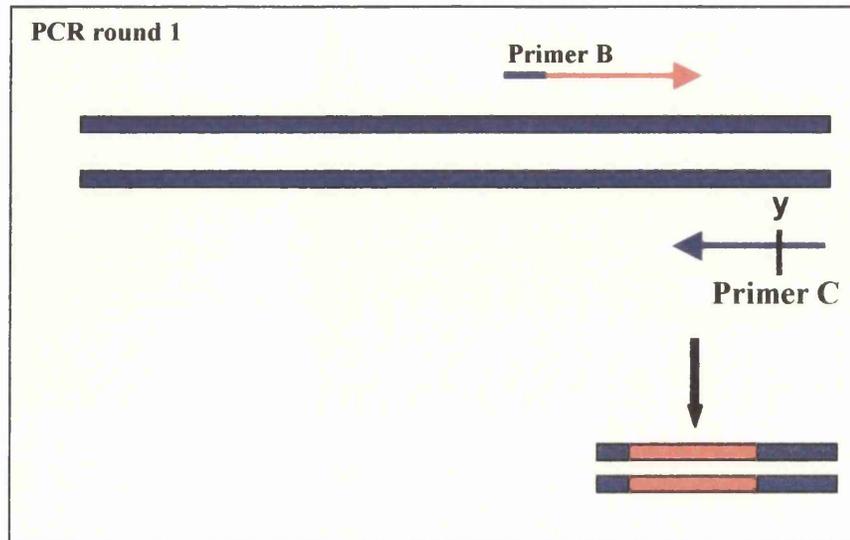


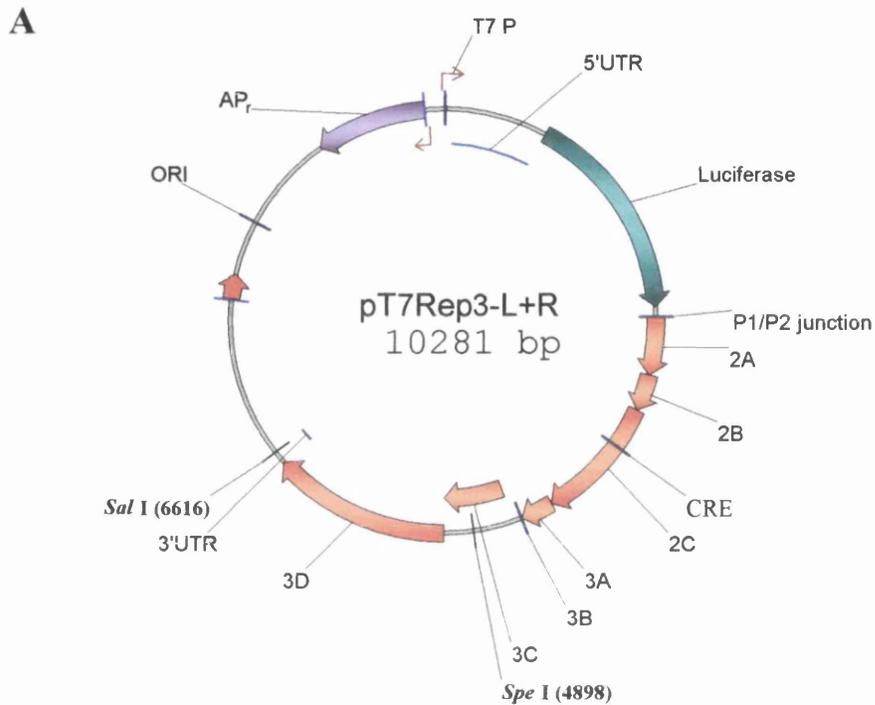
Figure 3.11 Mutagenesis using asymmetric PCR.

Figure 3.11 outlines the requirements of each stage of the asymmetric PCR method to create a full-length PCR product. Basically, the first round product can be used as a primer in a second PCR reaction to generate the full-length product. The PCR product can then be cloned into the vector of choice using the restriction enzyme sites X and Y. In Figure 3.11 the parental DNA sequence is shown in blue while the area of mutagenesis is highlighted in red

3.6 Construction of subgenomic replicons bearing mutations in 3D^{pol} structure I and II

Point mutations that were to be built into the luciferase replicon pT7Luc+R were introduced into the region 6773- 7173 using the PCR mutagenesis techniques described previously. In each case the non-mutagenic primers used were the forward primer IG48 (see Table 6.2 materials and methods) that binds upstream of the *Spe I* site and the reverse primer Rep3rev (see Table 6.2 materials and methods) that binds downstream of the *Sal I* site in pT7Luc+R (Figure 3.12). The final PCR product and the plasmid vector were prepared for cloning by digestion with *Spe I* and *Sal I* restriction enzymes. The restricted insert and vector were purified from agarose and ligated together. The ligations were cleaned up prior to transformation into electrocompetant *E. coli* ER278 cells. Transformed cells were plated onto antibiotic selective media. DNA was prepared from putative clones and analysed for the presence of distinct restriction digestion patterns that identified that the mutagenesis had been successful.

To verify the presence of the designed mutations and to rule out the introduction of unwanted mutations that introduced coding changes the cloned 2 kb region was sequenced, by automated sequencing at the in-house facility, using the primers O-AD-7B, O-AD-8C, IG24 and IG26 (see Table 6.2 materials and methods). With the exception of pT7Luc3D^{Apol}+R this sequencing verified that no coding changes had been introduced into the region by the PCR and that only the desired mutations were present. pT7Luc3D^{Apol}+R will be discussed in more detail later.



B T7 polymerase RNA transcript



Figure 3.12 Luciferase replicon pT7Luc+R.

(A) The vector map of replicon pT7Luc+R. (B) The transcript produced by the bacteriophage T7 RNA polymerase contains the information for expression of the luciferase gene fused to the message encoding the viral replication proteins at the P1/P2 junction. The presence of a ribozyme in the vector results in a fully authentic viral RNA-like 5' end.

3.7 Introduction of mutations in 3D^{pol} structure I and II into an infectious cDNA

The use of secondary structures as signals for the selective packaging of genomes into the viral particle had been identified in a number of viruses including HBV and HIV-1 (Schlesinger et al., 1994). Though earlier work using subgenomic replicons had argued that a packaging signal in the genome of picornaviruses must exist (Barclay et al., 1998, Jia et al., 1998, Porter et al., 1998) the first picornavirus encapsidation signal was identified, in Aichi virus, only recently (Sasaki & Taniguchi, 2003). In contrast to Aichi virus, where the packaging signal was identified in the 5'UTR the accumulated evidence provided by studies using genetically-modified viruses (Alexander et al., 1994, Todd et al., 1997) and subgenomic replicons (Barclay et al., 1998) indicates that the poliovirus packaging signal is located within the region of the genome encoding the non-structural proteins. Given that Nugent *et al* (Nugent et al., 1999) had observed that packaging in poliovirus is coupled to replication the possibility was raised that though a replication phenotype might not be observed using the replicon-based system, a severe phenotype might be observed when the mutations were analysed in the context of a full virus lifecycle.

To investigate the possibility that structure I and structure II influenced such as packaging or uncoating the synonymous changes, that had given a subtle replication phenotype in the replicon system, were introduced into an infectious clone available in the laboratory. This work was carried out prior to obtaining the results of the translation and processing assay discussed in section 3.8.2.2.

To construct pT7FLC3D^{Apol} from the subgenomic replicon pT7Luc3D^{Apol}+R the restriction enzyme sites *Not I* and *BssHIII* (Figure 3.13) were utilised. Both the restricted insert (VP1-4 derived from pT7FLC) and cDNA backbones, bearing the changes, were purified from agarose prior to being ligated together.

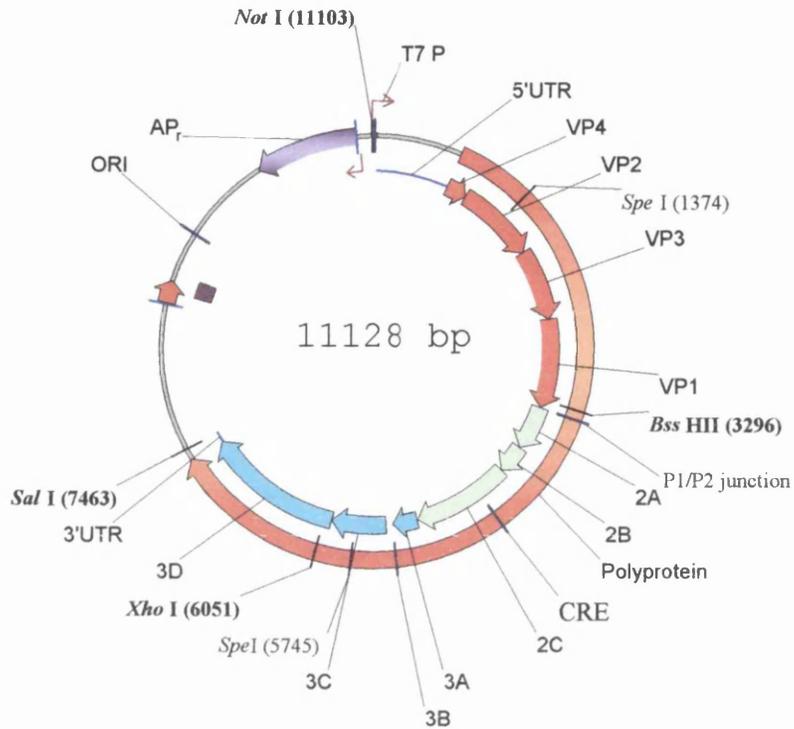


Figure 3.13 Vector map of the infectious clone pT7FLC

The infectious clone is based on poliovirus P3/Leon/37 (type3). The infectious clone has been engineered to contain a T7 promoter and a unique *Sal I* site. The proteins of the P1 region (VP1-4) are shown in red, the proteins of the P2 region are highlighted in green and the replication proteins (P3 region) are shown in blue. Restriction enzyme sites used in the construction of infectious clones are in bold.

pT7FLC3D^{NheI} was constructed in a slightly different manner. pT7Luc3D^{NheI}+R was digested with *Xho I* and *Sal I* restriction enzymes. Both the insert obtained from pT7Luc3D^{NheI}+R and the vector, pT7FLC that had been digested in a similar way, were purified from agarose and ligated together. The ligations were cleaned up prior to transformation into electrocompetant *E. coli* ER278 cells. Transformed cells were plated onto antibiotic selective media. The presence of the capsid gene in putative pT7FLC3D^{Apol} cDNAs were screened for using a *Sma I*/*Xho I* digestion, pT7FLC3D^{NheI} was screened using the unique *Nhe I* site. Correct orientation of the cloned fragment in putative pT7FLC3D^{NheI} cDNAs was screened using *Xho I* and *Sal I* restriction enzymes.

3.8 Functional studies on the effect of mutations in 3D^{pol} structure I and II

3.8.1 Analysis of the influence of structure I and II mutations on replication

Naturally occurring defective interfering (DI) particles of poliovirus have been shown to replicate their genomes efficiently though they can have up to 13 % of the parental genome deleted (Kajigaya et al., 1985, Omata et al., 1986). All naturally occurring DI particles of poliovirus contain in-frame deletions in the capsid coding region of the genome (Cole & Baltimore, 1973, Kuge et al., 1986). DI genomes constructed from cloned cDNAs have been demonstrated to be translation and replication competent (Hagino Yamagishi & Nomoto, 1989, Percy et al., 1992).

The ability of the DI genomes to be translated and replicated in the absence of the capsid sequences has been used to construct self-amplifying subgenomic replicons. In addition to poliovirus (Kaplan & Racaniello, 1988, Percy et al., 1992) subgenomic replicons have been constructed from the genomes of several picornaviruses including hepatitis A virus (Yi & Lemon, 2002), foot and mouth disease virus (McInerney et al., 2000), rhinovirus 14 (McKnight & Lemon, 1996) and Aichi virus (Sasaki et al., 2001).

pT7FLC/Rep is a subgenomic replicon that was constructed from the poliovirus type 3 Leon strain. In pT7FLC/Rep VP4, VP2 and the amino terminus of VP3 have been deleted and replaced by the CAT gene, in-frame with the rest of the viral genome. Monitoring of the production of CAT by enzymatic assay following transfection of the RNA transcript into mammalian cells showed that the RNA synthesised from pT7FLC/Rep was amplified following transfection. Addition, post-transfection, and maintenance throughout the incubation period of guanidine hydrochloride in the cell media inhibited the CAT signal. As a known inhibitor of poliovirus RNA synthesis (Baltimore et al., 1963, Tershak, 1982), the ability of guanidine hydrochloride to inhibit the production of CAT enabled a direct correlation to be made between the production of CAT and the replication-competent nature of the RNA transcripts tested. The observation that RNA derived from pT7FLC/Rep could be packaged *in trans* by helper virus enables this system to be extended to study packaging signals (Percy et al., 1992).

In order to investigate whether the introduced mutations had an effect on the replication of the viral genome, the mutations had been introduced into the replicon pT7luc +R.

pT7luc+R is a subgenomic replicon that was derived from pT7FLC/Rep3, a replicon 25% smaller in size than the full length parental genome (Barclay et al.). Expression of the luciferase gene from the pT7luc+R construct has been shown in previous studies, using replicons based on the poliovirus type 1 genome, to be a readily quantifiable and reproducible marker for viral genomic amplification when used in this way (Andino et al., 1993).

Replicons were linearised using *Sal I*. RNA was transcribed from the linearised template using the bacteriophage T7 polymerase. Normalised levels of RNA were transfected into HeLa T4 cells using electroporation. The cells were incubated in DMEM/10%FCS at 37 °C/ 5% CO₂. Cellular extracts from the cell were harvested at 6 hours post transfection and assayed for luciferase activity. Guanidine hydrochloride was added to the DMEM/10%FCS of the negative control at a concentration of 4 mM post transfection and maintained throughout incubation. The production of luciferase is presented as an average of three independent transfections with error bars calculated to show the variation between the independent samples.

3.8.1.1 Influence on replication of structure I

Following six hours incubation, significant luciferase signals could be obtained following transfection of HeLa cells with RNA derived from pT7luc3D^{NheI}+R, pT7luc3D^{NdeI}+R or pT7luc3D^{BglII}+R. Analysis of the production of luciferase produced showed that the RNA synthesised from pT7luc3D^{NheI}+R was reduced by approximately 0.5 log₁₀ with respect to RNA derived from the parental cDNA (Figure 3.14A). A decrease in replication of approximately 0.5 log₁₀ with respect to RNA derived from the parental cDNA was also observed with RNA derived from pT7luc3D^{BglII}+R (Figure 3.15A) or pT7luc3D^{NdeI}+R (Figure 3.15B). As a result functional characterisation of RNA derived from pT7luc3D^{BglII}+R or pT7luc3D^{NdeI}+R went no further. Consistent with the luciferase signal being an accurate measure of replication, no signal was obtained from cells transfected with RNA that had been incubated in the presence of guanidine hydrochloride (Figure 3.14 and 3.15). Although the difference between the RNA derived from the parental cDNA and pT7luc3D^{NheI}+R was subtle, it was shown to be consistently reproducible.

In order to try and ascertain the point of the replication cycle that was affected by the mutations the transfections were repeated and cellular extracts were harvested at 0, 1, 2, 4, 6 and 8 hours post transfection. The efficiency of the RNA synthesised from pT7luc3D^{NheI}+R to replicate in HeLa cells was reduced in comparison to RNA synthesised from the parental cDNA at an earlier stage of the viral replication cycle. Figure 3.14B shows that at 4 hours the replication of RNA derived from pT7luc3D^{NheI}+R is reduced by approximately 0.5 log₁₀ when compared to the parental clone. At 8 hours post transfection luciferase production from pT7luc3D^{NheI}+R appears to be equivalent with that produced from the parental clone. Although subtle, the differences in replication observed between pT7luc3D^{NheI}+R and the parental clone over the duration of the time-course were consistently observed.

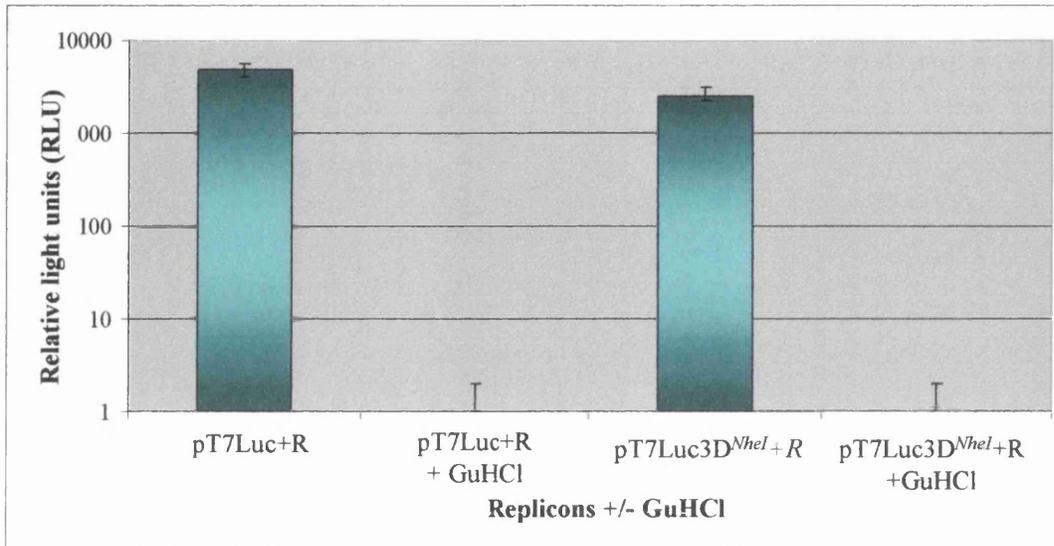
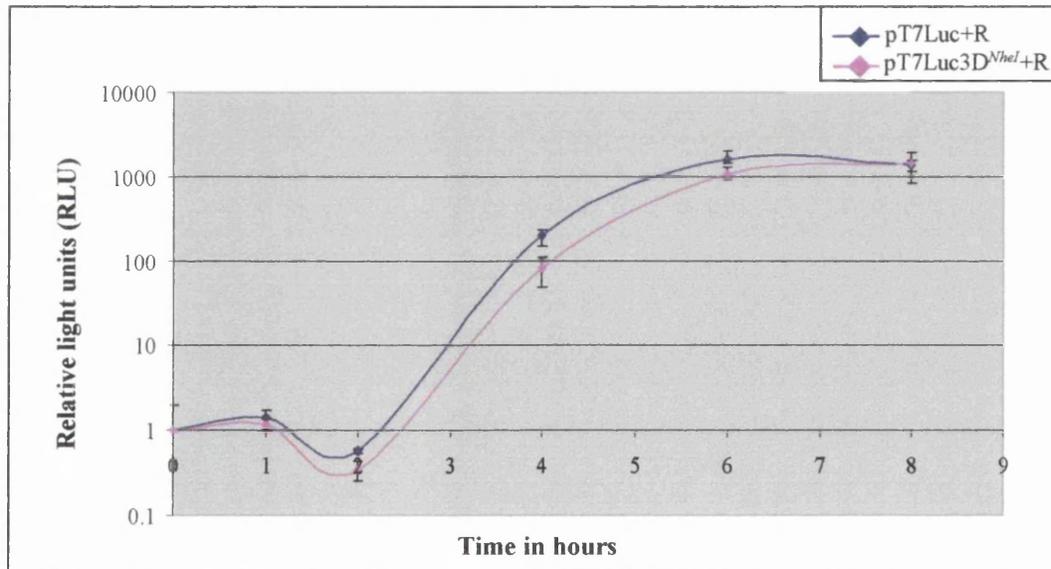
A**B**

Figure 3. 14 (A) Replication of pT7Luc3D^{NheI}+R, using the luciferase system.

Normalised levels of RNA were transfected into HeLa T4 cells. Cellular lysates were harvested at 6 hours post transfection and assayed for luciferase activity. Guanadine hydrochloride a known inhibitor of poliovirus replication was used as a negative control for replication. RNA derived from the parental replicon pT7Luc+R was used as a positive control. **(B) Time course of replication of RNA derived from pT7Luc3D^{NheI}+R.** Normalised levels of RNA were transfected into HeLa T4 cells. Cellular extracts were harvested at 0,1,2,4,6 and 8 hours post transfection and assayed for luciferase activity. RNA derived from the parental replicon pT7Luc+R was used as a positive control. Error bars were calculated to show the variation between the individual samples.

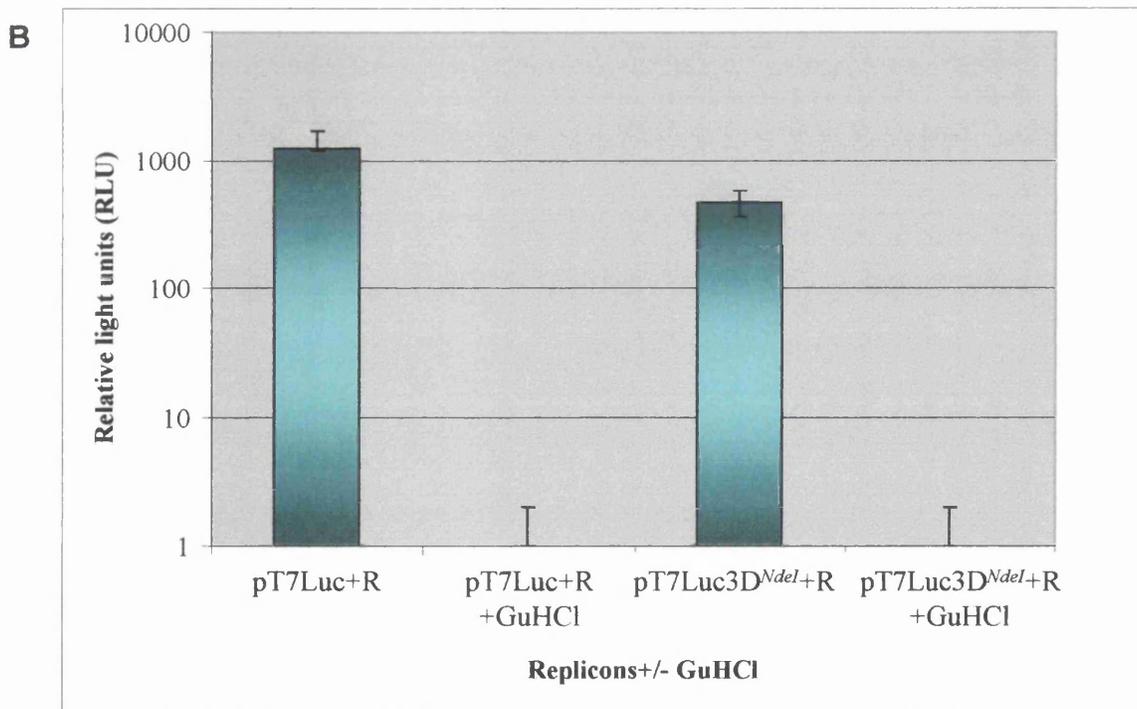
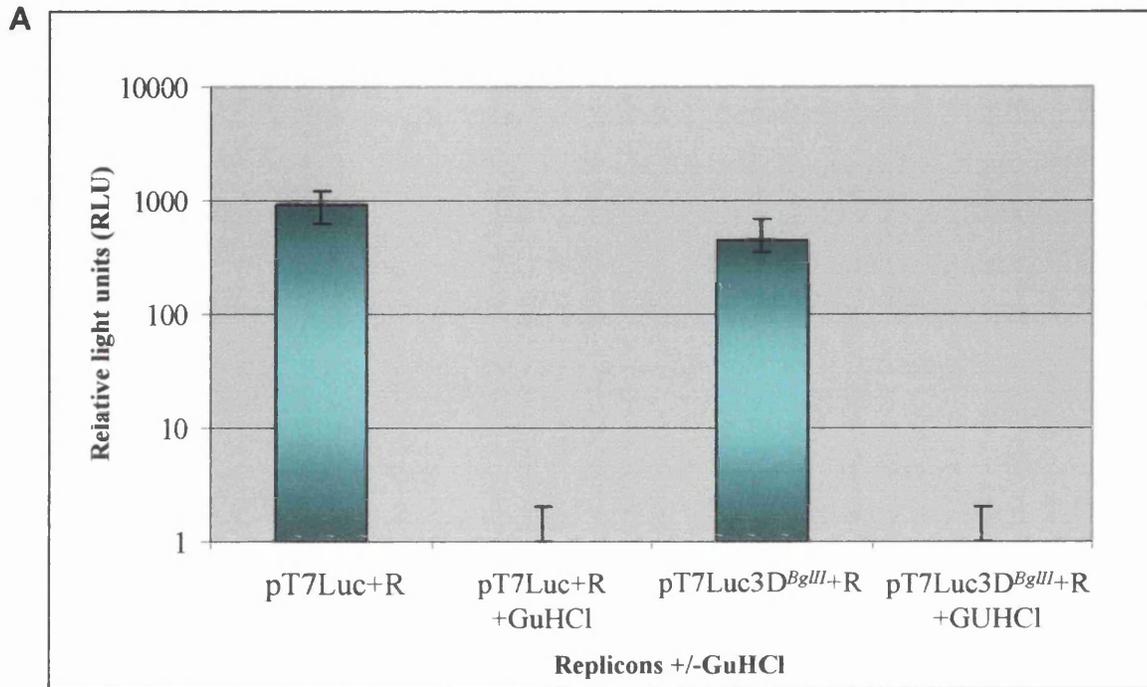


Figure 3.15. Replication assay using the luciferase system.

(A) Replication of pT7Luc3D^{BglII}+R. (B) Replication of pT7Luc3D^{NdeI}+R. Normalised levels of RNA were transfected into HeLa T4 cells. Cellular lysates were harvested at 6 hours post transfection and assayed for luciferase activity. Guanidine hydrochloride a known inhibitor of poliovirus replication was used as a negative control for replication. RNA derived from the parental replicon pT7Luc+R was used as a positive control. Error bars were calculated to show the variation between the individual samples.

3.8.1.2 Influence on replication of structure II

Following six hours incubation significant luciferase signals could be obtained following transfection of HeLa cells with RNA derived from pT7luc3D^{Nsi I}+R and the parental replicon pT7luc+R, but not pT7luc3D^{Apo I}+R. Consistent with the luciferase signal being an accurate measure of replication no signal was obtained from cells, transfected with RNA derived from either pT7luc3D^{Nsi I}+R or pT7luc+R, that had been incubated in the presence of guanidine hydrochloride (Figure 3.16A). A decrease in replication of approximately 0.5 log₁₀ with respect to RNA derived from the parental cDNA was observed with RNA derived from pT7luc3D^{Nsi I}+R.

In contrast, to the subtle replication phenotype observed when analysing the replication efficiency of RNA derived from pT7luc3D^{Nsi I}+R no significant luciferase signal could be detected above that of the negative control when the replication efficiency of transcripts derived from pT7luc3D^{Apo I}+R was assessed (Figure 3.16B). Thus while the RNA transcribed from pT7luc3D^{Apo I}+R appeared to be competent as a template for translation it appeared that RNA derived from pT7luc3D^{Apo I}+R was replication-incompetent. The reason as to why RNA derived from pT7luc3D^{Apo I}+R is replication-incompetent while RNA derived from pT7luc3D^{Nsi I}+R is replication competent will be discussed in depth in section 3.8.2.

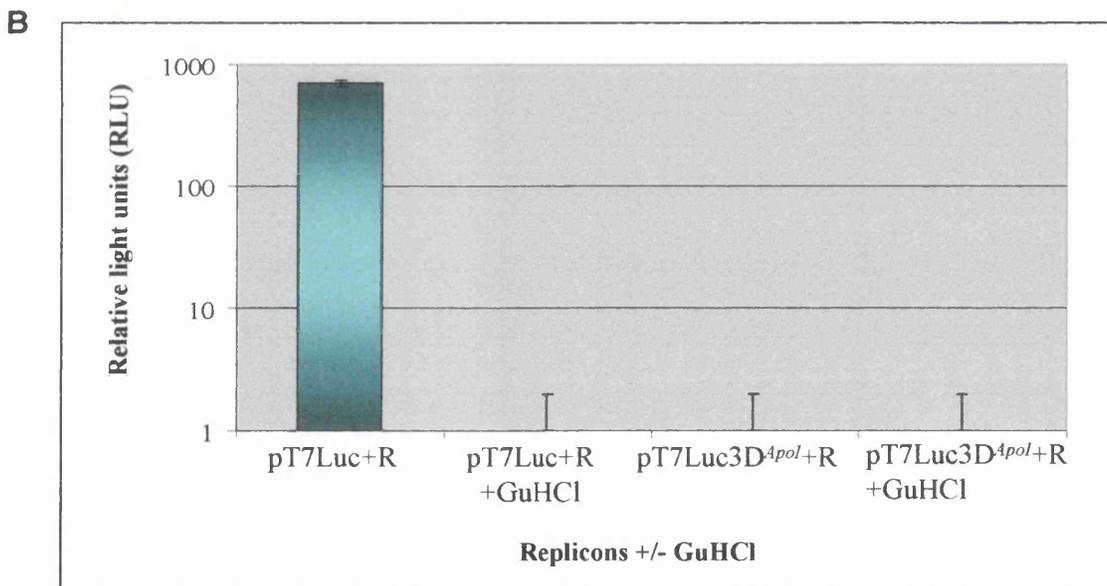
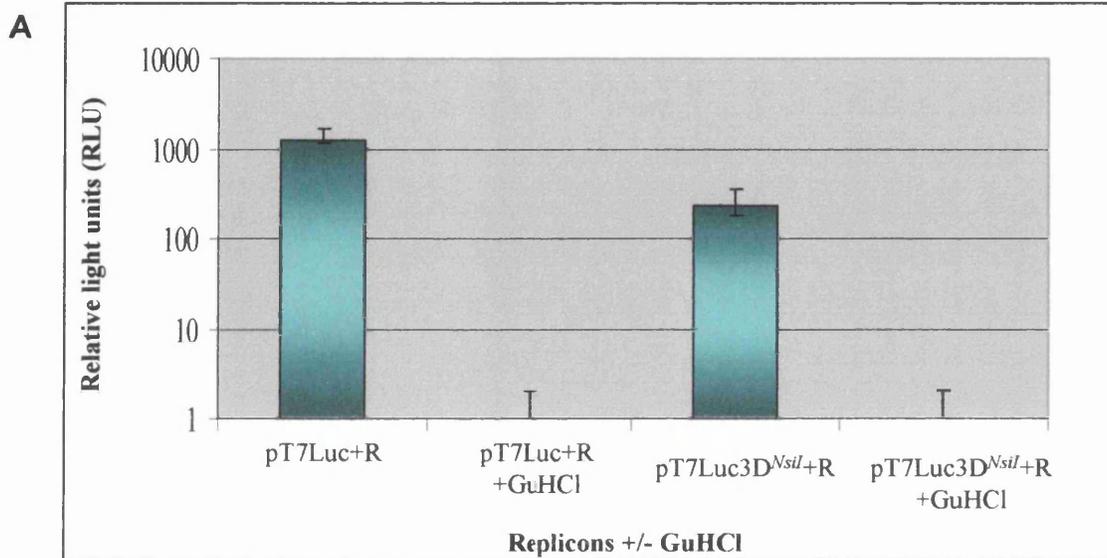


Figure 3.16. Replication assay using the luciferase replicon

(A) Replication of pT7Luc3D^{Nsil}+R. **(B)** Replication of pT7Luc3D^{Apol}+R.

Normalised levels of RNA were transfected into HeLa T4 cells. Cellular lysates were harvested at 6 hours post transfection and assayed for luciferase activity. Guanadine Hydrochloride a known inhibitor of poliovirus replication was used as a negative control for replication. The parental replicon pT7Luc+R was used as a positive control. Error bars were calculated to show the variation between the individual samples.

3.8.2 Influence on translation and processing of mutations in 3D^{pol} structure I and II

In order to confirm the RNA bearing changes was being translated and processed correctly an *in vitro* translation and processing reaction was carried out using a FLEXI® rabbit reticulocyte lysate system (Promega) supplemented with HeLa S10 extracts. HeLa S10 extracts have been demonstrated to contain all the cellular factors required to support the translation and replication of infectious poliovirus *in vitro* (Barton & Flanagan, 1993, Molla et al., 1991). The HeLa S10 extracts used in this experiment were made and donated by Dr. I. Goodfellow. The *in vitro* translation and processing assay can be used to analyse the correct processing of the protein product translated from RNA synthesised off either subgenomic replicons or full-length cDNAs.

In order to generate the large quantities of good quality RNA required to carry out a translation and processing assay the RIBOMAX™ transcription kit (Promega) was used. 2 µg of purified RNA was used per translation reaction. 2 mM guanidine hydrochloride was added to each translation reaction to prevent the replication of the transcribed viral RNA. The radiolabelled protein products of the translation and processing reaction were visualised using a phosphoimager and analysed using the Personal FX computer program (BioRad) following SDS-PAGE gel electrophoresis.

3.8.2.1 Translation and processing of RNA derived from the subgenomic replicons pT7luc3D^{Ndel}+R and pT7luc3D^{Nsil}+R

The *in vitro* translation and processing reaction of RNA derived from pT7luc+R was used as the positive control (Figure 3.17 Lane 2) as to whether the viral proteases 2A^{pro}, 3C^{pro} and 3CD^{pro} were able to correctly process the polyprotein derived from subgenomic RNAs, bearing changes in structure I or structure II. As a further control S³⁵-methionine labelled poliovirus type 3 infectious lysate (Figure 3.17 Lane 1) was run alongside the samples from the *in vitro* translation and processing reactions on the SDS-PAGE electrophoresis gel. Visualisation of the products of the *in vitro* translation and processing assay confirmed that there was no difference in the processing of the polyprotein expressed from RNA derived from the cDNAs pT7luc3D^{Ndel}+R (Figure 3.17 Lane 3) and pT7luc3D^{Nsil}+R

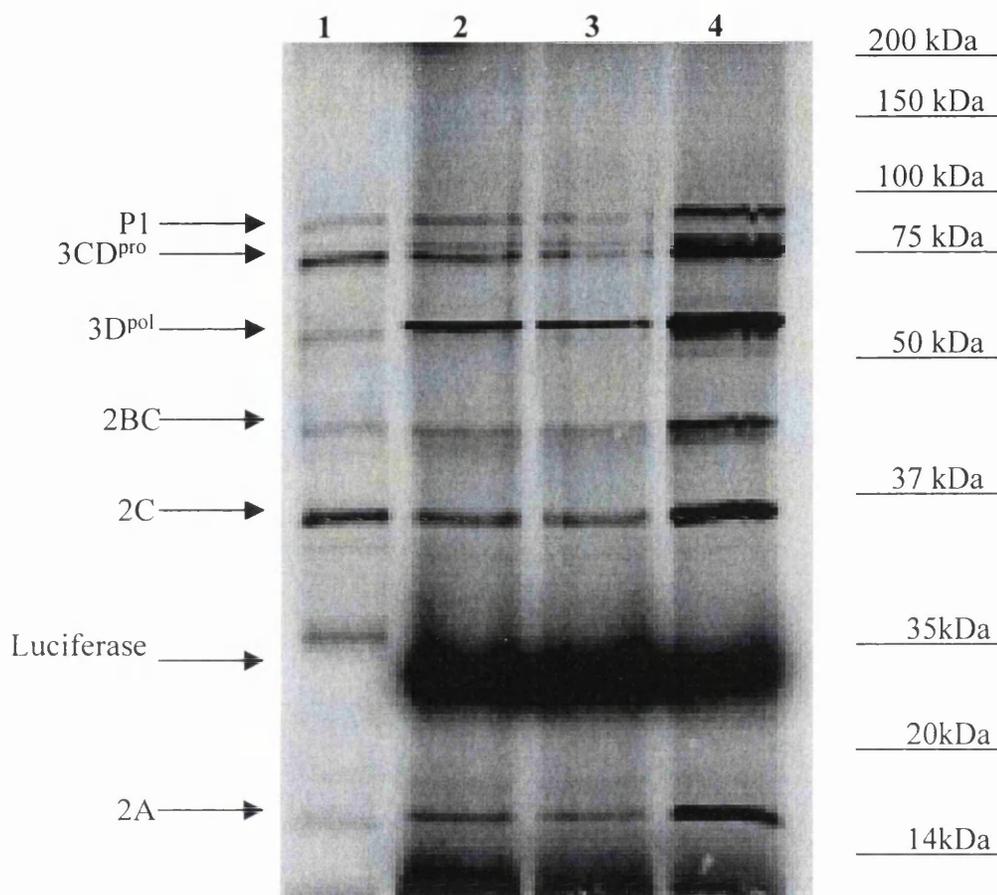


Figure 3.17 *In vitro translation and processing assay.*

Transcribed RNA was translated in rabbit reticulocyte supplemented with HeLa S10 extracts. The translated products derived from RNA transcribed from the luciferase replicons pT7Luc3D^{NdeI}+R and pT7Luc3D^{NsiI}+R are shown in Lanes 3 and 4 respectively. The translated products derived from the parental cDNA pT7Luc+R was used as a positive control (Lane 2). As a further control a sample of S³⁵-methionine labelled poliovirus type 3 infectious lysate was also run on the SDS-PAGE gel (Lane 1)

(Figure 3.17 Lane 4) with that observed with the parental polyprotein (Figure 3.17 Lane 2). The *in vitro* translation and processing reaction appears to suggest that luciferase is produced in greater abundance from the transcripts than the other protein products. However, in picornaviruses the RNA is translated as a single transcript and the encoded proteins are released from the resultant polyprotein via a proteolytic processing cascade. It is therefore infeasible that luciferase is produced in greater amounts than the virus proteins instead it most likely reflects an artefact of the system of labelling the protein products via the incorporation of S³⁵-methionine.

3.8.2.2 Translation and processing of RNA derived from full-length cDNAs pT7FLC^{NheI} and pT7FLC^{Apol}

Whether RNA derived from the pT7luc3D^{NheI}+R and pT7luc3D^{Apol}+R was capable of being correctly translated and processed was not analysed using the *in vitro* translation and processing reaction. Instead these mutations were assessed using the *in vitro* translation and processing reaction in the context of the virus genome. The translation and processing of RNA derived from the infectious cDNA pT7FLC was used as the positive control as to whether the viral proteases 2A^{pro}, 3C^{pro} and 3CD^{pro} were able to correctly process the polyprotein derived from full-length cDNAs bearing mutations in structure I (pT7FLC^{NheI}) or structure II (pT7FLC^{Apol}). As a negative control the HeLa S10 supplemented rabbit reticulocyte lysate extract was incubated in the absence of RNA. In the absence of RNA no radiolabelled products were observed (Figure 3.18 Lane 1) confirming that the bands observed in lanes 3-5 are all products of the translation of the virus RNA. Analysis of the results of translation shows that translation occurred efficiently in all three clones. However closer inspection of the products resulting from polyprotein processing by the viral proteases indicates that the viral polymerase (3D^{pol}) and the polymerase precursor (3CD^{pro}) in pT7FLC^{Apol} (Lane 4 Figure 3.18) were smaller in size than the counterpart bands in clones pT7FLC and pT7FLC^{NheI} (Figure 3.18 Lanes 5 and 3). Although this suggests the presence of a truncation within the region of the polymerase, sequence analysis clearly showed the absence of a frame-shift leading to the introduction of a stop codon. However, the sequence of pT7FLC^{Apol} indicated that, the earliest version of the computer program designed to introduce synonymous changes to the RNA sequence without affecting the reading frame, had through a programming error introduced 6 coding changes to the viral polymerase. The coding changes that were introduced to the viral polymerase were M₃₅₄L, P₃₅₆K, A₃₅₇T, D₃₅₈N, S₃₆₀F, and A₃₆₁T. The number given for each coding change refers to its position in the polymerase i.e. where 1 is the first codon of the virus polymerase following cleavage from the polyprotein. The introduction of the coding changes is postulated to alter the isoelectric point (PI) of the viral proteins 3D and 3CD, thereby causing a difference in behaviour in comparison to the wild-type protein during SDS-PAGE electrophoresis.

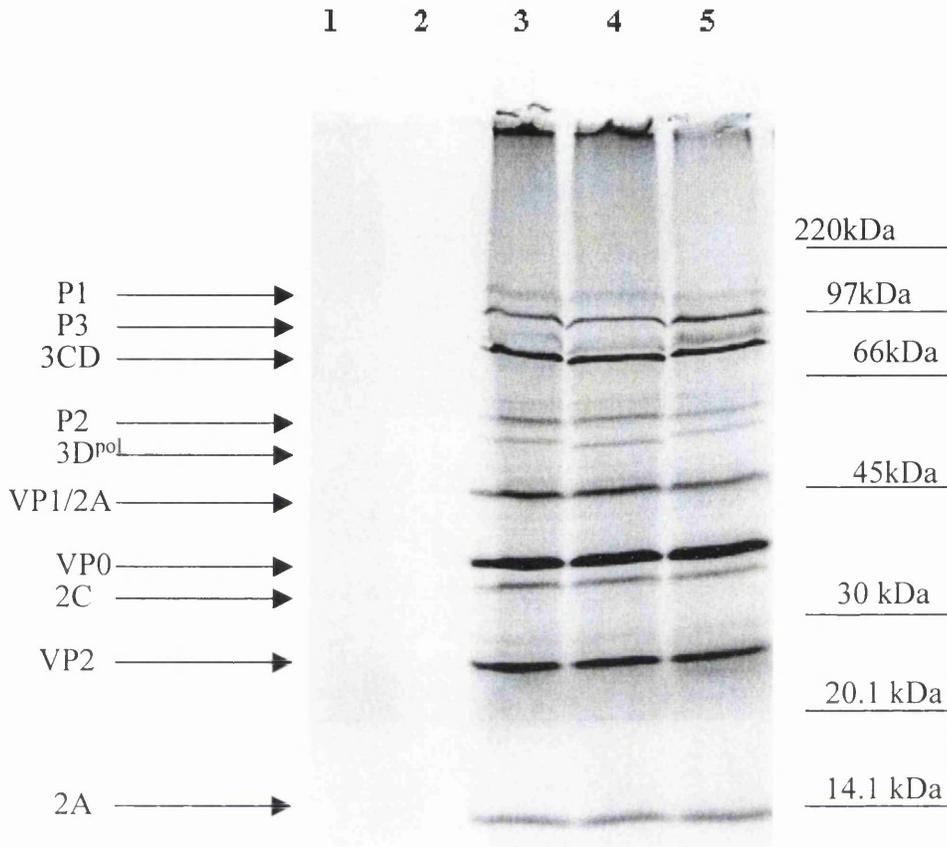


Figure 3.18 *In vitro translation and processing assay.*

Transcribed RNA was translated in rabbit reticulocyte supplemented with HeLa S10 extracts. The translated products derived from RNA transcribed from the infectious clones pT7FLC^{NheI} and pT7FLC^{Apol} are shown in Lanes 3 and 4 respectively. The translated products derived from the parental infectious clone pT7FLC was used as a positive control (Lane 5). As a negative control the supplemented rabbit reticulocyte were incubated in the absence of RNA (Lane 1). Lane 2 was left empty.

The introduction of coding-changes that inactivate the viral polymerase explain why RNA derived from pT7luc3D^{Apol}+R was demonstrated to be replication-incompetent but RNA derived from a different structure II mutant (pT7luc3D^{Nsil}+R) was replication-competent.

3.9 Infectivity assay

In order to investigate whether structure I controlled an aspect of the virus lifecycle not tested by the subgenomic replicon system i.e. uncoating or packaging of the viral genome, the ability of the virus RNA to successfully infect mammalian cells was tested. Consistent with the results of the subgenomic replication assay and the translation and processing assay virus (P3/leon/37^{Nhel}) could be successfully recovered from transcripts synthesised from pT7FLC3D^{Nhel}. In accordance with the previous data no virus could be recovered from transfection into cells of RNA transcribed from pT7FLC3D^{Apol} (data not shown).

Initially infectivity assays were carried out in HeLa cells but inconsistencies in the plaquing of both P3/leon/37^{Nhel} and P3/leon/37 led to the use of a different cell line for this assay. Due to time limitations, the repetition of the infectivity assay using P3/leon/37^{Nhel} and P3/leon/37 in Rhabdomyosarcoma (RD) cells was carried out by Dr.H.Harvala. Normalised levels of RNA, as determined by C¹⁴ incorporation, transcribed from linearised pT7FLC or pT7FLC3D^{Nhel} were diluted 10 fold serially in dH₂O. The diluted RNA was transformed into RD cells using DEAE-dextran. The plaque morphology of the virus and its titre was determined after two days growth by staining the cells with crystal violet. The plaque phenotype of P3/leon/37^{Nhel} (Figure 3.19B) was uniformly smaller (diameter of ~0.5 mm) than that of the parental virus (diameter of ~1 mm) (Figure 3.19A), a phenotype consistent with a replication defect. Calculation of the titre showed that P3/leon/37^{Nhel} was reduced in titre in comparison with the parental virus by approximately 0.5log₁₀. These results are consistent with the data obtained using the subgenomic replicon based system. Subsequently Dr.H.Harvala has shown that the same plaque phenotype and difference in titre between P3/leon/37^{Nhel} and P3/leon/37 that was obtained in RD cells can be observed when the same virus RNAs are analysed in Vero cells (data not shown).

The stability of the mutations introduced, into structure I was investigated by passaging P3/leon/37^{Nhel} on HeLa cells. Plaque assays carried out on the harvested virus showed that the plaque phenotype of P3/leon/37^{Nhel} remained stable after 3 passages (Figure 3.20 A+B) though the plaque phenotype was mixed. The plaque phenotype that was observed in RD cells was smaller and more uniform in nature than that observed in HeLa cells for P3/leon/37^{Nhel}. The difference in plaque phenotype of both P3/leon/37^{Nhel} and the parental virus observed between HeLa and RD cells may have arisen in part from the different overlay medium (2% bactoagar and 0.5% carboxymethylcellulose diluted in 5% FCS/DMEM) that was used.

Recent characterisation of P3/leon/37^{Nhel} by Dr.H.Harvala has shown that the replication defect is temperature-sensitive. At 39 °C replication of P3/leon/37^{Nhel} was reduced by 2 log₁₀ when compared with the parental virus P3/leon/37 (Figure 3.21A). The reduction in growth of 2 log₁₀ of P3/leon/37^{Nhel} at 39 °C was observed over a range of starting M.O.I's (0.1, 1 and 5) (Figure 3.21B).

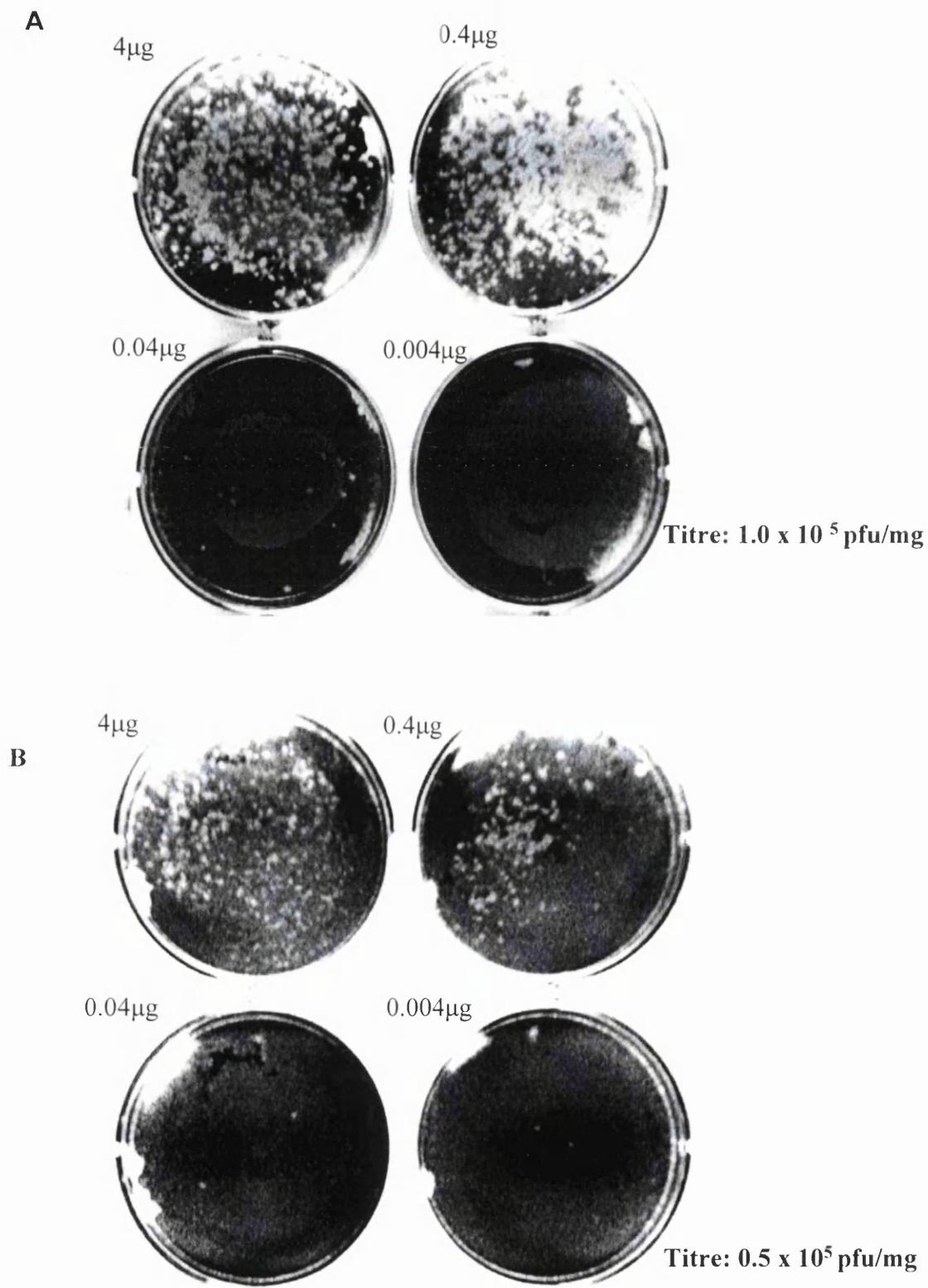


Figure 3.19 Infectivity assays stained after 48 hours incubation at 37 °C..

Ten-fold dilutions of **(A)** P3/Leon/37 (type 3) and **(B)** P3/Leon/37^{Nhel} (type 3) were used to transfect RD cells. The infection was allowed to proceed for 48 hours before the cells were stained with crystal violet.

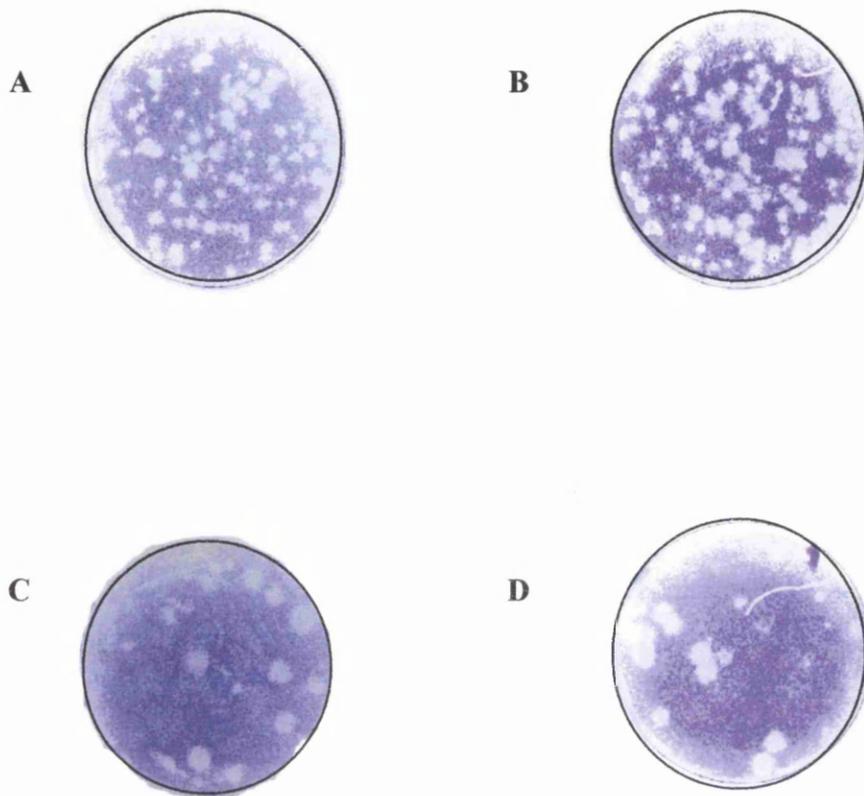
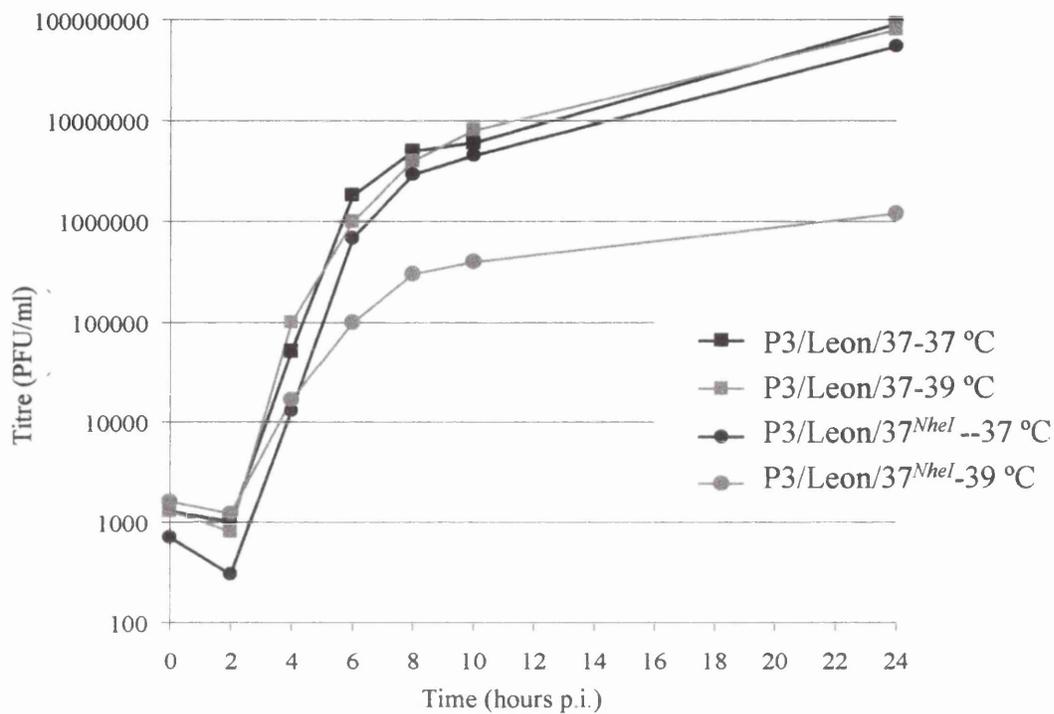


Figure 3.20 Stability of mutations

Poliovirus P3/Leon/37^{Nhe1} was passaged three times in HeLaT4 cells. The plaque phenotype of poliovirus P3/Leon/37^{Nhe1} (A) at pass 0. The plaque phenotype of poliovirus P3/Leon/37^{Nhe1} at pass 3 (B). Poliovirus P3/Leon/37 at pass 0 and at pass 3 (C and D respectively) are shown for comparison. Staining of the cell sheet was carried out 48 hours after virus had been applied to cell layer.

A



B

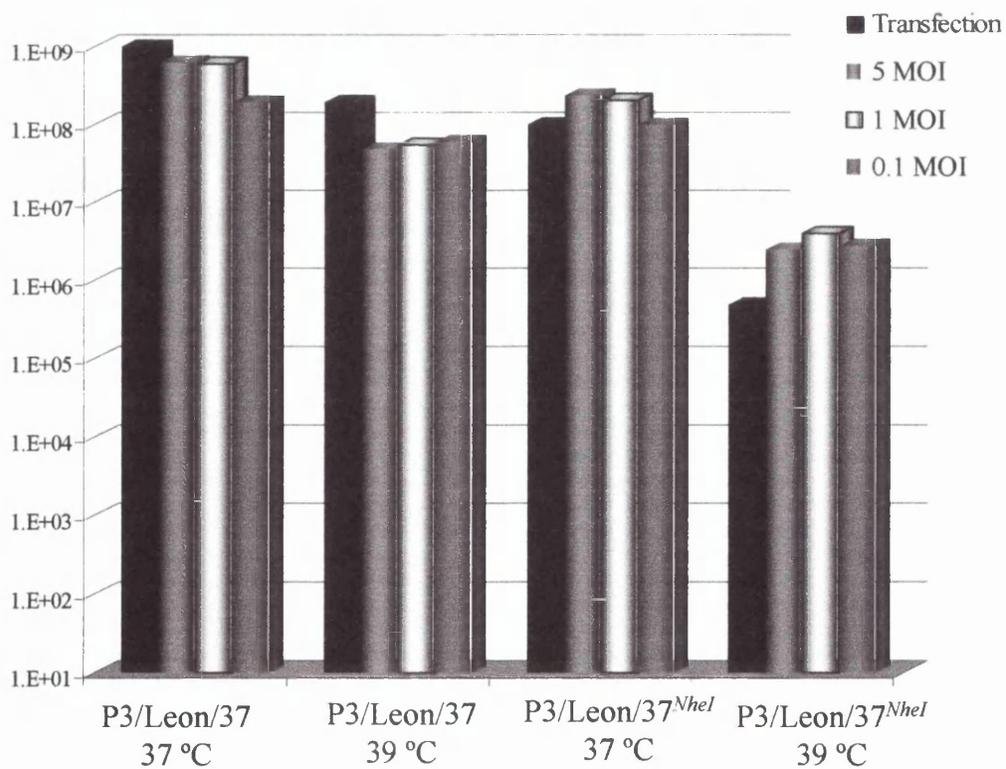


Figure 3.21 Temperature-sensitive growth phenotype of P3/Leon37^{Nhel} in RD cells. (A) One step growth curves of P3/Leon/37 and P3/Leon/37^{Nhel} at 37 and 39° C from a starting M.O.I of 10. (B) Titres of P3/Leon/37 and P3/Leon/37^{Nhel} at 37 and 39° from a starting M.O.I of 0.1, 1 and 5.

3.10 Discussion of results

Evidence has steadily accumulated that though RNA viruses with a positive-sense single-stranded genome are a highly diverse group a basal replication strategy exists that has been preserved from a common progenitor. The use of *cis*-acting elements to signal functions like initiation of genomic replication is a phenomenon associated with many viruses including the coronaviruses (Fosmire et al., 1992, Repass & Makino, 1998), and tombusviruses (Park et al., 2002, Pogany et al., 2003). Within the *Picornaviridae* three *cis*-acting elements have previously been identified to be essential for efficient genomic RNA replication: cloverleaf, CRE and 3'UTR. A further *cis*-acting element, the IRES, functions in the initiation of polyprotein synthesis.

The use of SSSV as an indicator of the presence of RNA signals of functional importance has previously been reported for HCV (Tuplin et al., 2002, Walewski et al., 2001) and poliovirus (Tuplin et al., 2002). Tuplin *et al* were able to demonstrate that in both poliovirus and HCV areas of SSSV correlated well with thermodynamic predictions that are an established method of predicting areas of localised secondary structure (Tuplin et al., 2002). In their analysis of poliovirus Tuplin *et al* were able to identify a strong region of SSSV centred on the CRE, which is the only known essential *cis*-acting RNA element to be located in the ORF of the poliovirus genome. The identification of a well-characterised *cis*-acting replication element by SSSV provided evidence that this type of analysis was capable of identifying RNA sequences or structures of functional importance from aligned sequences. Further support for this view has been supplied recently by a research group investigating the mechanics of HCV replication. Tuplin *et al* described a region in the gene encoding the virus RNA-dependent-RNA polymerase (NS5B) that showed strong SSSV (Tuplin et al., 2002). This analysis identified two new stem-loops 5BSL1 (nt 111-128) and 5BSL2 (nt 333-342) as well as three structures that had previously been identified by Smith and Simmonds using sequence inspection methods (Smith & Simmonds, 1997). The structures identified by Smith and Simmonds: 5BSL3.1, 5BSL3.3 and 5BSL3.4 constitute part of a cruciform structure (nt 539-591) which includes a fourth stem-loop structure 5BSL3.2, that has recently been identified as a context dependent essential HCV *cis*-acting replication element (You et al., 2004). Concrete evidence therefore exists that SSSV provides a good predictor of functionally important RNA sequences or structures.

The investigation, presented in the course of this chapter, was concerned with assessing the functional relevance of an area of SSSV in the gene encoding the poliovirus polymerase that was initially identified by Tuplin *et al* (Tuplin et al., 2002). The effect of introducing destabilising mutations into structures within this region was analysed using both subgenomic replicon and viral systems. In the subgenomic replicon system a reduction in the level of replication, compared to the parental replicon, of approximately $0.5\log_{10}$ was consistently observed. The growth phenotype of virus recovered, that contained the same nucleotide changes, was comparable with that observed in the subgenomic replicon system. Thus, while the data from the virus characterisation studies strongly argues against the RNA secondary structure under investigation, functioning as an uncoating or packaging signal it does provide support for the argument that it contains a determinant of replication. There are a number of possible explanations for why only a $0.5\log_{10}$ decrease in replication was observed, which will be discussed below.

Investigations into the replication phenotype of the mutagenized cDNAs over time using a luciferase reporter gene system suggested that the observed difference between the parental cDNA and the mutagenized cDNA pT7luc3D^{Nhel} was as a result of an effect within the first six hours of initiation of genomic RNA synthesis. No difference was observed between the parental cDNA and the mutagenized cDNA pT7luc3D^{Nhel} at 8 hours post transfection. Due to the sensitivity associated with the luciferase reporter gene (in the range of 8 orders of magnitude) this is unlikely to be due to saturation of the assay. One explanation of the difference in phenotype observed over time between the parental cDNA and pT7luc3D^{Nhel} was that the RNA differed in stability. In this model the mutagenized RNA would be more sensitive to the degradative enzymes present within the cell. In a cell infected with poliovirus the replication of genomic RNA occurs exponentially between 4-6 hours post-infection. At 6-8 hours post infection the priority of the virus switches from replication viral genome to assembly of the virus particle as the ability of the cell to support translation rapidly decreases.

The ability of the virus to support translation occurs as a result of the cleavage by the protease 2A of eIF4E, the cap-binding factor, from the transcription factor eIF4G. This cleavage “shuts-down” the host cell by preventing the translation of the majority of cellular mRNAs as a consequence of which cellular proteins are degraded and not replaced. The accumulated effect of the degradation of host cell proteins is that the cell gradually loses the ability to provide the host cell factors that poliovirus requires to support the translation and replication of its genome. As there is an inherent requirement for the genome of poliovirus to be translated before it can be replicated it would be expected that a subtle decrease in the stability of the genomic RNA would only be observed at the peak of the cycle of translation and replication. Between 6 and 8 hours the cell will no longer be capable of supporting the translation and replication of the virus.

One unusual feature of the replication of poliovirus is the selective inhibition of nuclear import via the degradation of the nuclear pore complex (NPC) proteins Nup 153 and p62 (Gustin & Sarnow, 2001). These proteins form what is termed the classical import pathway. In 2001 Gustin and Sarnow demonstrated that the glucocorticoid receptor import and CRM1 export pathway remain functional in infected cells (Gustin & Sarnow, 2001). The inhibition of the classical import pathway in cells leads to the cytoplasmic accumulation of cellular factors that are predominantly nuclear in uninfected cells. The cellular factors that have been shown to accumulate in the cytoplasm of infected cells are summarised in table 3.2 along with their reported cellular functions.

Among the cellular factors that are accumulated in infected cells are the La autoantigen and PTB that are both known to be involved in translation of the viral genome (Hellen et al., 1993, Meerovitch & Sonenberg, 1993). In addition Sam68 and Nucleolin have been shown to interact with the viral polymerase and the 3'UTR of poliovirus respectively (McBride et al., 1996, Waggoner & Sarnow, 1998). Although functions have been described for La and PTB, no definitive functions in viral replication have been ascribed to the remainder of the proteins known to accumulate during the course of infection. Whilst inhibition of nuclear import has been postulated to provide a mechanism by which the virus can down regulate the antiviral interferon response (Gustin, 2003) this is unlikely to be the only benefit of inhibiting nuclear import.

Nuclear protein	Cellular functions ^a	Viral targets	Function in viral replication	References ^b
Nucleolin	rRNA transcription and processing, pol II transcription, m RNA stability	3'UTR of poliovirus genome	RNA synthesis?	Waggoner and Sarnow (1998) Gustin and Sarnow (2002)
hnRNP K	Translational silencing, mRNA stability	?	?	Gustin and Sarnow (2001 and 2002)
PTB ^c	Splicing, polyadenylation	IRES of entero-, rhino-, and cardiovirus genomes	Translation	Back et al (2002)
La	Pol III transcription, processing	IRES of entero-, rhino-, and cardiovirus genomes	Translation	Meerovitch et al (1993), Gustin and Sarnow (2002)
hnRNP A1	mRNA transport, splicing	?	?	Gustin and Sarnow (2001 and 2002)
hnRNP C	Splicing, mRNA stability	?	?	Gustin and Sarnow (2001 and 2002)
Sam68	Cell cycle control? mRNA transport	3D ^{pol} of poliovirus	?	McBride et al (1996) Gustin and Sarnow (2002)

Table 3.2 Cellular nuclear protein that accumulate in the cytoplasm of poliovirus and rhinovirus cells.

Taken from Gustin, K.E., 2003. ^a reported functions of indicated protein in uninfected cells ^b Reference demonstrating nuclear accumulation of nuclear protein in infected cells ^c PTB has not been shown to accumulate in the cytoplasm of rhinovirus-infected cells

The cellular proteins that accumulate as a consequence of inhibition of nuclear import all have known functions in mRNA transport, mRNA stability, translation or polyadenylation (Table 3.2). As poliovirus has a limited coding capacity it is probable that these proteins have functional roles in the replication of the virus.

The cellular proteins Nucleolin, hnRNP K and hnRNP C have all been reported to be involved in stabilisation of mRNA. The interaction of Nucleolin with the 3'UTR of poliovirus has previously been documented (Waggoner & Sarnow, 1998). As of yet no interaction has been described between hnRNP K and C with a viral protein or RNA structure. However, it is probable that these proteins are involved in the stabilisation of the viral genome through interaction with RNA secondary structure or sequences present within the genome. Disruption of this interaction would correlate with a decrease in the infectivity of the virus as a consequence of a reduction in translation and replication of the viral genome. *In vivo* this phenotype would be expected to result in a less severe pathogenic outcome. As these cellular proteins are ubiquitously expressed the use of these proteins to stabilise the viral genome, and the sequence or structure that would facilitate the interaction would presumably be a highly conserved mechanism used by picornaviruses.

A second explanation for the phenotype observed in the replication time course would be that in the RNA containing mutations in structure I the defect affected the initiation of RNA synthesis but not the elongation stage of the process. One possibility with regards to the apparent defect in initiation observed with the cDNA pT7luc3D^{Nhel} is that the mutations involved affect the ratio between positive and negative strands. It is probable that *cis*-acting replication elements that regulate the process of genomic RNA synthesis are located within the genomes of all positive single-stranded RNA viruses. Genome replication in positive single-stranded RNA viruses occurs in two distinct stages. Initially there is the production of a negative strand complementary to the incoming genome. The synthesis of an anti-genome strand results in the formation of a double-stranded replication intermediate (RF) that is subsequently used as a template for the synthesis of large amount off positive single-stranded RNA strands. In a poliovirus-infected cell the ratio is approximately 50:1 ratio in favour of the positive-sense strands to the negative-sense RNA strands (Novak & Kirkegaard, 1991). This is consistent with a model in which numerous positive strands are synthesised of the RF template. Amongst the tombusviruses this bias is approximately 100:1 in favour of the synthesis of positive-sense strands. Recent work on

the tomato bushy stunt virus (TBSV), the prototypic tombusvirus, (Pogany et al., 2003) has identified a small RNA element located at the 3' terminal end of the viral genome that negatively regulates the initiation of minus strand synthesis. The down-regulation of negative strand synthesis thus biases RNA synthesis towards the accumulation of positive-sense strands. Pogany *et al* propose that the presence of a “replication silencer” element in the genome of TBSV is beneficial to the virus as it can minimize the wasteful production of non-infectious template. In addition Pogany *et al* proposed a second benefit of replication silencing. As a consequence of the downregulation of negative sense strands the level of RF templates within the cell would also be reduced. Pogany *et al* propose that in TBSV infections the reduced amount of RF templates delay recognition of the infection by the host post-transcriptional gene silencing (PTGS) response (Vance & Vaucheret, 2001, Waterhouse et al., 2001). The PTGS response has a similar outcome to the double-stranded RNA-induced protein kinase R (PKR) response in mammalian cells. A delay in the onset of these antiviral responses would enable the virus infection to take hold of the cell. In the model proposed by Pogany *et al* an increase in the yield of RF templates would increase the rapidity of the innate immune response to the viral infection. One consequence of this may be a reduction in the pathogenesis of the virus. Any reduction in pathogenesis could be tested using the transgenic mouse models available to study poliovirus pathogenesis (Ren et al., 1990).

If the *cis*-acting replication element identified within this study was to function as a replication silencer then destabilisation of this element through the introduction of nucleotide changes would not necessarily introduce a dramatic replication phenotype *in vitro* as determined by reporter gene quantification. This is due to the requirement of the poliovirus RNA to be translated prior to its use as a template for genomic synthesis. Alternatively, it therefore remains possible that a severe replication phenotype was not observed because the mutations that were introduced into the cDNAs to disrupt structure I or structure II did not fully prevent a long-range RNA-RNA interaction, constituting the poliovirus “replication silencer” from occurring. After all, a known limitation of using computational analysis is that the thermodynamics favouring formation of tertiary structure or long-range interactions have not been clearly elucidated (reviewed in Tinoco & Bustamante, 1999). If either of these scenarios is correct an alteration in the ratio of genomic and anti-genomic strands synthesised by P3/leon/37^{Nhel} in comparison with the parental virus should be detectable using radiolabelled oligonucleotide probes.

The method by which replication silencing functions works in tombusviruses is by promoter masking (Figure 3.22). In poliovirus infected cells, if RNA silencing is used to promote asymmetric synthesis an RNA-RNA interaction would have to occur between this element (region encompassing structure I and structure II) and the Ori R, which is involved in the initiation of negative-strand synthesis. The 3'UTR of all enteroviruses appears to fold into either two or three stem-loop domains in which part of the poly-A tract is included (Figure 1.21). This core structure (Pilipenko et al., 1992b) has been supported by thermodynamic (Jacobson & Zuker, 1993, Melchers et al., 1997) and phylogenetic studies (Pilipenko et al., 1992b). However research using chimeric viruses has strongly suggested that the 3'UTR of all enteroviruses forms a common single stem-loop tertiary structure (Rohll et al., 1995) that arises as a result of an intramolecular "kissing interaction" between the stem-loop domains of the 3'UTR (Mirmomeni et al., 1997, Pilipenko et al., 1996). Consistent with a role in the regulation of picornavirus replication is the observation that virus, though severely compromised, can be recovered that contains a complete deletion of the 3'UTR (Todd et al., 1997). How the 3'UTR mediates the efficiency of virus replication remains unclear, though one possibility is that it functions to negatively regulate the production of anti-genomes. To investigate the possibility that an interaction between structure I or structure II and the 3'UTR might function as a replication silencer, the sequences of these elements was analysed for regions of complementarity. The search for complementarity between the elements was concentrated on the sequences of the loops as it was believed that these were the likeliest regions of the structures to be involved in any action. Figure 3.23 shows that no regions of sequence complementarity exist between the loop structures of the 3'UTR and either structure I or structure II. In my opinion, this clearly suggests that if this element is involved in replication silencing it makes it unlikely that poliovirus employs a promoter masking mechanism.

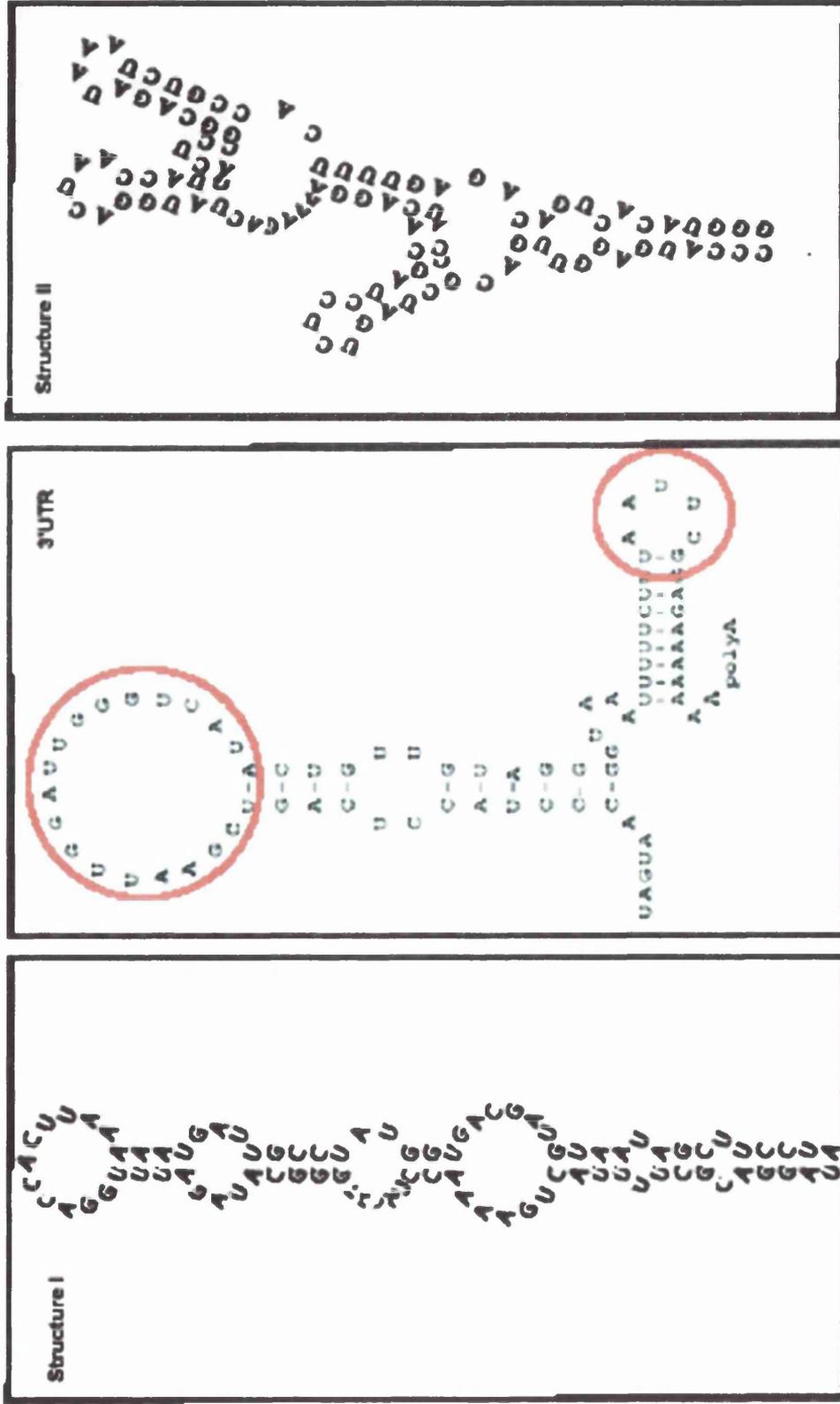


Figure 3.23 Comparison of the sequences of the 3'UTR with Structure I and II. Loops of 3'UTR available for interaction with RNA are highlighted in red. No complementary sequences (GCUUAACCUAACCCAGUUAU OR UUAAG) are found in loops of Structure I or II.

3.10.1 Identification of a species-specific sequence

The genomic sequences were re-evaluated, following the identification of a subtle phenotype, to identify whether the changes that were introduced might lie up or downstream from a significant sequence or structure that may have been missed during the initial analysis. Re-analysis of the region using MFOLD of individual enterovirus sequences identified a small stem loop structure that was present in all strains of the poliovirus and non-poliovirus HEV-C viruses analysed (Figure 3.24 A and B). The small stem-loop structure for the purposes of this thesis has been labelled motif X. Figure 3.25 shows that the sequence of motif X was 100% conserved in coxsackievirus A21 (accession number NC 001428), coxsackievirus A24 (accession number CXA24CG) and a number of other HEV-C with respect to the sequence present in poliovirus type 3. However analysis by eye suggested that the sequence was absent in viruses, more distantly related to poliovirus and coxsackie A viruses (HEV- C) that are classified within the HEV-B species. To determine whether the sequence/structure was a species-specific determinant, SSSV analysis was carried out on the HEV- C and HEV- B species independently. The results as illustrated in Figure 3.26 show that whilst the CRE is identifiable in both the HEV-B and C species the region of SSSV in the viral polymerase is only observed in the HEV-C species viruses. In addition the region of SSSV that has been studied in this chapter (purple arrow centred on nucleotide 6467 in figure 3.26) amongst the HEV-C shows a greater degree of SSSV than the CRE. This indicates that the region of SSSV must be of great importance to the HEV-C.

Motif X is situated between the two major stem-loop structures predicted by the MFOLD analysis. In 2001 Witwer *et al* demonstrated using computer algorithms (alidot and pfrali) that search large RNAs for secondary structures using a combination of thermodynamic structure prediction and phylogenetic analysis, that no structural feature apart from the CRE was shared among the picornaviruses. In the computational analysis undertaken by Witwer *et al* the potential for the formation of structures within the coding region of the viral polymerase in all picornavirus genera was observed. In six of the genera the potential structure lies at the end of the coding region of the polymerase adjacent to the 3'UTR (Figure 3.27). According to the analysis carried out by Witwer *et al* a structure of significance would be predicted to form at nt 7414-7430 of the enteroviral viral genome (Witwer *et al.*, 2001). This structure is downstream of the sequence identified using SSSV

analysis in this study. The data presented by Witwer *et al* does raise an important question regarding whether the SSSV analysis has identified an important species-specific sequence rather than an important RNA structure. This is a definite possibility. The SSSV algorithm is calculated independently of considerations for the formation of RNA secondary structure. As a consequence though SSSV analysis can be demonstrated to identify known RNA structures of functional importance i.e. the CRE, the analysis will also identify sequences of functional significance with the same frequency. If the motif identified by SSSV analysis was a sequence-specific recognition motif for a protein then disruption of the surrounding structure would only affect the function if the important sequence was masked, or the non-coding changes introduced altered critical residues of the sequence motif.

The identification of a putative species-specific replication determinant within the genome of the HEV- C species raises some interesting questions regarding its function. Although it has been proposed previously in this discussion that the element described in this chapter could function as a “replication silencer” and therefore promote asymmetric replication this is at odds with the identification of motif X as a species-specific sequence. Given the degree of similarity within the picornavirus family a shared mechanism for this feature of RNA synthesis would be expected. The identification of a species-specific sequence is also at odds with the case presented for a possible role in genomic stabilisation via the binding of ubiquitous cellular factors.

To investigate the function of motif X the aspects of the enterovirus replication cycle that could possibly be influenced by a species-specific sequence need to be considered. To my mind there are three possible areas, which will be discussed in turn. These are pathogenesis, RNA recombination and membrane targeting.

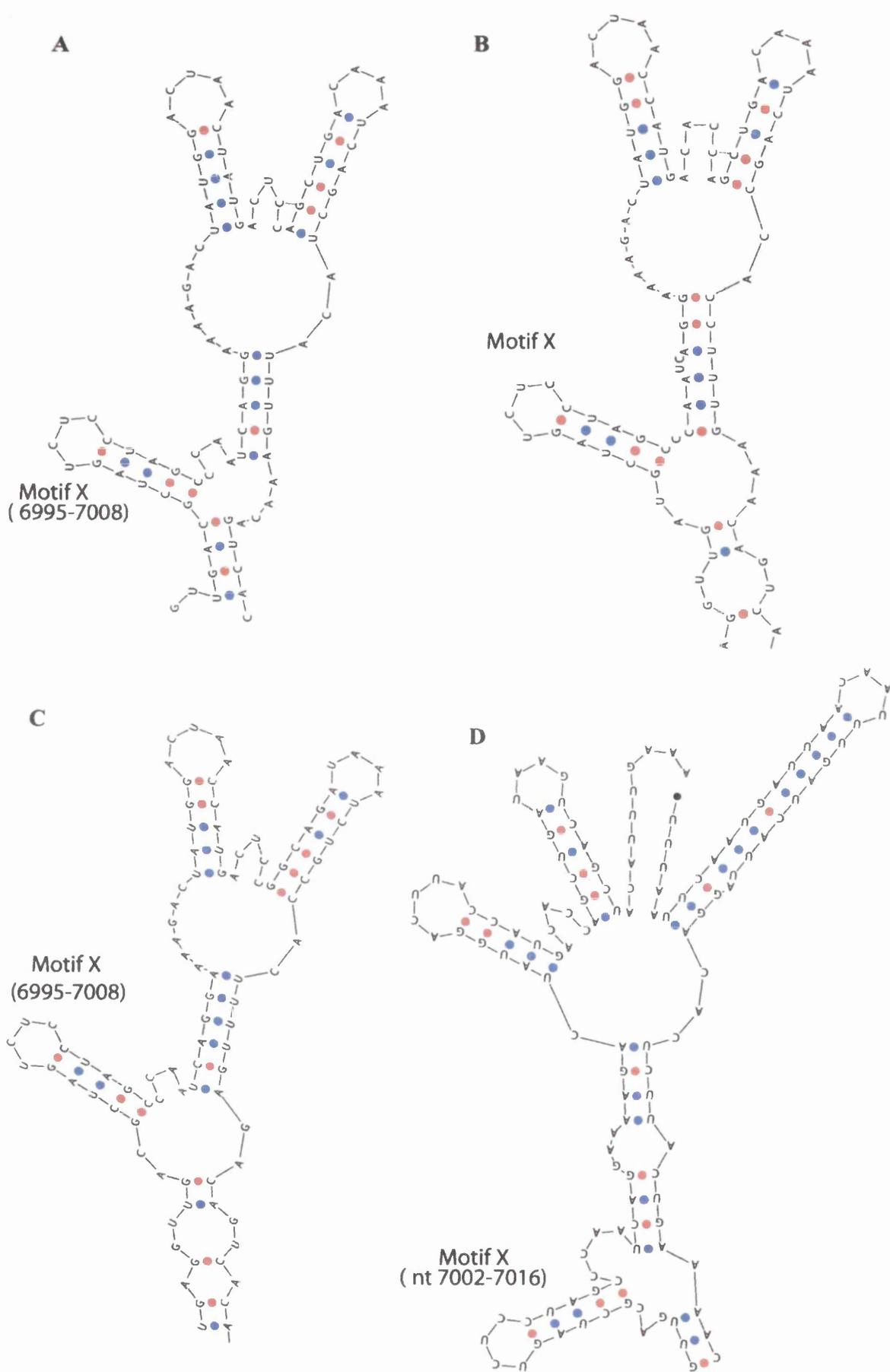


Figure 3.24 Structure of motif X as determined by MFOLD analysis
 (A) poliovirus type 1 (Mahoney), (B) poliovirus type 2 (Lansing),
 (C) poliovirus type 3 (Leon) and (D) coxsackievirus A24.
 MFOLD analysis was carried out using the MFOLD server.

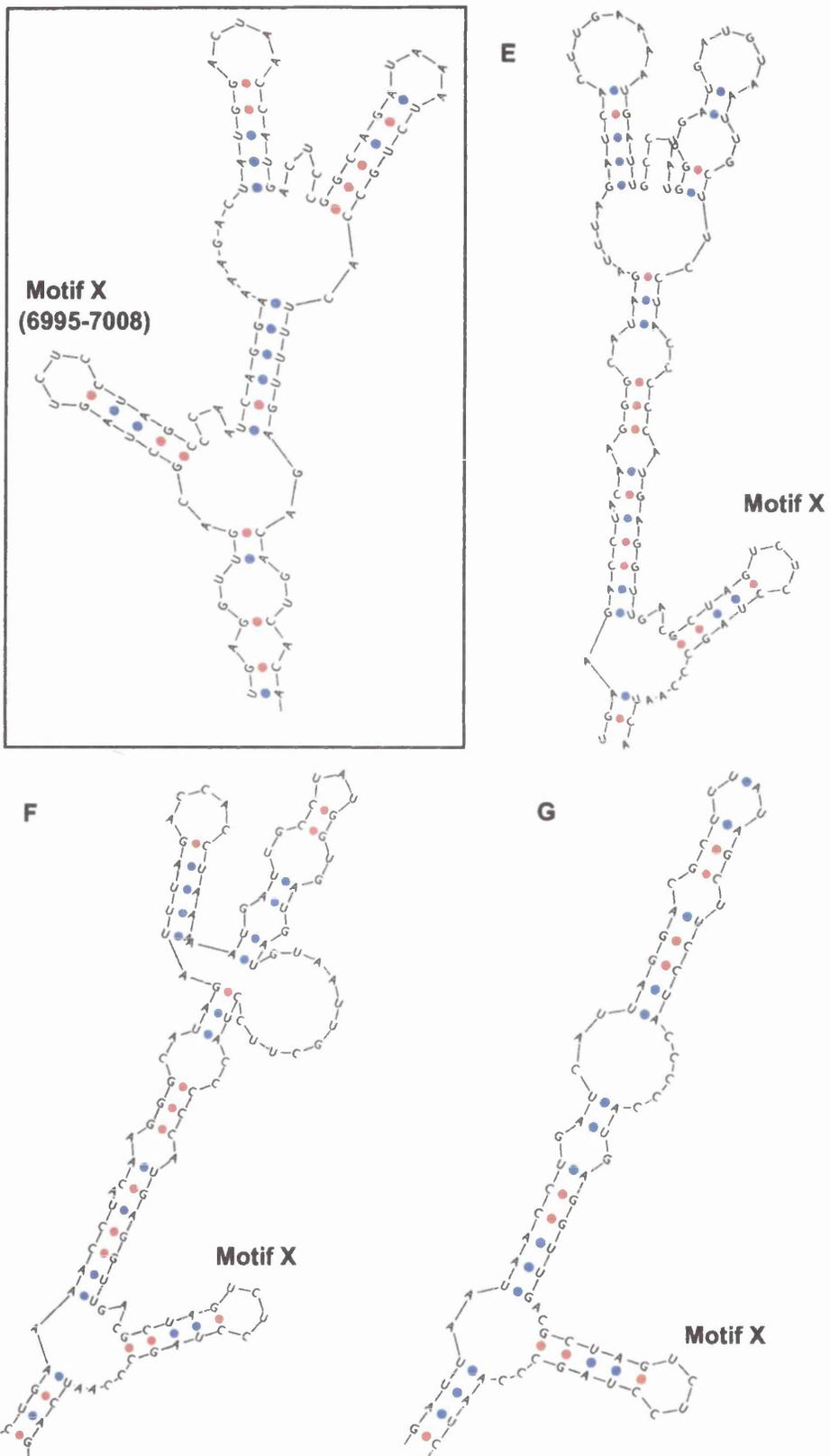


Figure 3.24B Structure of motif X as determined by MFOLD analysis
 (E) Coxsackievirus A13, (F) Coxsackievirus A20 and (G) Coxsackievirus A21.
 Poliovirus type 3 (Leon) is shown as an inset. MFOLD analysis was carried out using the MFOLD server.

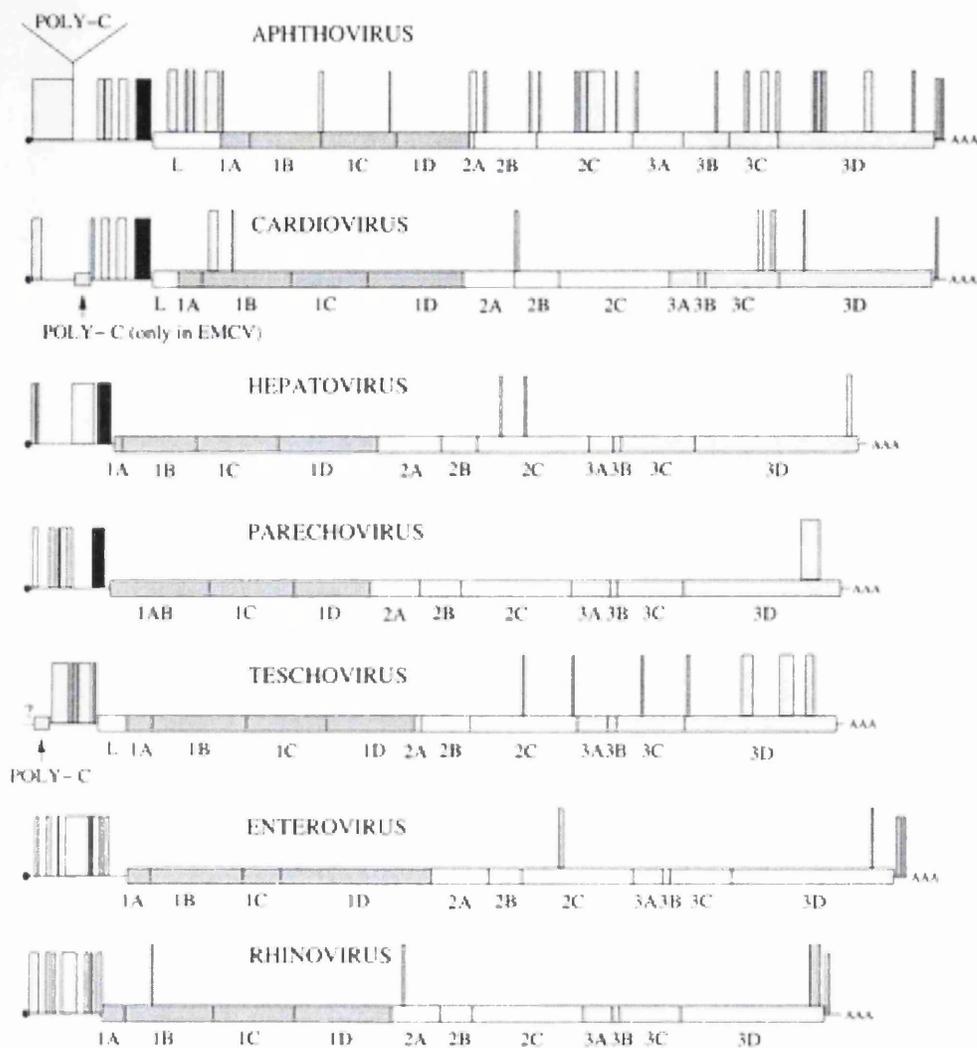


Figure 3.27 Overview of Picornavirus genomes. Putative conserved RNA elements are indicated above the diagrams of the reading frames. The black boxes indicate the J,K-element. Proteins: leader protein L (only present in aphthovirus, cardiovirus and teschovirus), capsid proteins 1A–1D, viral protease 2A, proteins involved in RNA synthesis 2B, 2C, unknown function 3A, VPg 3B, major viral protease 3C, RNA-dependent RNA-polymerase 3D. Figure 3.27 taken from Witwer et al, 2001.

3.10.1.1 Motif X as a determinant of viral pathogenesis?

One aspect of the viral lifecycle that can be influenced in a species-specific manner is pathogenesis. Previous work using chimeric viruses has demonstrated that within the enteroviruses species-specific blocks do exist that limit the tropism of the virus. As has been described elsewhere (see introduction Table 1.3), poliovirus and HEV-C species viruses are causal agents of flaccid paralysis, a disease associated with replication of the virus in the CNS. The HEV-B viruses, on the other hand, whilst known to cause meningitis are not associated with flaccid paralysis (Melnick, 1996a).

Research carried out using chimeric viruses to analyse the effect of the IRES as a virulence determinant in a mouse model of myocarditis showed that a species-specific block existed. A chimera was constructed using coxsackie B3 (HEV-B), the agent responsible for the majority of viral myocarditis, but containing the IRES of poliovirus (Chapman et al., 2000). The chimeric virus was shown to have abrogated the cardiovirulent phenotype *in vivo*. Likewise chimeric poliovirus containing the 5'UTR of rhinovirus 2 has been shown to reduce the neurovirulent phenotype observed with poliovirus (Gromeier et al., 1996). It is important to note that these effects are observed only within certain cell types. Gromeier *et al* demonstrated that whilst significant defects were observed in both neuroblastoma cells and in transgenic models of poliomyelitis the rhinovirus/poliovirus chimera replicated efficiently in HeLa cells (Gromeier et al., 1999). Similar cell-type specific effects were observed when the replication ability of the coxsackie B3 chimeras in different cell-types was investigated. In both these cases the reduction in virulence correlates with the ability of the viral IRES to recruit cellular factors required for IRES function.

In addition to the 5'UTR, the Z domain of the enterovirus B 3'UTR has been shown to play a role in the pathogenesis of coxsackie B3 virus. *In vivo* it was demonstrated that while a recombinant virus lacking the Z domain was able to replicate in the pancreatic tissue, infection of the heart tissue was abortive. In addition it was shown that removal of the Z domain did not affect the ability of the virus to replicate in tissue culture (Merkle et al., 2002).

The first indication that motif X might not be a cell-type specific replication determinant was provided by the temperature-sensitivity studies using P3/leon/37^{N_{hel}}. Dr.H.Harvala observed that a 2log₁₀ decrease in growth of P3/leon/37^{N_{hel}} could be obtained, using a standard cell line, at higher temperatures. More recently Dr.H.Harvala has characterised the growth of a virus containing mutations in motif X. The results obtained have shown that the introduction of mutations into motif X significantly debilitates the growth of the virus in HeLa cells at 37 °C (2 log₁₀ decrease in titre relative to the parental virus) and 39 °C (4 log₁₀ decrease in titre relative to the parental virus). Thus, the accumulated evidence argues against the function of motif X being a determinant of neurovirulence but rather infers that it performs a role fundamental to the replication of the HEV-C viruses.

3.10.1.2 A putative role for motif X in RNA recombination?

Recombination of poliovirus occurs frequently within the human enteric tract. The rate at which recombination happens has been calculated to be 10⁻⁴ per base pair per replication cycle (Jarvis & Kirkegaard, 1992). It has since been postulated that the frequency of recombination provides one way for the virus to maintain general fitness in the face of the high mutagenic rate of the virus polymerase.

Investigations have identified both replicative (Jarvis & Kirkegaard, 1992, Kirkegaard & Baltimore, 1986) and non-replicative (Gmyl et al., 1999) mechanisms of recombination using poliovirus. Experimental data suggests that *in vivo* the predominant mechanism of recombination in poliovirus is replicative. Replicative recombination occurs when the polymerase prematurely terminates during the elongation of the nascent strand. Elongation of the nascent transcript is completed following re-association of the nascent strand-polymerase complex with a different template (Figure 3.28). A consequence of this is that replicative recombination is more commonly referred to as template-switching. Investigations into the process of replicative recombination have suggested that template-switching occurs with high frequency during the synthesis of the anti-genome (Kirkegaard & Baltimore, 1986). Using EM it has recently been identified that in the presence of RNA the poliovirus polymerase forms a highly-ordered lattice structure (Lyle et al., 2002a). Lyle *et al* proposed that the high levels of template-switching could be explained by the presence of multiple templates undergoing RNA synthesis at the same time on a shared polymerase-lattice (Lyle et al., 2002a). The presence of a polymerase lattice ordered in the presence of RNA may also explain the observation made by Duggal *et al* that replication-

deficient genomes were competent to act as a template for recombination (Duggal et al., 1997).

The basic model for template-switching argues that recombination can either be homologous (as shown in figure 3.28) i.e. the polymerase-donor strand complex switches to an analogous position on the acceptor template, or non-homologous i.e. the polymerase-donor strand complex switches to a different position on the acceptor strand thereby leading to a deletion or insertion of genomic material. According to this basic model of recombination the cross-over points are random so homologous or non-homologous recombination should occur with equal frequency. This however is not supported by analysis of poliovirus recombinants which suggests that homologous recombination is favoured (Jarvis & Kirkegaard, 1992, Wimmer et al., 1993). This observation has previously led to the proposal that alignment of parental strands involved in recombination is mediated by RNA secondary structure (Romanova et al., 1986, Tolskaya et al., 1987) or specific sequence motifs (King, 1988). The presence of specific sequences or RNA structures within the genome that promote recombination at specific sites in the genome and between closely related viruses would be advantageous to the virus. Random recombination, of genomes by strand-switching during minus strand synthesis, lays the virus open to the possibility that it may recombine with an incompatible partner. For example, experiments have shown that the proteinase 3C of HRV14 and coxsackievirus B3 can process the non-structural proteins of the respective virus but not the precursor of the structural proteins (Dewalt et al., 1989). The existence of specific sequence or structural recombination motifs would provide a mechanism for preventing any virus from recombining with viruses that are incompatible with itself at the level of protein function. A further advantage to the virus of specific sequence or structural recombination motifs would be the reduction in the probability that the recombinants generated would be replication-incompetent due to deletions or insertions in its genome.

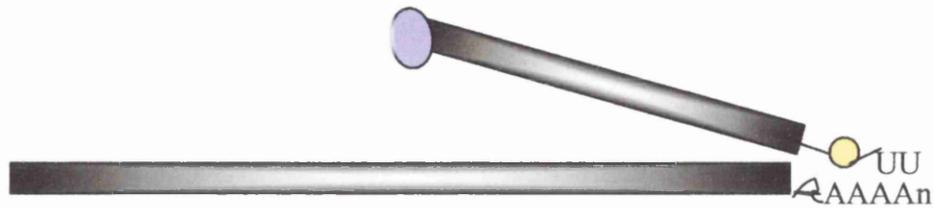
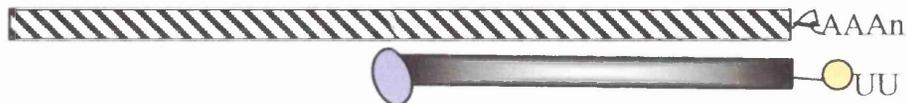
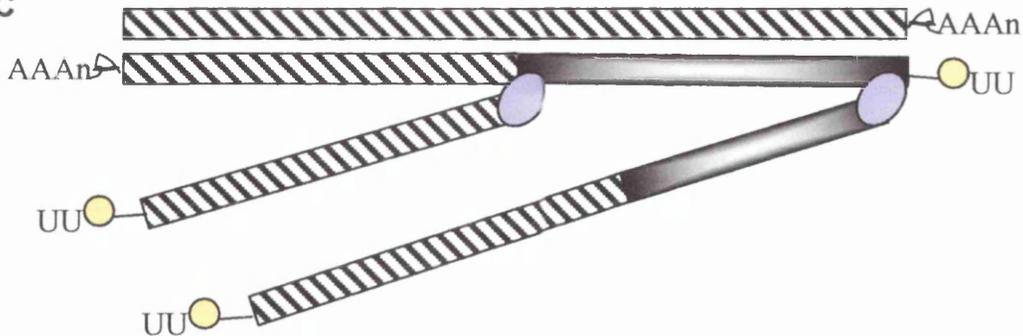
A**B****C**

Figure 3.28 *Strand-switching model for homologous recombination during picornavirus replication.*

The polymerase carrying a nascent negative strand detaches from the donor strand (shaded rectangle) during negative strand RNA synthesis (**A**) and reattaches to a second genomic RNA (**B**), termed the acceptor strand (rectangle with diagonal lines), within the replication complex. For clarity in this diagram the nascent strand is shown diffusing up to the acceptor strand. After hybridization to the acceptor strand RNA synthesis is completed. The resultant full length negative strand is then copied (**C**). Homologous recombination occurs when the polymerase complex switches to the same position on the acceptor template. Heterologous recombination occurs when the polymerase switches to a different position on the acceptor template.

If there is a requirement for secondary structures or specific sequences to form recombination structures then it would be expected that genomic “hotspots” of recombination would exist. Analysis of enterovirus recombinants has shown that the majority of these sites are found within the region of the genome encoding the non-structural proteins. Recent research carried out analysing naturally circulating enteroviruses has shown that recombination occurs frequently in the 3CD region (Lindberg *et al.*, 2003, Lukashev *et al.*, 2003) in HEV-B and HEV-C species viruses (Brown *et al.*, 2003). Amongst the analysed recombinants of poliovirus and HEV-C viruses recombination sites have been identified at nucleotides 6824/6825 of Sabin 3 (Blomqvist *et al.*, 2003) and 6379/7441 of poliovirus 1 (Liu *et al.*, 2003). The region of SSSV, which was analysed for a putative function, was centred on nt 6841 of the full-length genome. This lies just downstream of the cross-over point identified by Blomqvist *et al.* and within the region identified by Liu *et al.* as the crossover point. It can therefore not be ruled out as a possibility that the region of SSSV identified is involved in the alignment of HEV-C species viral genomes prior to recombination. The absence of these sequences in HEV-B species would make it less likely that HEV-B and HEV-C species would become closely enough aligned to make recombination a strong possibility.

Lukashev *et al.* and others have shown that intraspecies recombination amongst the HEV-B enteroviruses exists (Lindberg *et al.*, 2003, Lukashev *et al.*, 2003, Oberste *et al.*, 2004). Like the recombination observed between HEV-C enteroviruses, this occurs predominantly in the region of the genome encoding the non-structural proteins. On the basis of a model where the species-specific sequences are involved in promoting recombination at specific regions of the genome it would seem reasonable to suggest that HEV-B enteroviruses would have evolved sequences that would enable close alignment of HEV-B genomes. SSSV-analysis suggests no such species-specific sequences exist (figure 3.26). In my opinion, this argues against the species-specific motif being a determinant of HEV-C recombination.

3.10.1.3 A putative role for motif X in membrane targeting?

Fluorescent *in situ* hybridisation (FISH) used to trace the migration of viral genomes following uncoating of the viral particle showed the presence of a few distinct replication sites on the perinuclear membrane. Within cells that had been infected with poliovirus type 1/Mahoney and type 2 Sabin it was shown that at early and late stages of infection > 85% of replication complexes at these distinct sites contained both genomes (Egger & Bienz, 2002). It is still unknown how and why the genomes migrate to the perinuclear sites of replication. Amongst circulating enteroviruses it has been well documented that only intraspecies recombination is observed (Blomqvist et al., 2003, Brown et al., 2003, Liu et al., 2003, Lukashev et al., 2003). While it is possible that the apparent absence of interspecies recombination is due to differences in receptor usage alone it is tempting to speculate that the reason for this is the absence of a common sequence or structural recognition factor between enterovirus species that functions *in cis* to target the viral genome to specific areas of the perinuclear region. In this model, targeting of the genomic RNA would occur through hijacking of cellular RNA transport pathways. The specific targeting of mRNA to specific regions of cells via molecular “zipcodes” has been described in plants (reviewed in Okita & Choi, 2002) and mammalian cells (reviewed in Oleynikov & Singer, 1998). Amongst viruses the use of RNA transport systems to facilitate nuclear export of spliced and unspliced RNA has been observed in retroviruses (HIV, MPMV, RSV), Hepadnaviruses (HBV) and Herpesviruses (HSV-1).

Within mammalian cells, one RNA transport pathway that has been characterised is the hnRNP-A2 transport pathway. The cellular protein hnRNP-A2 has had numerous functions assigned to it. Within the nucleus it has a role in the regulation of splice site selection (Mayeda et al., 1994) and nuclear export (Daneshmandi, 1999, Lall et al., 1999). In oligodendrocytes hnRNP-A2 has been demonstrated to mediate the anterograde transport of RNAs along microtubules (Ainger et al., 1997). A 21 nt *cis*-acting signal (GCCAAGGAGCCAGAGAGCAUG) originally identified in myelin basic protein (MBP) mRNA has since been identified as the hnRNPA2 response element (A2RE). A2RE-like sequences have been identified in a wide range of cellular RNAs (Ainger et al., 1993, Shan et al., 2003) and all complex retroviruses (Mouland et al., 2001).

In HIV, in addition to the REV/REX mediated export of unspliced RNA, hnRNPA1/A2 mediated export of viral RNA transcripts occurs. Mouland *et al* demonstrated the presence

of A2RE-like sequences in transcripts encoding *gag* and *vpr*. Mouland *et al* have proposed a model whereby the viral RNAs encoding the structural proteins are targeted for transportation to the periphery of the cell by the presence of the A2RE-like sequences. The targeting of the *gag* and *vpr* to the plasma membrane, via the hnRNPA1/A2 transport pathway, (Mouland *et al.*, 2001) restricts expression of the structural proteins to the site of virion assembly (Figure 3.29). This provides the virus with a mechanism by which the risk of the structural proteins being localised in different regions of the infected cells at the point of virion assembly can be minimised. Little homology is observed when the nucleotide sequence of 19 nt motif X (GACGCTAGTCTCCTAGCCC) is compared with the HIV-1 A2RE-like sequence (GACAAGGACCAAAAGAACCCU). However the A2RE/hnRNPA1 transport pathway is likely to be only one pathway out of many involved in intracellular RNA transport. Within plant cells, in addition to a constitutive pathway, two regulated RNA transport systems have been described that target mRNAs to distinct regions of the ER (Hamada *et al.*, 2003). It may be therefore that the poliovirus genomic RNA is transported via a different transport pathway to the distinct site of replication. If RNA transport elements present within genomes of enteroviruses do target the genomic RNA to different locations within the secretory pathway then recombination between different enterovirus species would be unlikely to occur, as recombination requires that the donor genomes are in close proximity. Whilst this proposal is highly speculative it is known from research using inhibitors of the cellular secretory pathway, such as brefeldin A, that differences do exist amongst the *Picornaviridae* in the cellular membranes from which the replication complexes are derived (Gazina *et al.*, 2002). Whether there is an inherent need for these differences in membrane utilisation in terms of productive viral infection is unknown. Recent research using Flock House virus (FHV), an alphadnavirus, has demonstrated that this virus, at least, does not have a specific requirement for its usual cellular target, the mitochondrial membrane, as the virus was still able to replicate after it had been re-targeted to the ER membrane (Miller *et al.*, 2003). This of course does not rule out the presence of species-specific sequence or structural motifs in picornavirus genomes that influences the choice of membrane used. After all recombination forms a natural part of viral evolution. The use of one particular membrane for viral replication by any given virus may allow a mechanism to increase the probability of two viral genomes capable of recombining coming in contact.

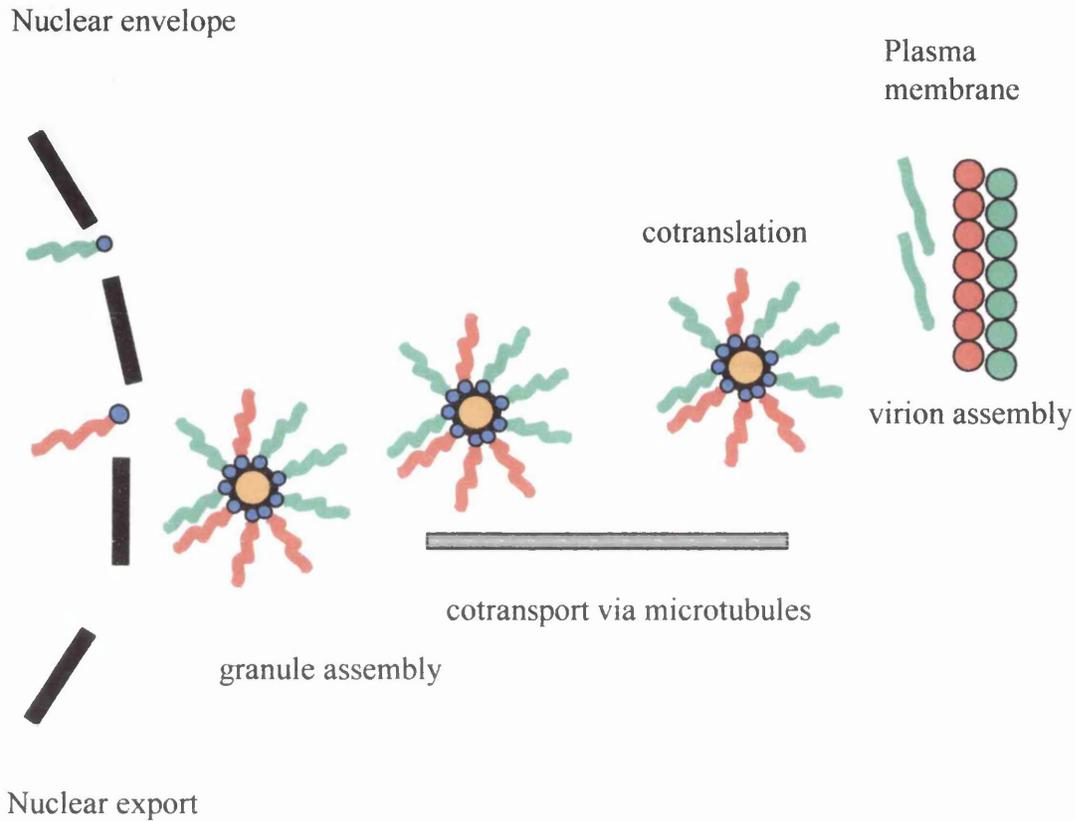


Figure 3.29 Putative model for hnRNP A2/A2RE-mediated in HIV-1

Adapted from Mouland, Xu et al, 2001. *gag* RNA (green squiggles) and *vpr* RNA (red squiggles) both contain A2RE-like sequences that bind to hnRNP A2 (blue spheres) in the nucleus. In the cytoplasm, multiple *gag* and *vpr* RNA molecules, with hnRNP A2 associated, assemble into granules (orange spheres). Granules containing both *gag* and *vpr* RNA are transported to the plus ends of microtubules in the periphery of the cell where the Gag protein (green spheres) and Vpr protein (red spheres) are co-translated. Gag and Vpr proteins co-assemble with HIV genomic RNA (green squiggles) at the plasma membrane.

3.11 Future work

The identification of a species-specific replication determinant raises questions regarding its possible functions, which due to time limitations could not be answered. As a priority, an attempt to select revertant viruses must be undertaken. If compensating mutations, mapping either to the motif X or to the non-structural proteins, can be identified this will provide valuable clues as to the function of motif X.

Conformational analysis of the structure of the RNA, using standard biochemical and enzymatic techniques, may also generate valuable information about the important nucleotides in the motif X sequence by the way it is presented for interactions with proteins or other RNA. Using the information provided by this analysis it may be possible to produce a virus with a “knock-out” phenotype.

The work carried out in this chapter raises the possibility that a second region of SSSV which also appears to be species-specific (Figure 3.26) may also be a determinant of replication or pathogenesis. When describing future work one possibility that must be considered is that the two regions of SSSV, identified as being species-specific, are somehow linked in function. In this study, the introduction of non-coding changes was targeted solely at the region of the SSSV that showed the most suppression. It may be that the regions can compensate for one another to a limited extent and that mutations are required in both regions of SSSV before a severe functional effect can be observed.

4 Preliminary characterisation of 3D^{N18Y}

4.1 Introduction

Work carried out previously by Rohll *et al* on the effect of the 3'UTR on picornavirus replication showed that a natural poliovirus 3'UTR (Figure 4.1A) (a two stem-loop structure) could be functionally replaced by the equivalent structure derived from HRV14 (single stem-loop structure) (Rohll *et al.*, 1995). During the course of their investigation Rohll and colleagues mutagenised a highly conserved stem located at the base of the rhinovirus 3'UTR (Figure 4.1C). One of the mutants identified in the study, designated mut4 (Figure 4.1B), though predicted to be as structurally stable as the wildtype rhinovirus 3'UTR, was shown to be replication deficient when analysed in the context of a subgenomic replicon containing a CAT reporter gene. Further characterisation of the mut4 3'UTR mutation, in the context of virus, identified a single compensating mutation in the N-terminus of the viral polymerase (N18Y) that could rescue virus with a mut4 3'UTR (Meredith *et al.*, 1999). The N18Y mutation was selected when the mut4 3'UTR was studied in the context of a poliovirus or rhinovirus backbone. This provided evidence that the rhino- and enterovirus polymerases interact with the 3'UTR in the same manner. By investigating the effect of N18Y on the biochemical attributes of the polymerase it was hoped that insight would be provided into the nature of the 3'UTR-RdRp interaction. To put these studies into context an overview of the RNA-dependent RNA polymerase (RdRp) encoded by poliovirus (3D^{wt}) will now be provided.

4.1.1 Crystal structure of 3D^{wt}

All RNA viruses require an RdRp to catalyse the replication of their genomes. The size of the RdRp encoded by RNA viruses varies according to the length and type of RNA genome. The first RNA-dependent RNA polymerase to be crystallised was the poliovirus polymerase 3D^{wt} (Hansen *et al.*, 1997).

The crystal structure showed that the 53 kDa polymerase, despite showing a lack of sequence homology with known DNA-dependent DNA polymerases (DdDp), DNA-dependent RNA polymerases (DdRp) and reverse transcriptases (RT), did share a common tertiary structure with these enzymes. The tertiary structure of 3D^{wt}, shown in figure 4.2, resembles an upturned right hand with fingers, palm and thumb motifs (marked with

arrows). In addition to the fingers, palm and thumb sub-domains, the poliovirus polymerase contains a structural element composed of residues N-terminal to the fingers sub-domain. Using a computer-based method of structural prediction Reilly and Kao suggested that the RdRp of viruses from the picorna- bromo-, tombus, and hepatitis C-like virus families these virus families shared this unique N-terminal extension of the fingers sub-domain (Reilly & Kao, 1998).

Six amino acid residues are conserved across all species of eukaryotic RNA viruses (Koonin, 1991). In the poliovirus RdRp these residues are lysine 159, glycine 289, aspartic acid 233 and 238, asparagine 297 and two highly conserved aspartic acid residues (aa 328 and 329) located in motif C (Hansen et al., 1997).

4.1.1.1 The poliovirus RdRp palm domain

The palm domain contains a core structure found in all classes of polymerase (motifs A-D). A fifth motif, known as motif E, has been shown to be unique to RdRp and RT enzymes (Poch et al., 1989) and is located between the palm and the thumb subdomains (Figure 4.3B). Five of the conserved residues found in all RdRp, aspartate 233 and 238 (motif A), asparagine 297 (motif B), aspartate 328 and aspartate 329 (motif C), are all found within the core structure. Through structural analogy with RT and mutational studies the functions of the RdRp conserved residues with respect to the process of RNA synthesis is slowly being elucidated. The structural core of the palm domain in all polymerases provides the platform for catalysis to occur. In the majority of polymerases this core structure consists of two α -helices located beneath four anti-parallel β -sheets (see Figure 4.3A Gorbalenya et al., 2002), a motif similar to the RNA-recognition motif of splicing and ribosomal proteins (Hansen et al., 1997). In the crystal structure obtained by Hansen *et al* aspartate 233, aspartate 328 and aspartate 329 are clustered closely together (figure 4.4). The presence of cations is an essential requirement for the process of catalysing the incorporation of NTP into nascent strands. *In vivo* RdRp's preferentially utilise magnesium cations. In 3D^{wt} aspartates 233 and 328 are believed to be involved in the chelation of magnesium cations *in vivo*. Recent analysis has indicated a role for asparagine 297 in the process of cation selection (Crotty et al., 2003).

In poliovirus 3D^{wt} aspartate 238 hydrogen bonds with asparagine 297, an interaction that is analogous to the Tyr 115–Phe160 interaction in HIV-1 RT. In Moloney murine leukaemia

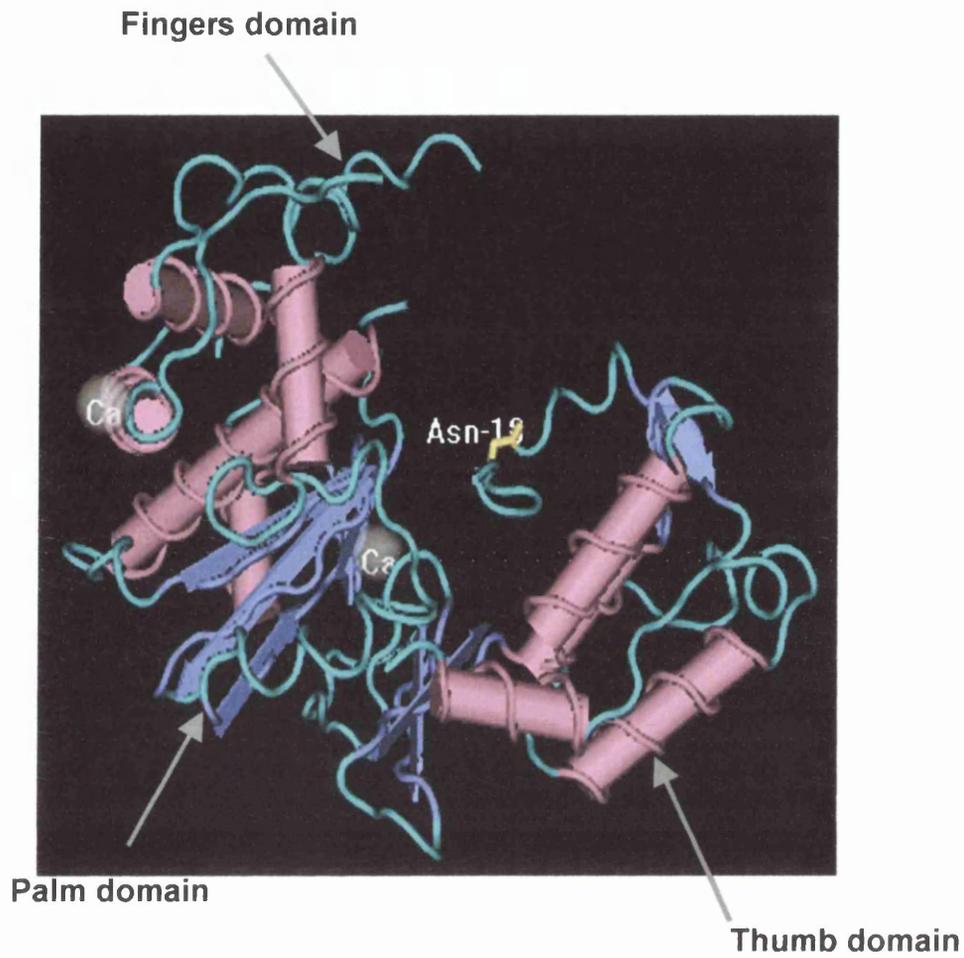


Figure 4.2 Crystal structure of Poliovirus type 1 polymerase

Location of residue 18, highlighted in yellow, within the crystal structure of Poliovirus type 1 polymerase obtained by Hansen et al., 1997. Location of residue 18 in the molecule was highlighted using 3nCD available via the NCBI web server

virus (MMLV) mutation of Tyr 115 to a phenylalanine residue resulted in an alteration in the preferred substrate (dNTP or NTPs) of the RT enzyme (Gao et al., 1997). It has been proposed that an aspartate at residue 238, which can be aligned with residue 115 of the HIV-1 RT, could favour the preferential binding of NTP compared to dNTP by RdRp.

4.1.1.2 The poliovirus finger and thumb domains

The fingers domain is composed of two polypeptide sections one N terminal (aa 97-194) and a second region located between the core motifs A and B (aa 240-285). Two of the conserved residues, lysine 159 and Glycine 289, map within this domain of the polymerase. The function of Lysine 159 has been proposed to be the stabilisation of the triphosphate moiety of the incoming NTP, while glycine 289 forms part of the NTP binding pocket. In the crystal structure of 3D^{wt} significant regions of the fingers sub-domain (residues 98-180 and 267-290) were disordered so no conclusions could be drawn regarding any structural similarities that might exist between 3D^{wt} and RT. However the ordered regions of the finger domain showed significant difference from the analogous region in the HIV-1 RT.

The poliovirus thumb domain is composed mainly of residues C-terminal of the palm domain and is largely alpha-helical in structure. The crystal structure of the 3D^{wt} showed that the thumb domain begins with a short beta-sheet that interacts with motif E. The remainder of the thumb is composed of a series of five alpha-helices. One of the five alpha-helices (termed alphaK) is positioned along the active site cleft. In crystal structures of HIV-1 RT, T4 DNA polymerase and T7 RNA polymerase an alpha-helix is located in an analogous position. The co-crystal structure of HIV-1 RT complexed with DNA has shown that the alpha-helix is positioned in the minor groove of the DNA. Data obtained on the primer/template requirements of the HCV polymerase supports the view that the thumb sub-domain of RdRp functions to correctly position the RNA template in the active site (O'Farrell et al., 2003, Zhong et al., 2000). In the crystal structure of 3D^{wt} the thumb domain appears to interact with the unique N-terminal region of the polypeptide (Figure 4.2). It remains unclear how this might affect the positioning of the template and nascent strand in 3D^{wt}.

Canonical RdRp

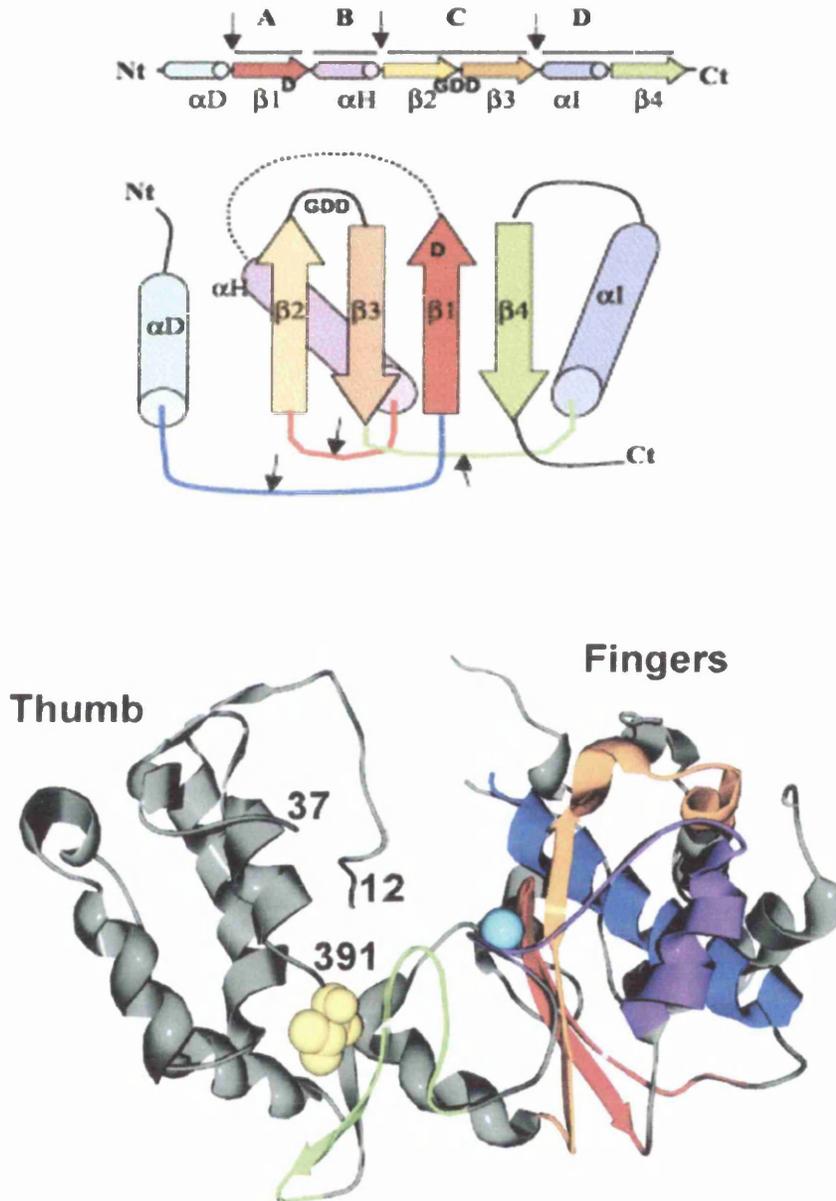


Figure 4.3 The structure of the palm domain of the, $3D^{pol}$. (A) Schematic of the poliovirus palm domain. Taken from Gorbalenya et al., 2002 (B) the location of the palm domain within the the three-dimensional structure of poliovirus polymerase. The five canonical RNA-dependent RNA polymerase domains are colored as follows: motif A, *orange*; motif B, *blue*; motif C, *red*; motif D, *purple*; and motif E, *green*. Val³⁹¹, identified as important for the binding of viral protein 3AB, is shown in *yellow*. The metal ion at the active site is shown in *aqua*. Taken from Lyle et al., 2002 and references within

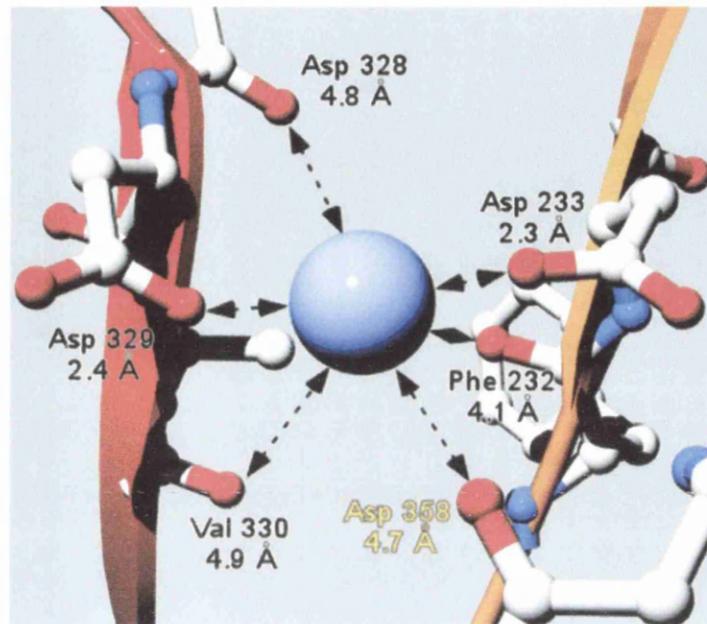


Figure 4.4. Amino acid residues that are located within 5 Å of the metal ion present at the active site of 3D^{wt}.

Figure 4.4 has been taken from Lyle et al., 2002. The Ca²⁺ ion present in the polymerase crystal is shown in *aqua*; this is presumably replaced by Mg²⁺ or Mn²⁺ during catalysis. The ribbon diagrams of the canonical polymerase domains are coloured as described in Fig. 4. x (motif A, *orange*; motif C, *red*). The location of Asp358 that has been demonstrated to be required for 3B uridylylation but not polymerisation is shown in *yellow*.

4.1.1.3 N-terminal region

The unique N-terminal region of the 3D^{wt} polypeptide chain is essential for the catalytic functions of the polymerase. Unfortunately, only two segments of the unique polypeptide region of the poliovirus polymerase were ordered in the 3D^{wt} crystal. The first ordered region (aa 12-37) extends upwards from the active site to the top of the thumb domain. The second ordered segment (aa 67-97) forms an α -helix that is positioned at the base of the fingers domain. An intramolecular connection between these two regions (shown as a blue line in Figure 4.5) would span a distance of > 44 Angstrom (\AA). It is for this reason that Hansen *et al* postulated that the N-terminal residues could be donated by an adjacent polymerase molecule. This type of intermolecular interaction would require the N-terminal amino acid residues to span a distance of < 32 \AA (Red line Figure 4.5).

4.1.1.4 Polymerase-polymerase interactions

Analysis of purified polymerase suggested that 3D^{wt}-3D^{wt} interactions were important for the RNA binding and polymerisation (Pata *et al.*, 1995). Glutaraldehyde cross-linking (Pata *et al.*, 1995) and yeast two-hybrid system (Hope *et al.*, 1997, Xiang *et al.*, 1998) analysis have provided further support to the idea that oligomerisation is important for 3D^{wt} function. The crystal structure of 3D^{wt} identified two interfaces between polymerase molecules: Interface I and II (Figure 4.6).

Interface I involves an interaction of residues located on the back of the palm with residues found on the side of the thumb domain. The Interface I interaction involves at least 23 amino acid side chains and through these interactions a head to tail polymerase fiber is formed in which each polymerase molecule is rotated 180° relative to the adjacent molecule (Hansen *et al.*, 1997). Mutational disruption of Interface I was demonstrated to be lethal to the virus (Diamond & Kirkegaard, 1994). Further analysis demonstrated that the Interface I polymerase-polymerase interactions were important in binding RNA (Hobson *et al.*, 2001, Lyle *et al.*, 2002a). This is in agreement with the model, based on the location of ds-DNA in the co-crystal structure of HIV-1 RT in which the RNA template would be bound along the oligomeric fiber (Figure 4.7).

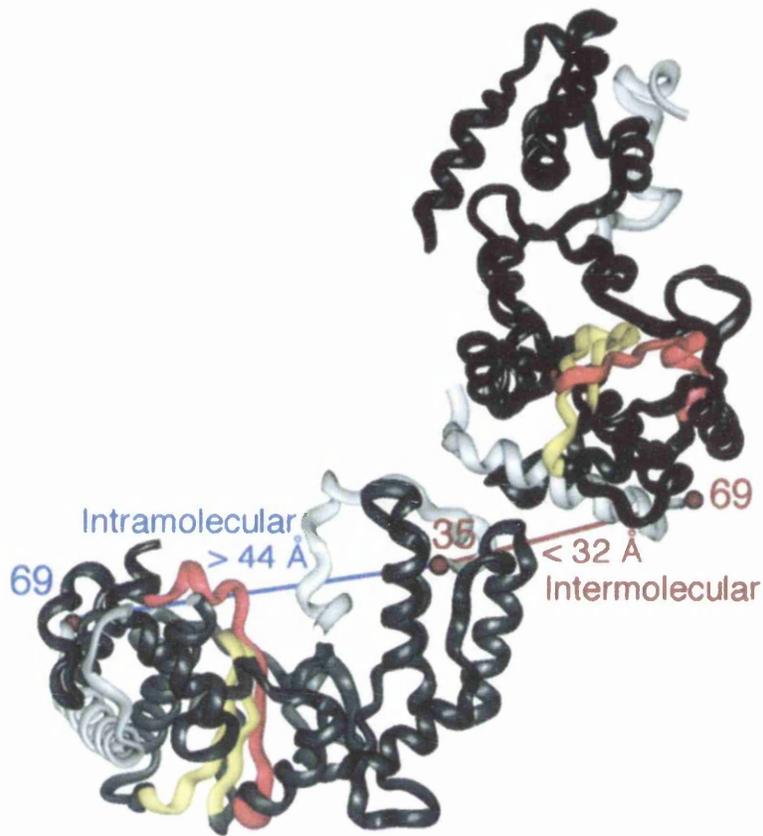


Figure. 4.5 Interactions between the N-terminal strand (light gray) and the thumb subdomain of poliovirus polymerase. Taken from Hobson et al., 2001 . The positions of mutations introduced at Interface II are labeled and shown in yellow, modelled as the mutated amino acid. Intra- (shown in blue) and intermolecular distances (shown in red) between residues 35 and 69 in crystals of poliovirus polymerase.

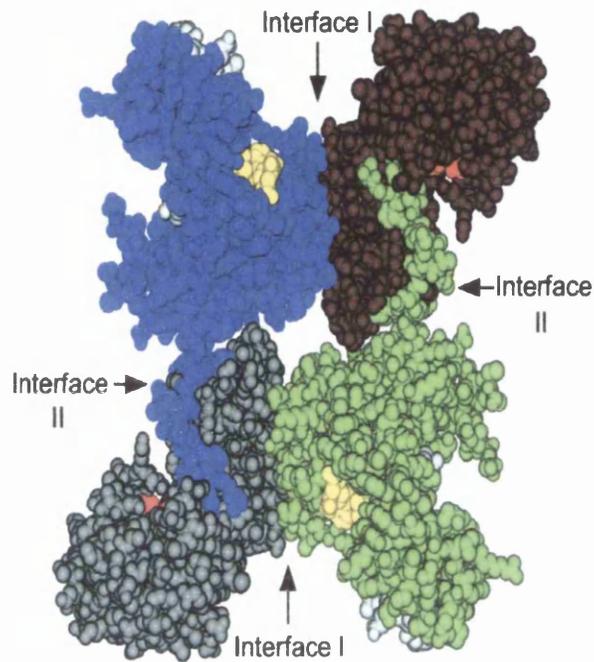


Figure 4.6 *Polymerase-polymerase interactions in the three-dimensional structure of poliovirus polymerase and location of the 3AB binding site.* Taken from Lyle et al., 2002. The unit cell of the crystal structure is shown, with each polymerase coloured distinctly. Residues that constitute the 3AB binding site are shown in *yellow*. The GDD sequence at the active site is shown in *red*. Interface I can be seen as the abutment of polymerase monomers in the horizontal dimension. Interface II is apparent as the abutment of monomers in the vertical direction and the intermolecular donation of the N-terminal strand of one monomer into the thumb of the adjacent monomer. The N-terminal strands that are donated by monomers outside the unit cell are shown in *white*.

In this model the intermolecular donation of the N-terminal domain of one polymerase to a region of the thumb near the active site of an adjacent polymerase molecule provides the second interface (Interface II) (Figure 4.5 and Figure 4.6). Support for this model has been obtained using disulphide cross-linking (Hobson et al., 2001). Interestingly, polymerase molecules unable to form Interface II interactions show no deficit in their ability to bind RNA when compared to wildtype polymerase (Hobson et al., 2001). However the N-terminal residues are critical for the ability of the polymerase to elongate templates (Hobson et al., 2001, Plotch et al., 1989).

Recent EM analysis of purified polymerase showed large sheets and tubes of polymerase molecules were formed. EM analysis of Interface I and Interface II polymerase suggests that both interfaces are required for the formation of the wild-type polymerase structures. Under conditions of low salt Lyle *et al* demonstrated that polymerase could be visualised by immunostaining on tubular structures derived from the membranous replication vesicles. These tubular structures were similar to those obtained using analysis of purified protein (Lyle et al., 2002a). On the basis of this Lyle *et al* have proposed a model in which sheets of polymerase coat the replication vesicles. The model proposed by Lyle *et al* is shown in Figure 4.8. As 3D^{wt} is not membrane associated, the 3D^{wt} are shown tethered to the membrane through the previously described interaction with 3AB (Lyle et al., 2002b). The head to tail arrangement of the polymerase molecules connected by Interface I interactions to form fibers, as described previously by Hansen *et al*, would enable 3AB to tether the 3D^{wt} fiber to the membrane every second molecule. In the model proposed by Lyle *et al* the individual polymerases fibers (shown horizontally across the model) are joined via Interface II interactions. Due to the requirement of the N-terminus for the catalytic activity of the polymerase catalytically active centres would be formed where polymerase fibers were connected by Interface II interactions.

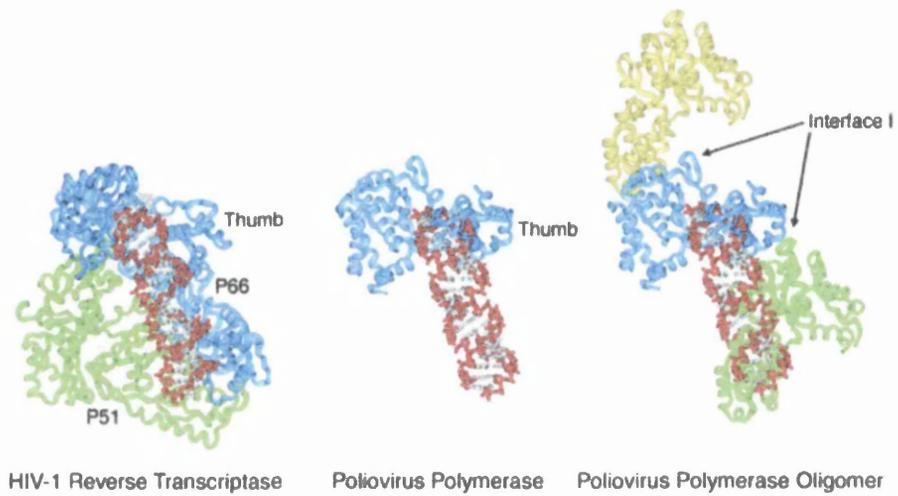


Figure 4.7 Model of a poliovirus polymerase–dsRNA complex.
 Taken from Hobson et al., 2001. The structure is based on the structure of HIV-1 RT complexed to dsDNA (Huang et al., 1998).

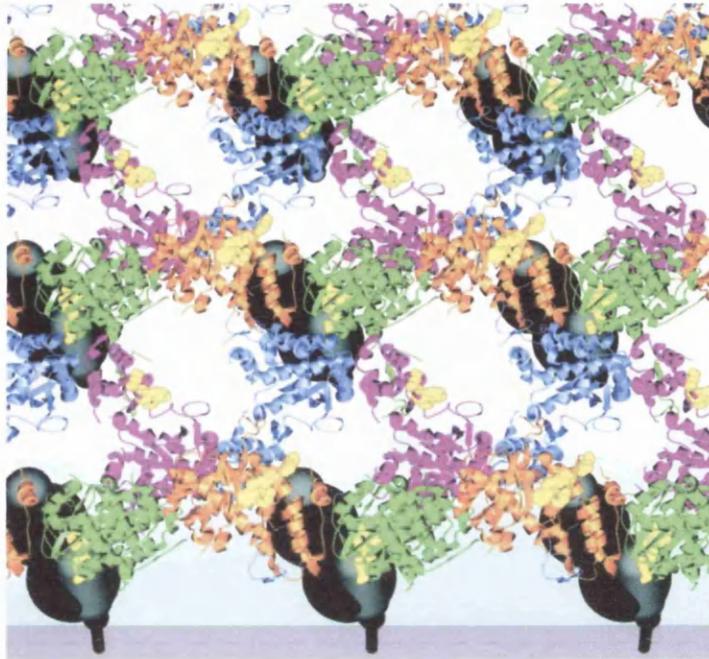


Figure 4.8 Model of higher order polymerase structure.

The model of polymerase array formation as proposed by Lyle et al., 2002 is shown. 3AB is represented as a globular integral membrane protein in black and grey. Fibers of polymerase connected across interface I lie horizontally. These fibers are connected to one another through interactions at interface II, modelled to extend the lattice as a planar array, shown in perspective. Contacts with 3AB through the 3AB binding sites (Hope et al., 1997; Lyle et al., 2002) rendered in space-filled yellow side chains could occur at every other polymerase along the axes defined by both interfaces.

4.2 HCV polymerase: structure-function model for 3D^{wt}

Recently crystal structures have been obtained of the RdRp (NS5B) of HCV (Bressanelli et al., 1999, and O'Farrell et al., 2003). In the crystal structures of the HCV polymerase the N-terminal residues are an extension of the fingers sub-domain and interact with the thumb domain to close the polymerase into a circular structure (Figure 4.9).

Three pieces of evidence have been used to argue that the crystal structure of HCV NS5B may offer a more accurate reflection of the complete structure of 3D^{wt}.

- The ordered region of the fingers sub domain of 3D^{wt} (residues 181-194 and 240-266) are structurally analogous to the HCV counterparts (Bressanelli et al., 1999)
- The β - sheet that connects the palm and thumb domain, that is often referred to as the “ primer grip” is similar in both HCV and poliovirus
- Similar crystal structure has been resolved for the calicivirus polymerase (Ng et al., 2002) and ϕ 6 (reviewed in Van Dijk et al., 2004) suggesting all RdRp may share the same overall structure.

4.2.1.1 Crystal structure of HCV NS5B

Figure 4.9 shows the structure of the HCV strain B4 strain polymerase (Bressanelli et al., 1999). From the crystal structure of HCV (Figure 4.9) it can be seen that the fingers sub-domain folds mainly into α -helices (labelled with numbers) located near the palm domain and β - sheets (labelled with numbers) distal to the palm domain. The distal region of the fingers domain, described by Bressanelli *et al* as the “fingertips” consists of a six-stranded β -sheet and an α -helix (Figure 4.9). Two loops originating from this structure have been described (Δ 1) and (Δ 2). The first of these, Δ 1, extends away from the fingertips and via a short α -helix (helix A formed from residues 24-30), interacts with residues in the thumb domain located between helices O and Q. No corresponding region to Δ 1 has been identified in the HIV-1 RT enzyme. However a similar loop to Δ 2 has been observed in the HIV-RT (Bressanelli et al., 1999). The interaction of the Δ 1 loop with the thumb domain

and the position of the $\Delta 2$ loop creates two tunnels within the structure: the NTP tunnel and the RNA-binding groove (Bressanelli et al., 1999).

4.2.1.2 Internucleotide bond formation: the HCV model

Although the exact positioning of the template and the nascent strand within 3D^{wt} has not been deduced, structural similarities suggest the formation of internucleotide bonds will occur in a similar manner between RdRp. Based on the complex structure of HIV-1 RT complexed to its template/primer and NTP (Huang et al., 1998) a model for HCV has been constructed (Figure 4.10). In this model a maximum of 5 nucleotides of template would be bound by the HCV^{pol} RNA binding-site. This is consistent with the *in vitro* biochemical data that has shown the minimal template length to support the initiation of RNA synthesis to be 5 nts (Zhong et al., 2000).

A β -hairpin structure that is part of the thumb domain is positioned facing towards the active site. No β -hairpin structure is found at the analogous position in the thumb domain of 3D^{wt}. In the computer model produced by Zhong *et al* the 3' of the template abuts the β -hairpin structure. This interaction could ensure that the template is in the correct position to enable the formation of the optimal structure between T+3 and the incoming NTP for catalysis of the nucleotidyl transfer reaction to occur. Alternatively, Zhong *et al* postulate that the β -hairpin structure may function to separate the nascent RNA strand from the template, as has been shown for the N-terminal domain of T7 RNA polymerase, which occupies the same spatial arrangement (Cheetham & Steitz, 1999). This suggestion was based on the fact that the formation of a second nucleotidyl transfer reaction (between T+4 and an incoming NTP) requires the translocation of the template towards the β -hairpin.

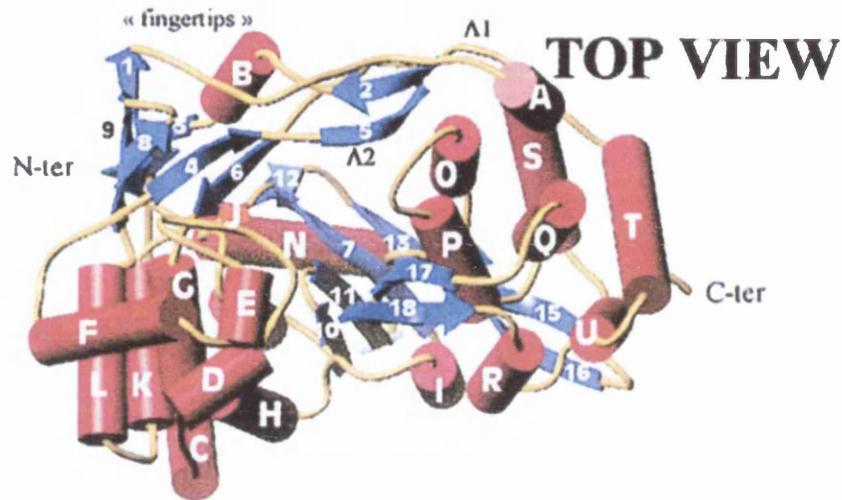
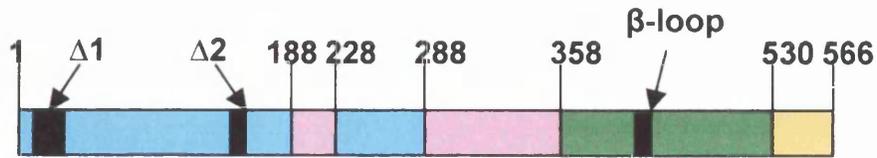


Figure 4.9 *The crystal structure of the HCV RdRp.*

In the schematic at the top, residues of the fingers sub-domain are highlighted in **blue**, the palm domain in **pink** and the thumb domain in **green**. The C-terminal residues are shown in yellow. The crystal structure of the RdRp of HCV. Taken from Bressanelli et al., 2003. In the crystal structure of HCV NS5B. α - helices are labelled with capital letters and β -sheets are labelled with numbers. Connecting loops are shown as yellow tubes. (B) The “fingertips” are composed off a mixed barrel of an α - helices and β -sheets. Two loops ($\Delta 1$ and $\Delta 2$) emanate from this barrel to close the enzyme by interacting with the back of the thumb.

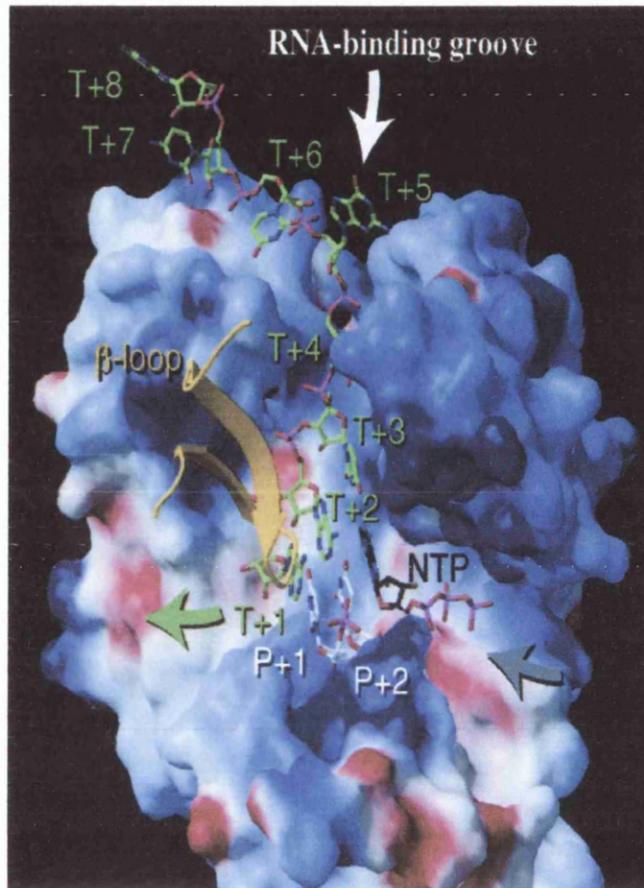


Figure 4.10 Model for internucleotide synthesis in HCV NS5B. Taken from Zhong et al., 2000. The HCV NS5B protein is depicted by a molecular surface colored by local electrostatic potential from red to blue as the potential ranges from negative to positive. The thumb subdomain is omitted from this structure. Atoms of RNA are shown and denoted T+1 through T+8 on the template strand and P+1 to P+2 on the primer strand. P+1 and P+2 base pair with T+1 and T+2, respectively. The incoming NTP base pairs with T+3. The β -hairpin is shown in yellow. The 3' of the template is abut the β -hairpin structure. An interaction between this structure and the template could ensure that the template is in the correct position to enable the formation of the optimal structure between T+3 and the incoming NTP for catalysis of the nucleotidyl transfer reaction to occur. Alternatively, the β -hairpin structure may function to separate the nascent RNA strand from the template, as has been shown for the N-terminal domain of T7 RNA polymerase, which occupies the same spatial arrangement. The direction of the motion of the nucleic acid and the NTP during nucleotidyl synthesis are indicated by arrows.

Recently a crystal structure of NS5B complexed with single-stranded RNA has been obtained that is in agreement with the computer predicted model (O'Farrell et al., 2003). The co-complexed structure obtained showed that the last base of the U₅-oligo interacted with the β -strand and residues of the C-terminus of the protein. In addition direct interactions between the U₅-oligo and residues 14, 93, 95, 97, 98, 139, 141 and 160 of the fingers domain were observed. It is proposed that these direct interactions play an important role in directing the RNA to the correct position in the active site thus enabling the phosphoryl transfer reaction to occur.

4.2.1.3 Oligomerisation of HCV NS5B

Analysis of HCV NS5B polymerase molecules have shown that the efficiency of RNA synthesis increased under conditions that promoted oligomerisation (Wang et al., 2002). Two interfaces have been described between NS5B molecules (Figure 4.11), both of which involve the “fingertips and $\Delta 1$ and $\Delta 2$ region of the HCV polymerase. Interface A is a head to tail interaction between the thumb domain (α - helices T and U) of one structure and the “fingertips” of the adjacent molecule (β -sheets 1, 4 and 8). The second interface is a head to side interaction between the back of the thumb domain in one molecule (loops between α - helices T and U) and an anti-parallel β -strand of the fingers sub-domain of an adjacent molecule (α - helix B and β -sheets 2 and 5 found in $\Delta 1$ and $\Delta 2$ loops) (Wang et al., 2002).

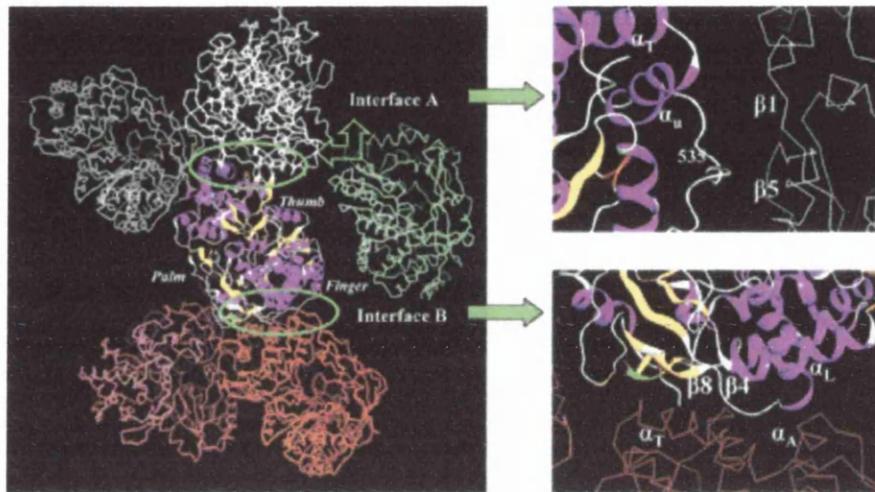


Figure 4.11. Interactions between molecules of the HCV RdRp in a crystal lattice. Taken from Wang et al., 2002. The crystal structure of HCV NS5B- Δ C21 was solved at a resolution of 2.9Å. The two interfaces identified from the N35B crystal lattice are highlighted and labelled Interface A and Interface B. Nomenclature for the key motifs is the same as in figure 4.9.

4.3 N18Y mutation

In the entero-, hepatoviruses and HRV14 residue 18 is conserved as an asparagine (N18). In the arrangement of the N-terminus in the crystal structure of 3D^{wt} the N18 residue is located towards the base of the active-site cleft (Figure 4.2). If the position of residue N18 is modelled onto a circular RdRp structure, based on the HCV NS5B crystal structure then it is positioned in an analogous position to the Δ 1 loop (Figure 4.12).

In the original characterisation of the N18Y mutation it was demonstrated that in the absence of the mut4 3'UTR the resulting virus could replicate 0.5 log₁₀ higher than the parental PV3-HRV14 chimeric virus (Figure 4.13). It was initially suggested that the increase in fitness of the PV3-HRV14 chimeric virus containing the N18Y was due to a non-specific enhancement of the polymerisation activity of the polymerase. However Meredith *et al* demonstrated that the N18Y was a specific coding change as the N18Y coding change was not one of the compensating mutations recovered from transfections of viral RNAs that had the entire 3'UTR deleted (Meredith *et al.*, 1999).

So what aspect of polymerase function could the N18Y mutation effect? The overviews of the current models about 3D^{wt} and HCV offer some insight into the potential functions affected by this change:

Oligomerisation - In the crystal structures of 3D^{wt} and HCV the N-terminal residues have been implicated in the oligomerisation of the polymerases. In both HCV and 3D^{wt} RNA synthesis and RNA-binding have been shown to be co-operative. An alteration in the ability of the protein to form oligomers could have effects on the formation of the polymerase lattice and indirect effects on RNA binding and polymerisation.

RNA-binding - The data published by Meredith *et al* suggests a direct interaction between N18Y and the virus 3'UTR. Interestingly, within the HCV-liganded complex (O'Farrell *et al.*, 2003) it was shown that residue 14 of the HCV polymerase directly interacts with the RNA template. Structural alignments of HCV and 3D^{wt} seem to indicate that N18 is structurally analogous to residue 14 of the HCV (Bressanelli *et al.*, 1999). If N18 is functionally analogous then it may be that the only difference between 3D^{wt} and 3D^{N18Y} might be an alteration in the affinity of binding of the protein to the 3'UTR.

To define how the N18Y coding change exerts its effect the polymerase must be characterised biochemically. This requires the expression and purification of large quantities of protein. The decision was taken to express 3D^{wt} and 3D^{N18Y}, using a bacterial expression system, in an untagged form to ensure the data obtained from the *in vitro* biochemical studies was as relevant as possible to what occurs *in vivo*. Preliminary characterisation using *in vitro* assays developed previously to study attributes of 3D^{wt} including polymerisation and uridylylation would ensure that the purified recombinant proteins were functional. It was also hoped that the preliminary characterisation would elucidate the biochemical basis for the phenotype observed by Meredith *et al*.



Figure 4.12 The location of N18Y within the polymerase structure is shown based on the crystal structure of HCV NS5B

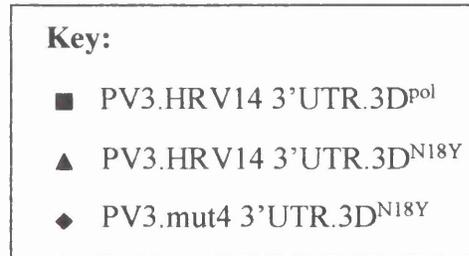
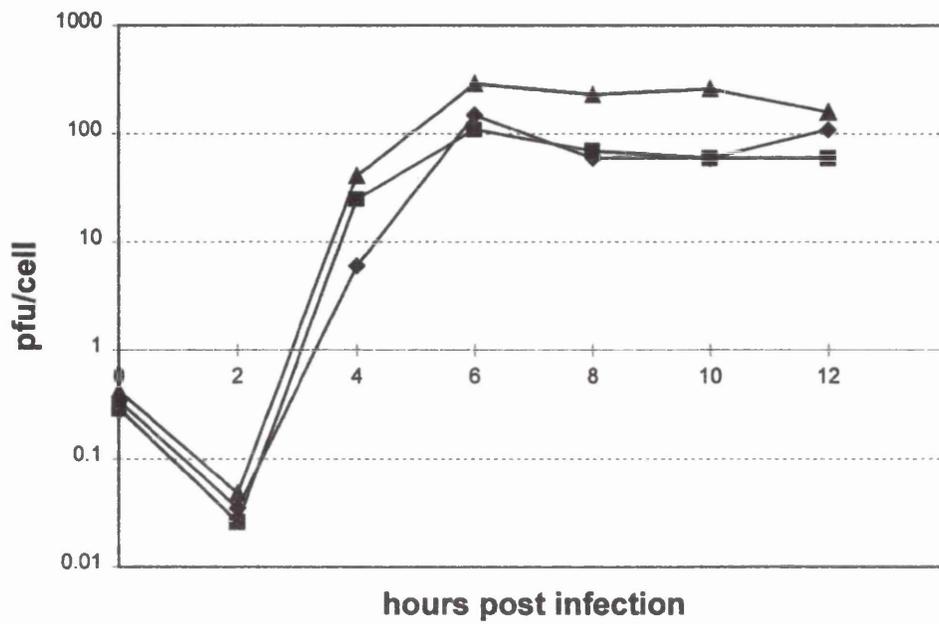


Figure 4.13 A one step growth curve to compare *N18Y* viruses. Taken from Meredith et al., 1999. The key indicates the chimeric viruses used in the growth curves.

4.4 Construction of DNA vectors expressing 3D^{wt} and 3D^{N18Y}

It has previously been documented that enzymatically active 3D^{wt} can be successfully expressed in and purified from *E. coli* (Morrow et al., 1987, Plotch et al., 1989). The plasmid that was used to express 3D^{wt} and 3D^{N18Y} in bacterial cells was pET16b (Novagen). Expression of proteins from the expression vector pET16b occurs from a T7 promoter (Figure 4.14). The expression vector pET16b encodes an N-terminal His-Tag sequence to enable purification. However it was decided to purify an untagged version of the polymerase. To achieve this, the protein was fused in-frame at the *Nco I* restriction enzyme site (blue letters in Figure 4.14) and *Bam HI* site (red letters in Figure 4.14).

Oligonucleotide primers IG24 and IG25, which contain the relevant restriction sites (Table 6.2 materials and methods), were used to amplify the cDNA sequence of 3D^{wt} using PCR from the subgenomic replicon pT7FLC/Rep3 (Barclay et al, 1998). Oligonucleotide primers IG24 and IG25 were also used to amplify the cDNA sequence of a 3D^{wt} containing the N18Y mutation (3D^{N18Y}). The cDNA sequence of the polymerase 3D^{N18Y} was amplified from pT7FLCN18Y a plasmid that contained a complete cDNA copy of the genome of PV3.mut4 3'UTR.3D^{N18Y}. The plasmid pT7FLCN18Y was constructed by Janet Meredith. The expression plasmid pET16b3D^{N18Y} was constructed by Dr.V.Cowton.

The PCR product and the vector were digested with *Nco I* and *Bam HI*. The digested vector and PCR product were purified from agarose prior to being ligated together. The ligations were purified and transformed into *E. coli* DH5 α using standard electroporation conditions. The transformants were screened for the presence of insert by redigesting the recovered plasmids with *Nco I* and *Bam HI*. Transformants positive for insert were identified by the presence of an additional band, when compared to linearised pET16b of approximately 1.5 kb in size. Sequencing was used to confirm that no additional coding changes had been introduced to the viral polymerase.

T7 promoter →

AGATCTCGATCCCCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGA
TAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATAACCAT
 GGGCCATCATCATCATCATCATCATCATCATCACAGCAGCGGCCATATCGAAGGTC
 GTCATATGCTCGAGGATCCGGCCTGCTAACAAAGCCCCGAAAGGAAGCTGGGAGTT
 GGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACG
GGTCTTGAGGGGTTTTTG ←

T7 terminator

Figure 4.14 pET16b cloning/expression region

The *lac* operon (bold black line) and His tag (bold grey line) are highlighted. The restriction enzyme sites used to clone the 3D^{N18Y} and 3D^{wt} cDNA sequence into pET16b were *Nco I* (Blue letters) and *Bam HI* (red letters). The T7 promoter and terminator sequences are shown with a black arrow. Information was obtained from Novagen.

4.5 Expression of 3D^{wt} and 3D^{N18Y}

Expression of the viral polymerase from the expression vector is driven by a T7 promoter. The T7 promoter in the expression vector is under the control of the *lac* operon (Figure 4.14). As a consequence expression of the protein of interest is suppressed in the presence of glucose. In the absence of glucose the expression of the protein of interest can be induced by isopropylthio- β -galactosidase (IPTG), a lactose analog.

To express 3D^{wt}, the expression construct pET16b3D^{wt} was transformed using calcium chloride into *E.coli* BL21 DE3 and *E.coli* BL21 DE3 (PlysS). Both these strains of bacteria contain a chromosomal copy of the T7 RNA polymerase under the control of an IPTG-inducible promoter. In addition, to the T7 RNA polymerase *E.coli* BL21 DE3 (PlysS) also expresses T7 lysozyme. Initially the transformed *E.coli* cells were grown as a starter culture overnight at 37 °C for 16 hours, with shaking. After carrying out small-scale expression analysis it was discovered that expression of the protein from the cDNA was greater if the starter culture had been grown in the presence of 2 % glucose. The small scale expression analysis also showed that the recovery of protein was greatest from *E.coli* BL21 DE3 (PlysS). All large scale expression studies were subsequently carried out using freshly transformed *E.coli* BL21 DE3 (PlysS).

The starter culture, which had been grown for 16 hours in the presence of 2 % glucose with shaking, was used to inoculate 3 litres of LB media to an OD₆₀₀ of 0.3. After growing the culture at 37 °C with shaking till the OD₆₀₀ was between 0.6 and 1.0 the expression of the protein was induced with the addition of IPTG, the final concentration of which was 1 mM. The IPTG was maintained in the culture for the duration of the 5 hour induction. 3D^{N18Y} was expressed in the same way as 3D^{wt}.

4.6 Purification of the 3D^{wt} and 3D^{N18Y}

Following a 5 hour induction the bacterial cells were harvested by centrifugation at 3000 rpm in a Sorvall GS-3 rotor. The supernatant was removed and the pelleted bacterial cells were frozen at -20 °C. The frozen cells were thawed and resuspended in 20 mls of resuspension buffer I per litre of original culture (for details see materials and methods 6.3.1). The expression of T7 lysosyme by *E.coli* BL21 DE3 (PlysS) aids the rupturing of the bacterial cells during thawing.

The lysed bacterial cells were then subjected to 5 cycles of sonication using a sonicator (Dawe instruments type 7532-1A). Each burst of sonication was carried out for 1 minute. Following sonication, the bacterial preparation was centrifuged for 30 minutes at room temperature at 20,000 rpm in a Sorvall SS34 rotor. Analysis of the soluble and insoluble fractions showed that the majority of the protein was located in the soluble fraction.

Purification of the polymerase was carried out using an sulphopropyl (SP)-sepharose column with a bed volume of 1.5ml. Prior to purification the column was equilibrated using 8 column volumes of resuspension buffer I. The soluble cell lysate fraction was then adsorbed onto the column. Prior to elution the sepharose was washed with four column volumes of wash buffer I. The proteins bound to the column were eluted sequentially by four column volumes each of elution buffers I-IV. Samples positive for 3D^{wt}, as determined by SDS-PAGE gel electrophoresis, were dialysed overnight at 4 °C in polymerase storage buffer. Purified polymerase was stored at -70 °C in polymerase storage buffer.

Figure 4.15 presents the data from the SP-sepharose purification of 3D^{wt} (Figure 4.15A) and 3D^{N18Y} (Figure 4.15B).

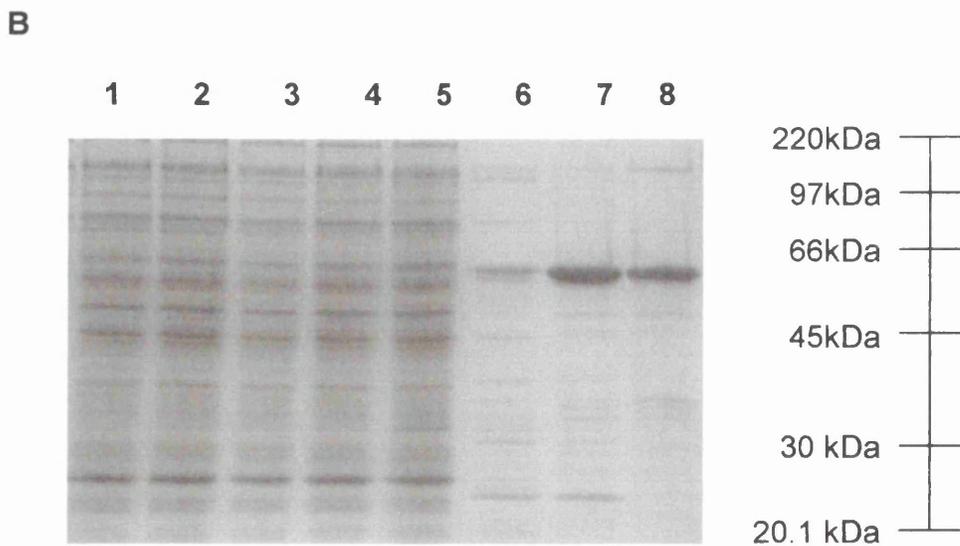
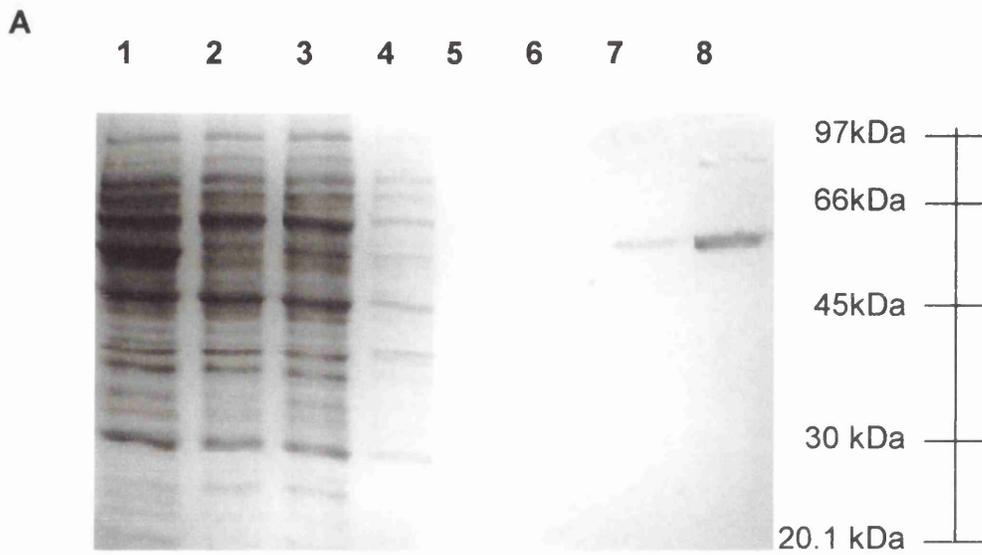


Figure 4.15 SDS-electrophoresis gel analysis of fractions from the purification of 3D^{wt} (A) and 3D^{N18Y} (B) using an SP-sepharose column. Soluble cell lysate obtained from cultures of BL21 DE3 (PlysS) expressing either 3D^{wt} or 3D^{N18Y} were loaded onto an SP-sepharose column that had been equilibrated with resuspension buffer I. For analysis purposes, fractions relating to the column flowthrough (lanes 1 – 4), column washes (lanes 5 and 6) and 100 (lane 7) and 200 mM (lane 8) elution steps of the purification procedure were run on an SDS-Page electrophoresis gel.

4.7 Preliminary functional characteristics of 3D^{wt} and 3D^{N18Y}

4.7.1 Uridylylation

Two mechanisms have been identified through which RdRp initiate RNA synthesis: primer independent (*de novo* initiation) or primer dependent. The RdRp of many viruses, including HCV and bacteriophage Q β , can initiate RNA synthesis *de novo* (Kao et al., 2001). Picornaviruses and viruses in the picorna-like superfamily, on the other hand, absolutely require a primer *in vivo* to initiate RNA synthesis. A virus encoded protein, designated VPg, which is found covalently attached to the 5' terminal-pUpUp, of the genome and anti-genome (Rothberg et al., 1978a) is utilised as a primer by picornaviruses following uridylylation. The covalent linkage of uridylylate residues to VPg is catalysed by 3D^{wt}, in a reaction templated for by a single adenylate residue located in the loop of the CRE (VPg-pUpU synthesised in this manner is utilised in the initiation of positive-strand), using a slideback mechanism (Paul et al., 2003), or by poly A (VPg-poly U synthesised in this manner is believed to be used to initiate negative-sense strands). The synthesis of VPg-pUpU by 3D^{wt} can be replicated *in vitro* using full-length genomic transcripts or transcripts of CRE. *In vitro* the uridylylation reaction is stimulated more than 20-fold when 3CD^{pro} is included in the reaction (Paul et al., 2000). Functional analysis carried out by Pathak *et al* on 3D^{wt} Interface I mutants identified that an interaction between the back of the thumb domain of the polymerase and 3C^{pro} was required for this stimulation to occur (Pathak et al., 2002).

To ensure that the purified polymerases were functional with respect to their ability to uridylylate VPg an *in vitro* uridylylation assay was carried out using the method described by Paul *et al* (Paul et al., 2000). The incorporation of [α -³²P] UTP in to VPg-pUpU by 3D^{wt} enables the products of the reaction to be visualised using a phosphoimager documentation system following electrophoresis on a 12 % Tris-tricine SDS-polyacrylamide gel. The ability of the 3D^{wt} and 3D^{N18Y} to uridylylate VPg was assessed over concentrations ranging from a maximum concentration of 2 μ g and 0.5 μ g. To obtain these concentrations 3D^{wt} and 3D^{N18Y} were serially diluted in 3D^{pol} storage buffer.

Figure 4.16 shows the developed image of products of the *in vitro* uridylylation assay as synthesised by either 3D^{wt} or 3D^{N18Y}. Figure 4.16 confirms that both the purified 3D^{wt} and 3D^{N18Y} are both functionally capable of catalysing the addition of uridylate residues to VPg. No significant difference in the relative abilities of 3D^{wt} and 3D^{N18Y} to catalyse the uridylylation of VPg to form VPg-pU can be observed at 0.5 µg and 2 µg. Figure 4.16 does suggest that a difference may exist in the ability of 3D^{N18Y} to catalyse the formation of VPg-pUpU compared to 3D^{wt}. An explanation for this phenotype is provided in section 4.8.1.

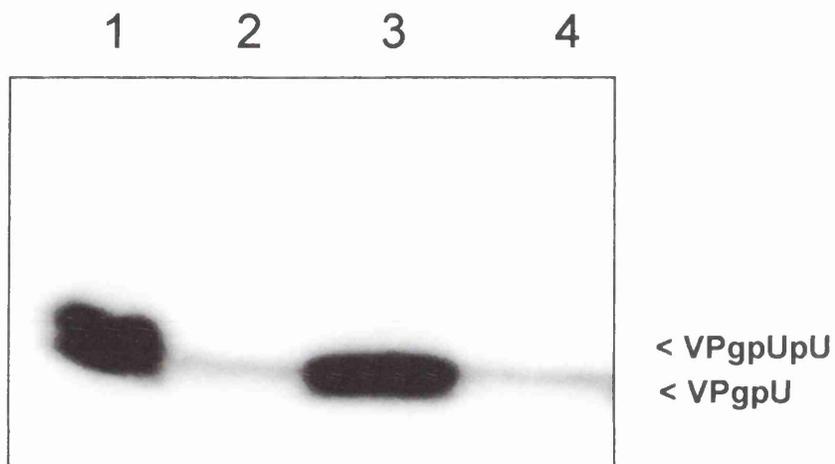
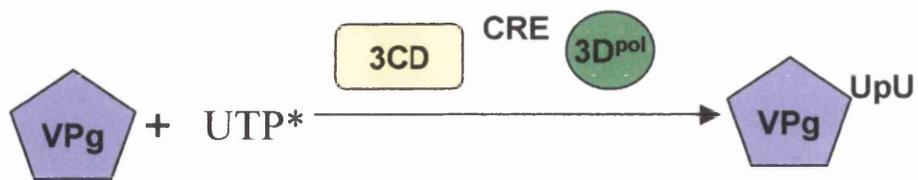


Figure 4.16 Visualised products of an in vitro uridylylation assay synthesised by $3D^{wt}$ or $3D^{N18Y}$

The ability of the $3D^{wt}$ (lanes 1 and 2) and $3D^{N18Y}$ (lanes 3 and 4) to catalyse the uridylylation of VPg was tested with 2 (lanes 1 and 3) and $0.5\mu\text{g}$ (lanes 2 and 4) of the purified protein.

4.7.2 Polymerisation

In vivo 3D^{wt} is highly specific with respect to the template that it copies. In contrast to the situation *in vivo* it has been demonstrated that *in vitro* 3D^{wt} can catalyse the elongation of artificial templates (Arnold et al., 1999, Morrow et al., 1987, Plotch et al., 1989). To investigate whether 3D^{wt} and 3D^{N18Y} retained functional polymerisation activity the ability of both polymerases to elongate a poly-(U) primed poly-(A) template was assessed. Elongation of the template was monitored by calculating the picomoles of [α -³²P] UTP that were incorporated over 30 minutes at 34 °C. The ability of the 3D^{wt} and 3D^{N18Y} to elongate a poly-(U) primed poly-(A) template was assessed over concentrations ranging from a maximum concentration of 20 μ M to a minimum concentration of 0.16 μ M. To obtain these concentrations the polymerase stocks of 3D^{wt} and 3D^{N18Y} were serially diluted in 3D^{pol} storage buffer. To ensure that the reactions were started at the same time the reaction mix (see materials and methods 6.10.9) was added to the pre-aliquoted polymerase on ice. The reactions were then transferred to the 34 °C incubator. Reactions were stopped with the addition of 0.5 mM EDTA. 5 μ l of each reaction was spotted onto a DE81 filter (supplied by Wakenan) that was washed several times in 0.5 M sodium bi-carbonate and airdried prior to be counted in a scintillation counter (Beckman model number LS5000 CE). Each polymerase concentration was analysed in triplicate independently and the average picomoles of [α -³²P] UTP for each concentration was calculated. Error bars were plotted as a measure of variation above and below the mean value. The polymerisation activities of 3D^{wt} and 3D^{N18Y} were tested at pH 7.5 and pH 5.5, a pH at which oligomerisation is promoted (Pata et al., 1995).

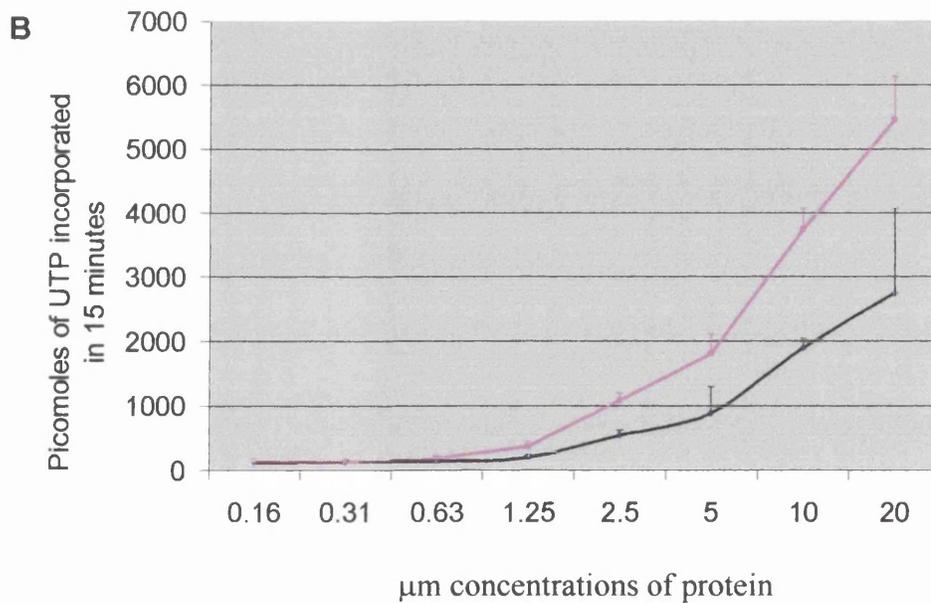
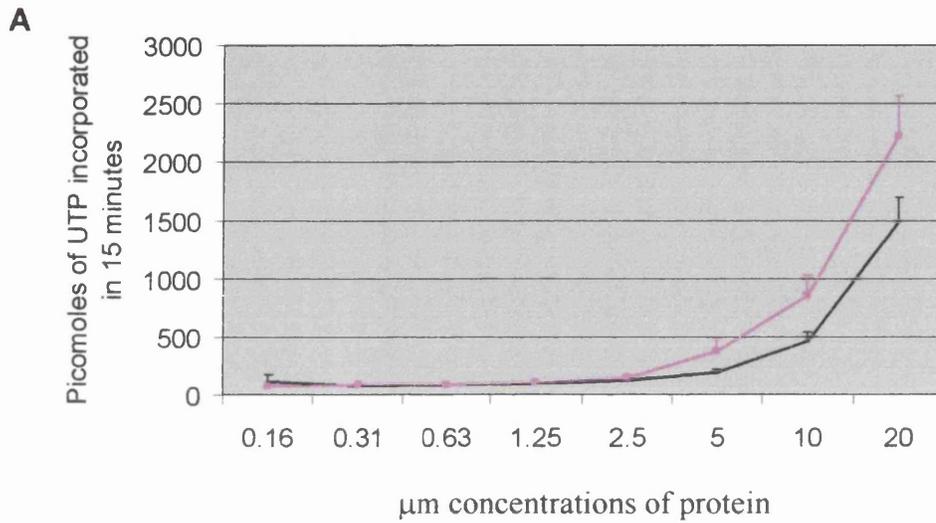
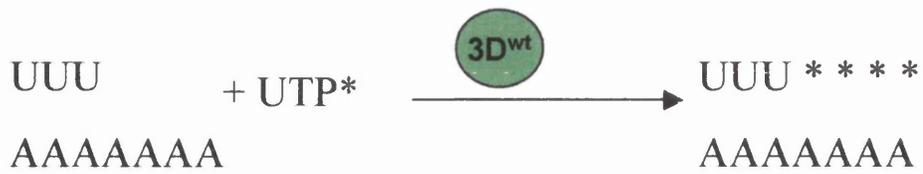


Figure 4.17 Poly r(U) polymerase activity of the 3D^{wt} and 3D^{N18Y} PV3 polymerase measured at at pH7.5(A) and a pH5.5 (B). In both A and B 3D^{wt} is represented as a black line and 3D^{N18Y} is represented with a purple line.

Figure 4.17 confirms that both the purified 3D^{wt} and 3D^{N18Y} have retained functional polymerisation activity. Using the poly-(U) primed poly-(A) template a difference in the rate of polymerisation activity between the 3D^{wt} and 3D^{N18Y} at both pH 7.5 (Figure 4.17A) and pH 5.5 (Figure 4.17 B) was observed. Figure 4.17 shows that the average picomoles of [α -³²P] UTP incorporated into the template by 3D^{wt} and 3D^{N18Y} at pH 5.5 is greater than that observed at pH 7.5. This is consistent with what has previously been reported (Pata et al., 1995). At both pH 5.5 and pH 7.5 the average picomoles of [α -³²P] UTP incorporated into the template by 3D^{N18Y} over the range of concentrations from 2.5 to 20 μ M is greater than that observed in reactions catalysed by 3D^{wt}. The observation that the average incorporation of [α -³²P] UTP into the template by 3D^{wt} and 3D^{N18Y} differ suggests that the kinetics of polymerisation of 3D^{wt} and 3D^{N18Y} are different. Using a different template these results could not be reproduced. Reasons for this are discussed fully in sections 4.9.1 and 4.9.3.

4.7.3 Terminal transferase activity

In vitro, 3D^{wt} has been reported as having terminal uridyl transferase activity (Neufeld et al., 1994) though whether the terminal uridyl transferase activity has any *in vivo* relevance remains unclear. To investigate whether 3D^{N18Y} retained terminal uridyl transferase activity the ability of 3D^{wt} and 3D^{N18Y} to catalyse the addition of [α -³²P] UTP to PV3 RNA at pH 7.5 was tested. The terminal transferase activity of the polymerases was tested at concentrations of 2, 1 and 0.5 μ g of the purified proteins. Figure 4.18 shows the developed image of products of the *in vitro* terminal uridyl transferase reaction as synthesised by either 3D^{wt} or 3D^{N18Y}. Figure 4.18 confirms that both the purified polymerases retain terminal uridyl transferase activity. No significant difference in the ability of 3D^{wt} and 3D^{N18Y} to catalyse the addition of a uridylate residue to the PV3 RNA was observed between the proteins at higher concentrations of protein (2 μ g). A subtle difference in terminal uridyl transferase activity was detected at lower concentrations (0.5 μ g). Determining the significance of this difference would require further investigation.

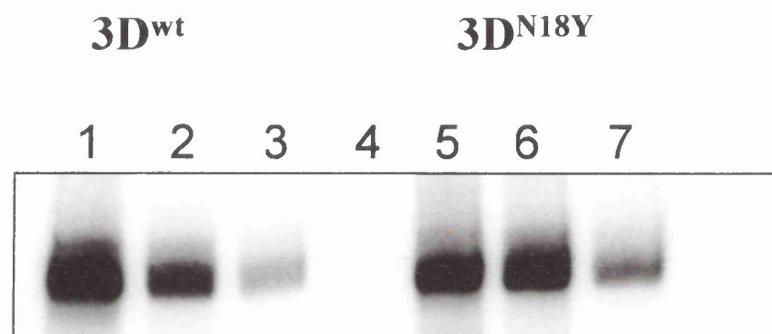
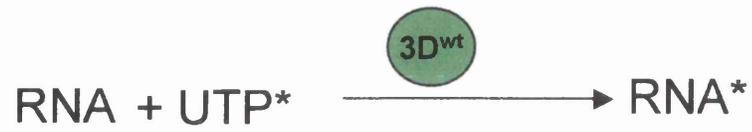


Figure 4.18 Terminal transferase activity of 3D^{wt} and 3D^{N18Y}PV3 polymerase measured at pH 7.5. The terminal transferase activity of 3D^{wt} and 3D^{N18Y}PV3 was tested at concentrations of 2 (lanes 1 and 5), 1 (lanes 2 and 6) and 0.5 (lanes 3 and 7) µg of purified proteins.

4.7.4 Oligomerisation

The crystal structure of 3D^{wt} implied that the N-terminal residues of the protein might be involved in the oligomerisation of the polymerase (Interface II interaction). Analysis of the N-terminal region has confirmed that mutations in the N-terminal 3D^{wt} can abolish polymerase function (Hansen et al., 1997). Biochemical analysis has shown that 3D^{wt}-3D^{wt} interactions are important for the polymerase functions of RNA binding and polymerisation (Pata et al., 1995). Research has since shown that the elimination of function does not result from an inability of N-terminal mutants to bind the RNA (Hobson et al., 2001). These observations have resulted in the proposal of a model in which catalytically active polymerase centres are formed, within a lattice structure, only where Interface II interactions occur. Any alteration in the N-terminal residues might result in an alteration in the ability of the RdRp to oligomerise. One consequence of this might be alterations in RNA binding and polymerisation. To identify whether any difference existed in the interactions of 3D^{wt} molecules compared with those observed between molecules of 3D^{N18Y}, it was decided to analyse the interactions of the polymerase using the yeast-two hybrid system. This system was chosen as it also provided a method by which the effect of the N18Y mutation could be studied on the interactions with the replication proteins 3AB and 3CD^{pro}. To investigate these interactions fully it was decided, in addition to 3D^{N18Y}, the N18Y should be expressed, in the yeast-two hybrid system as 3CD^{N18Y}.

4.7.4.1 Yeast-two hybrid system

The yeast-two hybrid system (Figure 4.19) was developed as a genetic method for studying protein interactions *in vivo* (Fields & Song, 1989). This system utilises the observation that the activation domain (AD) of cellular transcriptional factors does not have to be part of the same polypeptide as the DNA binding domain (BD) in order to reconstitute a functional transcriptional activator (Ma & Ptashne, 1988, Triezenberg et al., 1988). Figure 4.19 shows a standard yeast-two hybrid system. Basically, the proteins of interest are fused in frame with either the BD (BD/proteinX) or AD (AD/protein Y) of the transcription factor.

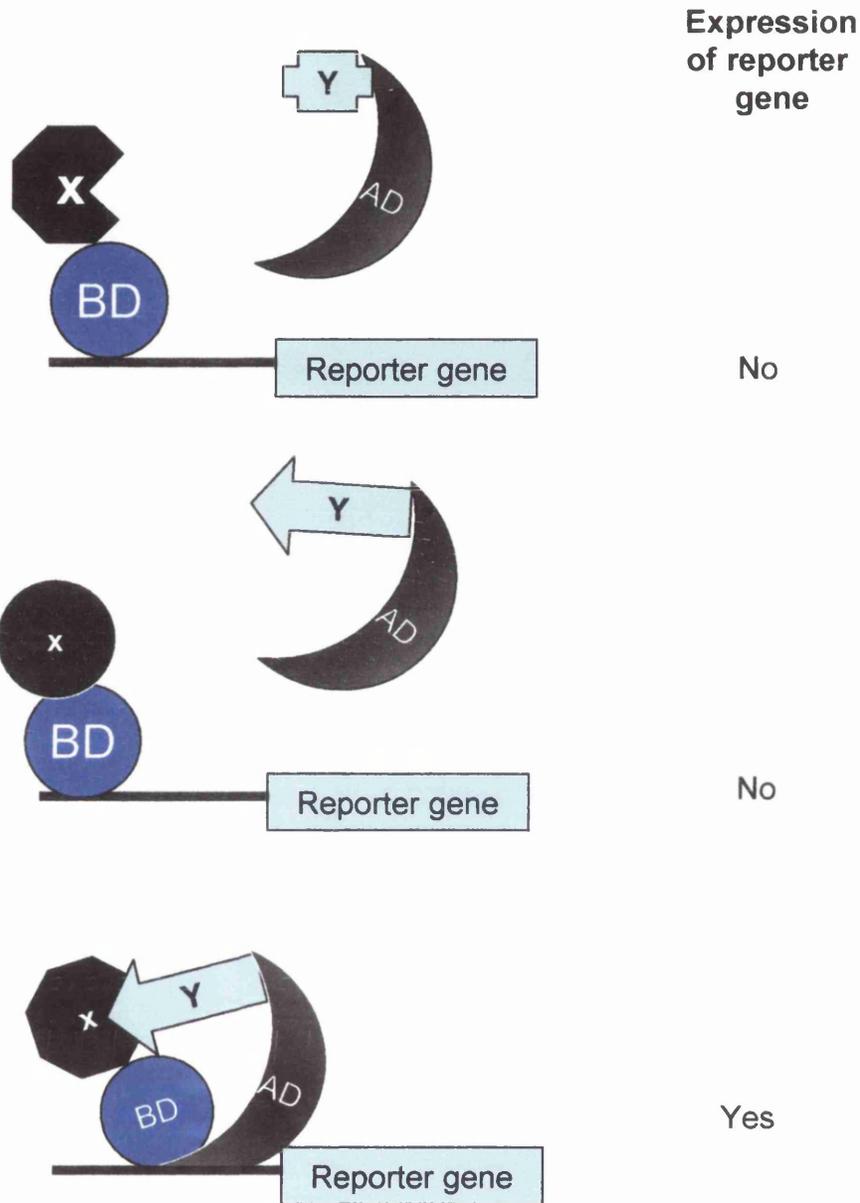


Figure 4.19 *The yeast-two hybrid system*

The yeast-two hybrid system consists of a bait protein containing a transcriptional binding domain (blue domain) fused to protein X. The system also contains a prey protein containing a fusion of a protein of interest, in this case protein Y with a transcriptional activator (black crescent moon shape). If the two proteins interact (C) transcriptional repression is relieved and the reporter gene is expressed. Transcriptional repression is not relieved if the two proteins fail to interact (A and B).

If protein X interacts with protein Y, in the nucleus, the AD will be brought together with the DNA BD to reconstitute transcriptional activation and hence gene expression. Positive interactions between proteins are identified by expressing the hybrid proteins within yeast strains that contain reporter genes, for example lacZ, that are under the control of regulatory regions that contain binding sites of the designated transcription factor. Although different transcription factors could be used, all available systems are based around the GAL4 or LexA transcription factors.

4.7.4.2 Constitutive GAL4-based yeast-two hybrid system

4.7.4.2.1 Construction of plasmids expressing 3D^{wt} and 3D^{N18Y} as fusion proteins with a GAL4- activation domain

To express 3D^{wt}, 3CD^{pro}, 3D^{N18Y} and 3CD^{N18Y} as fusions with the GAL4 activation domain (AD) the coding sequences were fused in-frame with the coding sequence of the GAL4AD in the plasmid vector pACTII (Figure 2.7). 3D^{wt} and 3CD^{pro} had previously been cloned into pACTII by Dr.I.Goodfellow.

The cDNA sequence of 3CD^{N18Y} was amplified from pT7FLCN18Y using the primers IG25 and IG26. The cDNA sequence of 3D^{N18Y} was amplified using the primers IG24 and IG25 (for primer details see materials and methods Table 6.2). Following amplification the purified PCR product was digested with the restriction enzymes *Not I* and *Bam HI*. The PCR product and pACTII, which had been digested in a similar way, were purified from agarose prior to being ligated together. The ligations were purified prior to transformation into *E.coli* JM109. The presence of insert was detected by digesting the plasmid recovered from the transformants with *Not I* and *Bam HI*. Plasmids positive for insert showed the presence of an additional band of 1.5 (3D^{N18Y}) or 2 (3CD^{N18Y}) kb in size in comparison to pACTII, which had been similarly digested with *Not I* and *Bam HI*, when visualised by agarose gel electrophoresis.

The pACTII constructs and the predicted molecular weights of the fusion protein expressed from the plasmids are summarized in table 4.1. The in-house sequencing service was used to confirm no additional coding changes had been introduced by the PCR amplification of the cDNA sequence of 3D^{N18Y} and 3CD^{N18Y}.

Plasmid name	Product expressed from plasmid	Predicted molecular Weight + AD or BD Tag (kDa)	Expression of fusion product	Laboratory database I.D. number
pGBKT7-53 ^a	p53 antigen fused to a GAL4 binding domain	68	Constitutive	P-AD-4A
pGBKT7-Lamin ^a	Lamin fused to a GAL4 binding domain	65	Constitutive	P-AD-5A
pGADT7-T ^a	SV40 T antigen fused to GAL4 activation domain	97	Constitutive	P-AD-3A
pGBKT7-3D ^{pol}	3D ^{pol} fused to a GAL4 binding domain	81	Constitutive	P-AD-6A
pGBKT7-3D ^{N18Y}	3D ^{N18Y} fused to a GAL4 binding domain	81	Constitutive	P-AD-7A
pACTII 3D ^{pol}	3D ^{wt} fused to GAL4 activation domain	72	Constitutive	P-AD-6B
pACTII 3CD ^{pol}	3CD ^{wt} fused to GAL4 activation domain	92	Constitutive	P-AD-3B
pACTII 3D ^{N18Y}	3D ^{N18Y} fused to GAL4 activation domain	72	Constitutive	P-AD-8A
pACTII 3CD ^{N18Y}	3CD ^{N18Y} fused to GAL4 activation domain	92	Constitutive	P-AD-9A
pHybLex/Zeo-Fos2 ^b	Fos2 antigen fused to a LexA binding domain		Constitutive	P-AD-1F

pHybLex/Zeo-Lamin ^b	Lamin fused to a LexA binding domain		Constitutive	P-AD-3F
pYESTrp2-Jun ^b	Jun antigen fused to B42 activation domain		Inducible	P-AD-2F
pHybLex/Zeo-3D ^{pol}	3D ^{pol} fused to a LexA binding domain		Constitutive	P-AD-9G
pHybLex/Zeo-3CD ^{pol}	3CD ^{pol} fused to a LexA binding domain		Constitutive	P-AD-1H
pHybLex/Zeo-3D ^{N18Y}	3D ^{N18Y} fused to a LexA binding domain		Constitutive	P-AE-5B
pHybLex/Zeo-3CD ^{N18Y}	3CD ^{N18Y} fused to a LexA binding domain		Constitutive	P-AE-4B
pYESTrp2-3D ^{pol}	3D ^{pol} fused to B42 activation domain		Inducible	P-AD-2H
pYESTrp2-3CD ^{pol}	3CD ^{pol} fused to B42 activation domain		Inducible	P-AD-8G
pYESTrp2-3D ^{N18Y}	3D ^{N18Y} fused to B42 activation domain		Inducible	P-AE-3B
pYESTrp2-3CD ^{N18Y}	3CD ^{N18Y} fused to B42 activation domain		Inducible	P-AE-1B
pYESTrp2-3AB	3AB fused to B42 activation domain		Inducible	P-AD-7G

Table 4.1. Yeast-two hybrid expression plasmids.

^a Positive and Negative controls for the constitutive Y2H system supplied by Clontech

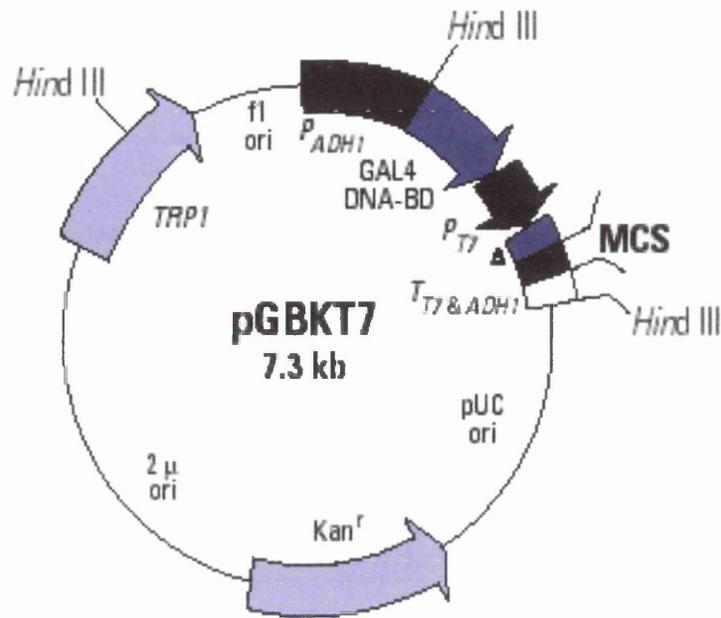
^b Positive and Negative controls for the inducible Y2H system supplied by Invitrogen

4.7.4.2.2 Construction of plasmids expressing 3D^{wt} and 3D^{N18Y} as fusion proteins with a GAL4 binding domain

To express as 3D^{wt}, 3CD^{pro}, 3D^{N18Y} and 3CD^{N18Y} fusions with the GAL4 binding domain (BD) the coding sequences were fused in-frame with the coding sequence of the GAL4 BD in the plasmid vector pGBKT7 supplied by Clontech (Figure 4.20). The cDNA sequences of 3D^{wt} and 3CD^{pro} were amplified from the template pT7FLC/Rep3 using the primer pairs IG24/IG25 and IG25/IG26 respectively. As has been previously described the cDNA sequences of 3D^{N18Y} and 3CD^{N18Y} were amplified from pT7FLCN18Y.

Following amplification the purified PCR product was digested with the restriction enzymes *Not I* and *Bam HI*. The PCR product and pGBKT7, which had been digested in a similar way, were purified from agarose prior to being ligated together. The ligations were purified prior to transformation into *E.coli* JM109. The presence of insert was detected by digesting the plasmid recovered from the transformants with *Not I* and *Bam HI*. Plasmids positive for insert showed the presence of an additional band of either 1.5 (3D^{wt} and 3D^{N18Y}) or 2 kb (3CD^{pro} and 3CD^{N18Y}) in size in comparison to pGBKT7, which had been similarly digested with *Not I* and *Bam HI*, when visualised by agarose gel electrophoresis.

The pGBKT7 constructs and the predicted molecular weights of the fusion protein expressed from the plasmids are summarized in table 4.1. Sequencing at the in-house facility confirmed that no coding changes had been introduced to the cDNA sequences of 3D^{wt} and 3CD^{pro}. The in-house sequencing service was also used to confirm no additional coding changes had been introduced by the PCR amplification of the cDNA sequence of 3D^{N18Y} and 3CD^{N18Y}.



▲ c-Myc epitope tag



Figure 4.20 Cloning vector pGBKT7. The information was supplied by Clontech. Restriction enzyme sites *Nco I* and *Bam HI* found in the multi-cloning site were used for the construction of all plasmids based on pGBKT7 (see table 4.1).

4.7.4.2.3 Expression of the GAL4 fusion-proteins in *S. cerevisiae* Y187

To confirm that the poliovirus fusion proteins were expressed in *S. cerevisiae* Y187 protein expression was determined by western blot analysis.

Though the viral proteins are fused to the (19.1 kD) GAL4AD, the pACTII vector also contained an influenza haemagglutinin epitope, YPYDVPDYAG, (Figure 2.7) that was expressed as part of the fusion protein. Similarly, though the coding sequences of the virus proteins 3D^{wt}, 3D^{N18Y}, 3CD^{pro} and 3CD^{N18Y} are fused to the sequence encoding the ~29 kD GAL4BD within the vector pGBKT7 a c-Myc epitope (EEQKLISEDL) present within pGBKT7 is expressed as part of the fusion protein (Figure 4.20). The plasmids, pACTII and pGBKT7 were co-transformed into *S. cerevisiae* Y187 using lithium chloride. Transformants were selected for using the nutritional markers encoded by these vectors. Cell extracts were prepared as described in materials and materials section 6.11.4.

To detect expression the fusion proteins immobilised on a nitrocellulose membrane were probed with a mouse monoclonal antibody that had been raised to the haemagglutinin epitope, diluted 1:5000 in 4% milk powder/PBS. The use of an antibody raised to the haemagglutinin epitope to detect the immobilised protein enabled the same antibody to be used to detect the expression of the different viral proteins. Following washing with PBS/0.1% Tween 20, the blot was incubated with HRPO conjugated anti-mouse IgG. The immunoreactive proteins was visualised after further washing using the Supersignal West chemiluminescent detection kit (Pierce). The nitrocellulose membrane onto which the proteins has been immobilised was then stripped and re-probed with a mouse monoclonal antibody that had been raised to the c-Myc epitope diluted 1:1500 in 4% milk powder/PBS. Following washing with PBS/0.1% Tween 20, the blot was incubated with HRPO conjugated anti-mouse IgG (Pierce). The immunoreactive proteins was visualised after further washing using the Supersignal West chemiluminescent detection kit (Pierce). No cross-reactivity of the monoclonal antibodies that had been raised to Haemagglutinin or c-Myc with *S. cerevisiae* Y187 proteins was observed.

The molecular weights of the fusions between the Gal 4 AD and 3D^{wt} was predicted to be 71.5 kDa. Monoclonal antibody raised to the haemagglutinin epitope could not detect the expression of 3D^{wt} or 3D^{N18Y} (Figure 4.21A lanes 3-6). A band of approximately 20.1 kDa was detected by the monoclonal antibody raised to the haemagglutinin epitope in the

cellular extract of *S.cerevisiae* Y187 that had been transformed with pACTII (Figure 4.21A lane 1). The 20.1 kDa band detected corresponds to the known molecular weight of the GAL4 activation domain. Immunoreactive bands corresponding to the predicted molecular weight of 3D^{wt} or 3CD^{pro} were however obtained when the nitrocellulose membrane, on which the proteins had been immobilised, was re-probed with mouse monoclonal antibody raised against the c-Myc epitope (Figure 4.21B lanes 3-6). An immunoreactive band of approximately 20.1 kDa was detected by the monoclonal antibody raised to the c-Myc epitope was detected in a cellular extract of *S.cerevisiae* Y187 that had only been transformed with pGBKT7 (Figure 4.21B lane 2).

From previous expression work using pACTII it was known that expression of 3D^{wt} by *S.cerevisiae* from this vector could be detected using western blotting (see section 2.3.1). Furthermore, it was shown that expression of 3D^{wt} was not toxic. The most likely explanation for these results, given that the GAL4 BD fusion proteins were detectable following stripping of the blot, is that the expression of the 3D^{wt}-GAL4 AD from pACTII is at a level beneath that which can be detected using immunoprobng. It is likely that it is this inequality in expression of the AD/3D^{pol} fusion proteins compared to the DNA BD/3D^{pol} hybrid, within *S.cerevisiae* Y187, detected via western blot, that explains why no interactions could be detected using a liquid β -galactosidase assay (data not shown).

Previously, it had been documented that interactions between the proteins of the P3 region could be successfully detected using a LexA-based yeast-two hybrid system (Xiang et al., 1998). As a consequence of this the decision was taken to introduce 3D^{N18Y}, 3D^{wt}, 3CD^{N18Y}, 3CD^{wt} and 3AB into a LexA-based yeast-two hybrid system. The LexA-system that was chosen was the Hybrid HunterTM supplied by Invitrogen that allows the inducible induction of AD fusion proteins.

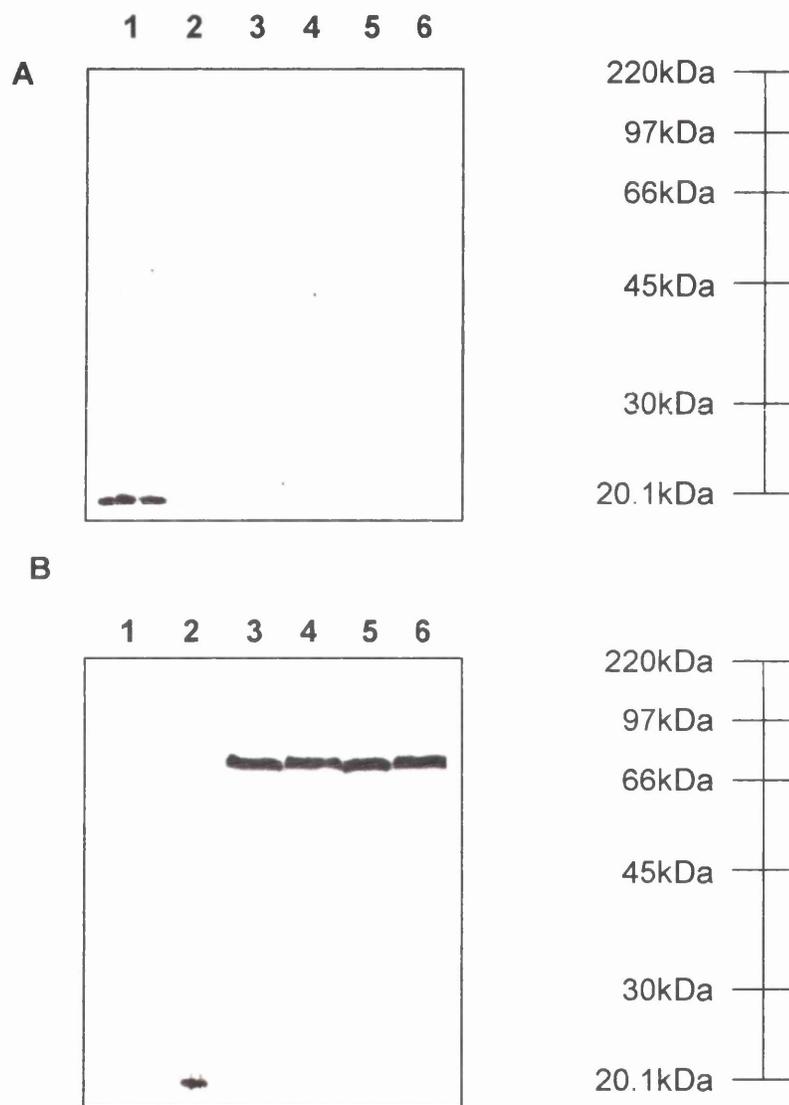


Figure 4.21 Western blot analysis of *S.cerevisiae* Y187 extracts

(A) Immunoreactive proteins detected by monoclonal antibody raised against the haemagglutinin epitope. (B) Immunoreactive proteins detected by monoclonal antibody raised against the c-Myc epitope. Cellular extracts of *S. cerevisiae* Y187 transformed with pACTII alone (Lane 1), pGBKT7 alone (Lane 2), pGBKT73CD^{N18Y} /pACT3D^{N18Y}(Lane 3) pGBKT73CD^{wt} /pACTII3D^{wt} (Lane 4), pGBKT73D^{wt} /pACTII3D^{wt} (Lane 5), pGBKT73D^{N18Y} /pACTII3D^{N18Y} (Lane 6).

4.7.4.3 Inducible LexA-based yeast-two hybrid system

4.7.4.3.1 Construction of plasmids expressing 3D^{wt} and 3D^{N18Y} as fusion proteins with a LexA binding domain

To express 3D^{wt}, 3CD^{pro}, 3D^{N18Y} and 3CD^{N18Y} as fusions with the LexA binding domain (BD) the coding sequences were fused in-frame with the coding sequence of the LexA BD in the plasmid vector pHybLex/zeo (supplied by Invitrogen) (Figure 4.22). The cDNA sequences of 3D^{wt} and 3CD^{pro} were amplified from the template pT7FLC/Rep3 using the primer pairs O-AC-2D/O-AC-3D and O-AC-1D/O-AC-3D respectively. Using the same primer pairs 3D^{N18Y} and 3CD^{N18Y} were amplified from the template pT7FLCN18Y (for primer details see materials and methods Table 6.2).

Following amplification the purified PCR product was digested with the restriction enzymes *Sac I* and *Sal I*. The PCR product and pHybLex/Zeo, which had been digested in a similar way, were purified from agarose prior to being ligated together. The ligations were purified prior to transformation into *E. coli* JM109. The *E. coli* JM109 transformed with pHybLex/Zeo were selected for using low salt LB agar containing zeocin at a concentration of 25µg/ml. The presence of insert was detected by digesting the plasmid recovered from the transformants with *Acc I*. Plasmids positive for 3D^{wt} and 3D^{N18Y} insert showed the presence of bands of 3.8, 1.5 and 0.82 kb while those containing 3CD^{pro} and 3CD^{N18Y} insert generated fragments of 3.8, 1.9 and 0.82 kb when visualised by agarose gel electrophoresis. Fragments of 3.8, 0.86 and 0.084 kb were obtained when pHybLex/Zeo, which had been digested with *Acc I*, was visualised by agarose gel electrophoresis

The pHybLex/Zeo constructs and the predicted molecular weights of the fusion protein expressed from these plasmids are summarized in table 4.1. Sequencing at the in-house facility confirmed that no coding changes had been introduced to the cDNA sequences of 3D^{wt} and 3CD^{pro}. The in-house sequencing service was also used to confirm no additional coding changes had been introduced by the PCR amplification of the cDNA sequence of 3D^{N18Y} and 3CD^{N18Y}.

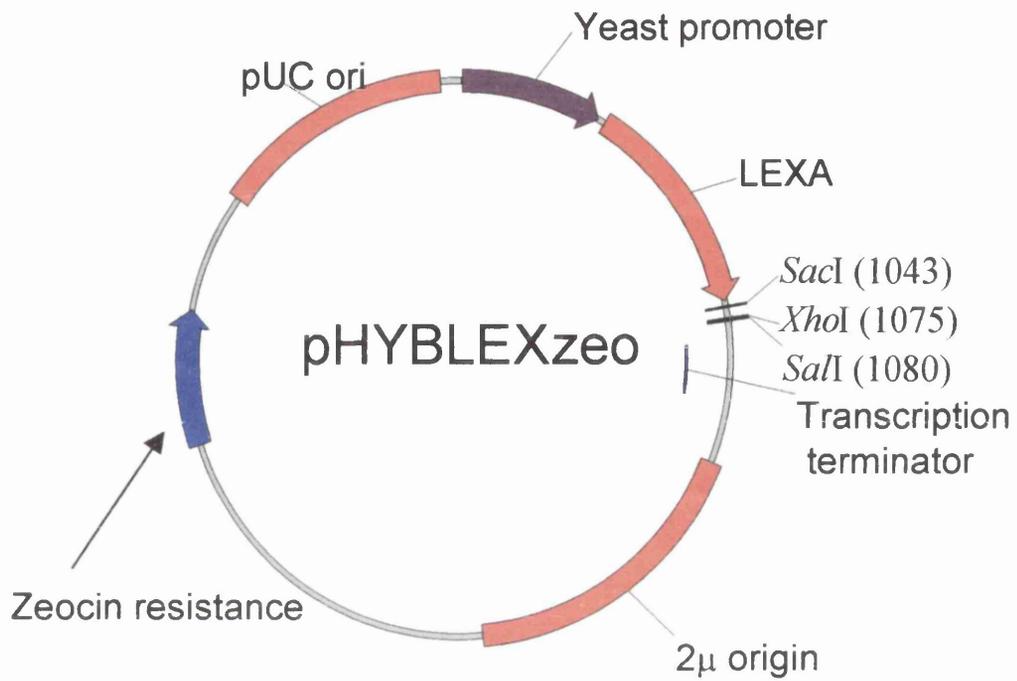


Figure 4.22 Vector map of pHYBLEXzeo

Restriction enzyme sites used in the construction of pHYBLEXzeo yeast two-hybrid expression plasmids are shown in italics.

4.7.4.3.2 Construction of plasmids expressing viral non-structural proteins as fusion proteins with a B42 activation domain

To investigate the intramolecular and intermolecular interactions of 3D^{wt} and 3D^{N18Y} the non-structural proteins 3AB, 3D^{wt}, 3CD^{pro}, 3D^{N18Y}, 3CD^{N18Y} were introduced into a plasmid pYESTrp2 (supplied by Invitrogen), that would express the proteins as a fusion with the B42 activation domain, a known activator of LexA transcription factor (Figure 4.23). The expression of fusion proteins from pYESTrp2 is constitutive in *S.cerevisiae* L40 but occurs in an inducible manner in *S.cerevisiae* EGY48 strains. In *S.cerevisiae* EGY48 induction of the expression of the fusion proteins from pYESTrp2 is controlled by the presence of glucose in the culture medium. In the absence of glucose, the addition of galactose to the culture medium releases the transcriptional repression of the GAL1 promoter controlling the expression of the fusion proteins.

As has been described before the cDNA sequences of 3D^{wt} and 3CD^{pro} were amplified from the template pT7FLC/Rep3 using the primer pairs O-AC-2D/O-AC-3D and O-AC-1D/O-AC-3D respectively. Using the same primer pairs 3D^{N18Y} and 3CD^{N18Y} were amplified from the template pT7FLCN18Y (for primer details see materials and methods Table 6.2). The cDNA sequence of 3AB was derived from pT7FLC/Rep3 through PCR amplification using the forward primer O-AC-9C and the reverse primer O-AC-5D.

Following amplification of 3D^{wt}, 3CD^{pro}, 3D^{N18Y} and 3CD^{N18Y} the purified PCR product was digested with the restriction enzymes *Sac I* and *Sal I*. The plasmid vector pYESTrp2 and the purified product of the PCR amplification of 3AB were digested with the restriction enzymes *Sac I* and *Xho I*. The individual PCR products and pYESTrp2 were purified from agarose prior to being ligated together. The ligations were purified prior to transformation into *E.coli* JM109. The presence of insert was detected by digesting the plasmid recovered from the transformants with *Ava I*. When the products of digestion were visualised by agarose gel electrophoresis plasmids positive for 3D^{wt} and 3D^{N18Y} generated fragments of, 4.2, 2.0 and 0.8 kb while fragments of 4.2, 2.0, 0.8 and 0.6 kb were obtained for plasmids positive for 3CD^{pro} and 3CD^{N18Y} insert. Bands of 2.9, 2.3 and 0.8 kb were observed in plasmids containing the 3AB insert. The fragments obtained were visualised and compared by agarose gel electrophoresis with those of pYESTrp2 which when digested with *Ava I* generates bands of 2.9, 2.0 and 0.8 kb in size.

The pYESTrp2 constructs and the predicted molecular weights of the fusion protein expressed from the plasmids are summarized in table 4.1. Sequencing at the in-house facility confirmed that no sequencing changes had been introduced to the cDNA sequences of 3AB, 3D^{wt} and 3CD^{pro}. The service was also used to confirm no additional sequencing changes had been introduced by the PCR amplification of the cDNA sequence of 3D^{N18Y} and 3CD^{N18Y}.

Due to time limitations, analysis of the intra- and intermolecular interactions of the viral polymerases (3D^{N18Y} and 3D^{wt}) using the LexA yeast-two hybrid system was not completed. Although expression of the fusion proteins has been detected by western blotting (data not shown) at the current time, the conditions required for the optimal expression of proteins using the Hybrid HunterTM system and the optimal time to assay expression of the lacZ reporter gene are still being deduced.

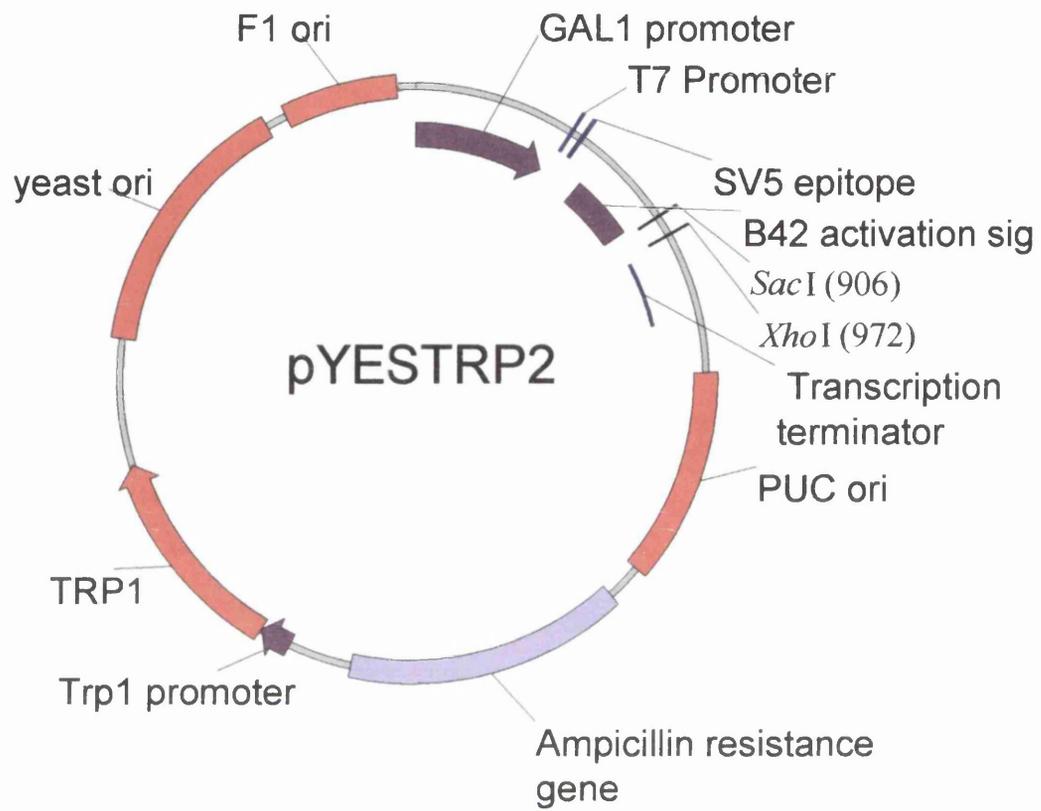


Figure 4.23 Vector map of pYESTrp2

Restriction enzyme sites used in the construction of pYESTrp2 yeast two-hybrid expression plasmids are shown in italics.

4.8 Discussion of results

All RNA viruses require an RdRp to catalyse the replication of their genomes. Aside from the fingers, thumb and palm domain common to all types of polymerase RdRp contain a unique N-terminal region not found in other types of polymerase (RT, DdDp and DdRp). This would argue that the function of the N-terminus has evolved in relation to the RdRp unique property of transcribing RNA from an RNA template.

In the work published by Meredith *et al* a coding change in the N-terminus region was found to be the sole change required to rescue a poorly replicating virus that contained a mutated 3'UTR. To characterise the mechanism by which the N18Y coding change can compensate the mut4 3'UTR a bacterial expression system was used to express recombinant poliovirus type 3 RdRp containing the N18Y mutation. The parental poliovirus type 3 RdRp was also expressed using this system. Consistent with the published literature the majority of the expressed protein was located in the soluble fraction. Both 3D^{wt} and 3D^{N18Y} were demonstrated to bind to SP-sepharose. Elution of 3D^{N18Y} was shown to be more sensitive to salt than 3D^{wt}. The preliminary characteristics of 3D^{wt} and 3D^{N18Y} confirmed that the purified proteins were functional.

4.8.1 Uridylylation

In picornaviruses, initiation of RNA synthesis requires a protein primer. In contrast to the picornaviruses, flaviviruses initiate RNA synthesis *de novo*. Reilly and Kao presented data that showed that amongst the virus families that contain a unique N-terminal region were the flavivirus, that includes HCV. This data suggested that the function of the N-terminus is related to a feature common to these virus families and so the N18Y mutation was unlikely to have an effect on the ability of 3D^{wt} to uridylylate VPg. Consistent with this hypothesis is the data that has been presented in this chapter. Using the uridylylation assay described by Paul *et al* (Paul et al., 2000) the preliminary data confirmed that no difference could be detected in the ability of purified 3D^{wt} and 3D^{N18Y} of catalysing the uridylylation of VPg to form VPg-pU. The preliminary data did suggest that a difference did exist between 3D^{wt} and 3D^{N18Y} in their ability to catalyse the formation of VPg-pUpU from VPg-pU. One possible explanation for this phenotype is that there is a difference in the

proportion of polymerase molecules that are “active” in uridylylation. This could have arisen as a consequence of the bacterial expression system used.

The expression of recombinant protein in bacteria requires the presence of polypurine Shine-Dalgarno sequence on the mRNA transcript (AGGAGG) that is capable of binding the 50s ribosomal subunit. The ribosome scans the RNA sequence until it reaches the first AUG downstream of the Shine-Dalgarno sequence where protein synthesis is initiated by the binding of N-formyl-methionyl-tRNA in the ribosomal P site. The process by which an mRNA transcript is translated within mammalian cells is similar to that which occurs in prokaryotes though it does not require the presence of a Shine-Dalgarno sequence. In a poliovirus-infected cell the translation complex is assembled on the IRES and the ribosomes scan the genome for the first AUG whereby translation is initiated. The translation of the poliovirus genome results in the production of a polyprotein that is co-translationally processed by the proteases encoded by the virus. The poliovirus polymerase is expressed by the processing of the polyprotein by 3C^{pro} that recognises a glutamate/glycine cleavage site. Under normal circumstances the first residue of 3D^{wt} is therefore a glycine rather than a methionine residue. In *E. coli* the formyl-methionine encoded by the initiation codon is removed in a two-step process involving the enzymes cellular deformylase and methionine aminopeptidase (Ben-Basset et al., 1987, Rajagoplan et al., 1997). However this process does not occur with 100 % efficiency (Sandman et al., 1995). As a result, a proportion of the purified recombinant protein will have retained the amino-terminal methionine. Recently, it has been reported that purified 3D^{pol} with a non-authentic amino-terminal shows a 25-fold reduction in polymerisation activity when compared with the wild-type activity (Gohara et al., 1999). Although, no data has been published on the effect an amino-terminal methionine has on the process of uridylylation it is possible that this influences the ability of the polymerase to catalyse the formation of VPg-pUpU. If this the case then the variation in the ability of 3D^{wt} and 3D^{N18Y} to catalyse the formation of VPg-pUpU reported in this preliminary characterisation may be explained by differences in the proportion of enzyme with an authentic amino-terminus within each preparation. To rule out this possibility, polymerase with an authentic 5' end should be purified and characterised using the *in vitro* uridylylation assay. The ubiquitin-based purification system developed by Gohara *et al* (Gohara et al., 1999) provides a method by which this could be achieved. This system will be discussed in more detail in section 4.8.3.

4.8.2 Oligomerisation

It is known that deletion of the N-terminal residues of the HCV polymerase and those of 3D^{wt} eliminates polymerase function (Gerber et al., 2001, Hansen et al., 1997, Lohmann et al., 1997). The observation that deletion of the N-terminal residues abolishes function of two distantly-related polymerases confirms that the N-terminal performs a role fundamental to the functioning of the RdRp of mammalian RNA viruses. One property common to the HCV polymerase (Wang et al., 2002) and 3D^{wt} (Hansen et al., 1997, Hobson et al., 2001, Pata et al., 1995) is the ability of the polymerase molecules to oligomerise. Support for the view that the N-terminal region of the polymerase may be an essential requirement for oligomerisation to occur is provided by the observation that the Q β -like virus polymerases, which are not known to form oligomers, do not contain a region N-terminal of the fingers domain (Reilly & Kao, 1998).

A number of approaches can be used to study the oligomerisation of polymerase molecules. The approach that was taken to study this aspect of the polymerase was the yeast two-hybrid approach. It was decided to take this approach as it remained a possibility that the N18Y mutation exerted its compensating effect through the polymerase-precursor 3CD^{N18Y}. Additionally it is equally possible that the compensating effect is exerted through the interactions of 3D^{N18Y} with 3CD^{N18Y} or as a result of an alteration in the interaction between 3AB and 3D^{N18Y} or 3AB and 3CD^{pro}.

Initially the decision was taken to investigate the homo- and heteromultimeric interactions of the polymerase 3D^{N18Y} and 3D^{wt} using a GAL4-based yeast two hybrid system. Interaction of the GAL4-binding domain and activation-domain fusion proteins within *S.cerevisiae* Y187 was not detectable by liquid β -galactosidase ONPG assay. Western blotting confirmed that the GAL4-binding domain fusion proteins were expressed by the *S.cerevisiae* Y187 strain. In contrast, the expression of the GAL4-activation domain fusion-proteins could not be detected. Expression of the GAL4-activation domain fusion-proteins from the plasmid pACTII is detectable in the *S.cerevisiae* R40coat strain. The failure to detect the expression of the GAL4-activation domain fusion-proteins in *S.cerevisiae* Y187 is most likely reflects low-level expression of the protein rather than a complete absence of synthesis of the GAL4-activation domain fusion-proteins within the cell. To overcome this problem 3D^{N18Y}, 3D^{wt}, 3CD^{N18Y}, 3CD^{pol} and 3AB have been built

into a LexA-based yeast two-hybrid system supplied by Invitrogen that enabled the inducible expression of the LexA-activation domain fusion-proteins. The decision was taken to introduce the proteins into a LexA-based system as this approach had been successfully used by Xiang *et al* (Xiang et al., 1998). Due to time constraints, this aspect of the project was not followed through to completion. It is hoped that following further optimisation of the conditions for the expression of the AD/polymerase fusion proteins and the β -galactosidase assay used to determine positive interactions this system will provide valuable insights into any differences that may exist in the intra- and intermolecular interactions of 3D^{N18Y} and 3D^{wt}.

Approaches other than the yeast-two hybrid system have been developed that enable oligomerisation to be studied. One such approach is to use the turbidity of a solution of purified polymerase as a measure of oligomerisation (Figure 4.24). Lyle *et al* showed that an increase in turbidity, as determined by measuring the optical density of 350 nm over time, of a solution of purified polymerase was dependent on the ability of the polymerase to form Interface I and II interactions (Lyle et al., 2002a). Therefore any disparity in the ability of 3D^{N18Y} to oligomerise, compared with 3D^{wt}, should be reflected by variations in the turbidity profiles of the solutions of purified polymerase. Furthermore, the use of EM analysis may highlight differences in the lattice structures formed by the 3D^{wt} and 3D^{N18Y} in the presence and absence of RNA.

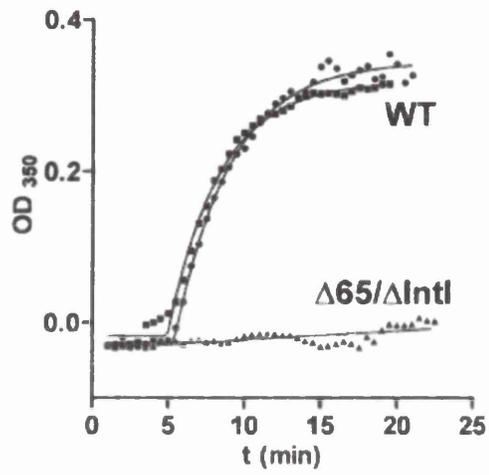


Figure 4.24 *The use of turbidity as a measure for oligomerisation.* Taken from Lyle et al., 2002. Optical density of solutions of wildtype (WT) and interface I and II mutants ($\Delta 65/\Delta \text{IntI}$).

4.8.3 Polymerisation

Two further functions common to all RdRp are the preferential binding of NTP and the catalysing of the incorporation of these nucleotides into a nascent strand. As a preliminary characterisation the ability of 3D^{wt} and 3D^{N18Y} to catalyse the elongation of a poly-U primed poly-A template was measured by incorporation of [α -³²P] UTP. The result of the elongation reaction suggested that 3D^{N18Y} catalysed the incorporation of [α -³²P] UTP at a higher rate than 3D^{wt}. However recent data obtained by Dr.I.Goodfellow has suggested that this result may be an artefact of the expression/purification system and template used in this study.

As has been discussed previously, the presence of a non-authentic 5' end was shown to reduce the activity of purified 3D^{wt}, as determined by poly (rU) polymerase assay (Gohara et al., 1999). Systems have been developed that allow protein with a fully authentic 5' end to be purified. One such system is the two-step purification system, developed by Gohara *et al* (Gohara et al., 1999), that is based around a ubiquitin-tag. In the first step the cell lysate is passed over a column that binds the ubiquitin-tagged recombinant protein. Following washing the bound protein is eluted from the column. In the second stage the ubiquitin-tag is removed from the 5' end of the purified protein by enzymatic treatment. 3D^{wt} and 3D^{N18Y} purified using the ubiquitin-tagged system was used in the polymerisation assays carried out by Dr. I. Goodfellow. One other difference between the polymerisation assay that was described here and that carried out by Dr.I.Goodfellow was the template. The template used by Dr.I.Godfellow was the sym/sub template previously described by Arnold *et al* (Arnold & Cameron, 2000). This template consists of a 10nt self-complementary RNA that forms a 6 nt duplex flanked by 4 nt overhangs. Each overhang of the sym/sub template is capable of templating the incorporation of a unique nucleotide. This allows single and multiple cycles of nucleotide incorporation by a polymerase to be calculated at stoichiometric levels. It has been reported that the template used to study polymerase activity can make a difference to the results obtained. One such report has shown that while a mutation in motif B of the palm domain only reduces polymerase activity 2-fold poly(rU) polymerase activity, the same mutation using the sym/sub system is reduced 15-fold in activity (Arnold & Cameron, 2000).

One possible explanation for this is the report that on heteropolymeric RNA primer-template 3D^{wt} can bind in two orientations. The first of these has the 3'OH of the primer in the catalytic centre of the polymerase. This is the “correct” orientation. The second way that 3D^{wt} can bind is with the 3'OH of the template strand in the catalytic centre, the so called “wrong” orientation. It has been reported by Arnold *et al* that 3D^{wt} bound in the “wrong” orientation can catalyse the addition of nucleotides to the blunt end of the template/primer (Arnold *et al.*, 1999). Characterising 3D^{wt} using a heteropolymeric RNA primer-template is therefore complicated by the fact that a proportion of the 3D^{wt} molecules in the reaction will be bound in the “wrong” orientation.

4.8.4 Recycling of the polymerase

Both the massive amplification of genomes observed in infected-cells and replication-dependent recombination require the recycling of polymerase molecules. The process of recycling of polymerase molecules i.e. the disassociation of the polymerase from one template to re-associate and re-initiate transcription on another is illustrated in Figure 4.25. In the absence of a limiting amount of enzyme in a reaction this cycle should ensure that all primer-template is extended. However a number of research groups have independently reported that *in vitro* only a low percentage of all primers are extended by 3D^{wt} under non-limiting conditions (Arnold & Cameron, 1999, Arnold & Cameron, 2000, Pata *et al.*, 1995, Rodriguez-Wells *et al.*, 2001). This can be explained by three *in vitro* biochemical observations.

- 3D^{wt} in the absence of nucleotides is turned over rapidly - $t_{1/2}$ of ~ 1 min at 42°C (Richards *et al.*, 1992)
- The affinity of 3D^{wt} for the primer-template is low ($K_d = \sim 1\mu\text{M}$) compared to the affinity of the majority of polymerases that are in the nanomolar range (Arnold & Cameron, 1999, Arnold & Cameron, 2000). The high K_d value ($K_d = k_{\text{off}}/k_{\text{on}}$) observed for 3D^{wt} indicates that the association of the polymerase with the template (low k_{on} value) and the formation of the active catalytic complex occurs slowly.

- The k_{off} value of 3D^{wt} is low consistent with slow release of the polymerase from the template (Rodriguez-Wells et al., 2001)

If these factors are looked at in combination the following model is suggested. The low affinity for the primer-template limits the association of 3D^{wt} to the primer-template. The enzyme that remains unbound following this initial association will rapidly become inactive. Following elongation of the genome (~ 7.5 kb in length), which *in vitro* biochemical data implies occurs in the majority of occasions, from a single binding event (Rodriguez-Wells et al., 2001), the recycling of the polymerase occurs slowly. Re-association of 3D^{wt} is subject to the same low rates of association and rapid inactivation of polymerase molecules. Thus the low-level of primer extension observed *in vitro* relates directly to the low rate at which the polymerase is recycled.

It is likely however that *in vivo* the rate of recycling of 3D^{wt} is greater than then the observed rate *in vitro*. *In vivo* 3AB is purported to tether the polymerase onto the membranes of the “rosette-like” replication complexes (Lyle et al., 2002a). *In vitro* analysis demonstrated that 3D^{wt} is stimulated by the multifunctional membrane-bound viral protein 3AB (Plotch et al., 1989). The stimulation of 3D^{wt} by 3AB corresponds to an increase in processivity (the average number of nucleotides added to the primer in a single-binding event between the primer-template and polymerase) of 3D^{wt} from 5400 to 18,000 nt (Rodriguez-Wells et al., 2001). Biochemical analysis of the stimulatory effect of 3AB has shown that the disassociation constant (k_{off}) molecules of 3D^{wt} ($k_{\text{off}} = 0.011 \text{ min}^{-1}$) was reduced by ~3-fold in the presence of 3AB ($k_{\text{off}} = 0.037 \text{ min}^{-1}$) (Rodriguez-Wells et al., 2001). In addition to stabilising the 3D^{wt}-template/primer complex experimental data obtained also showed that the 3AB/3D^{wt} interaction also enhanced the rate of initial association of the polymerase with the primer/template (Rodriguez-Wells et al., 2001). Furthermore, it is known that the concentration of nucleic acids and viral proteins in the vicinity of the replication complexes is high. The association of 3D^{wt} with these membrane structures and allied factors may substantially increase the stability of the polymerase *in vivo*.

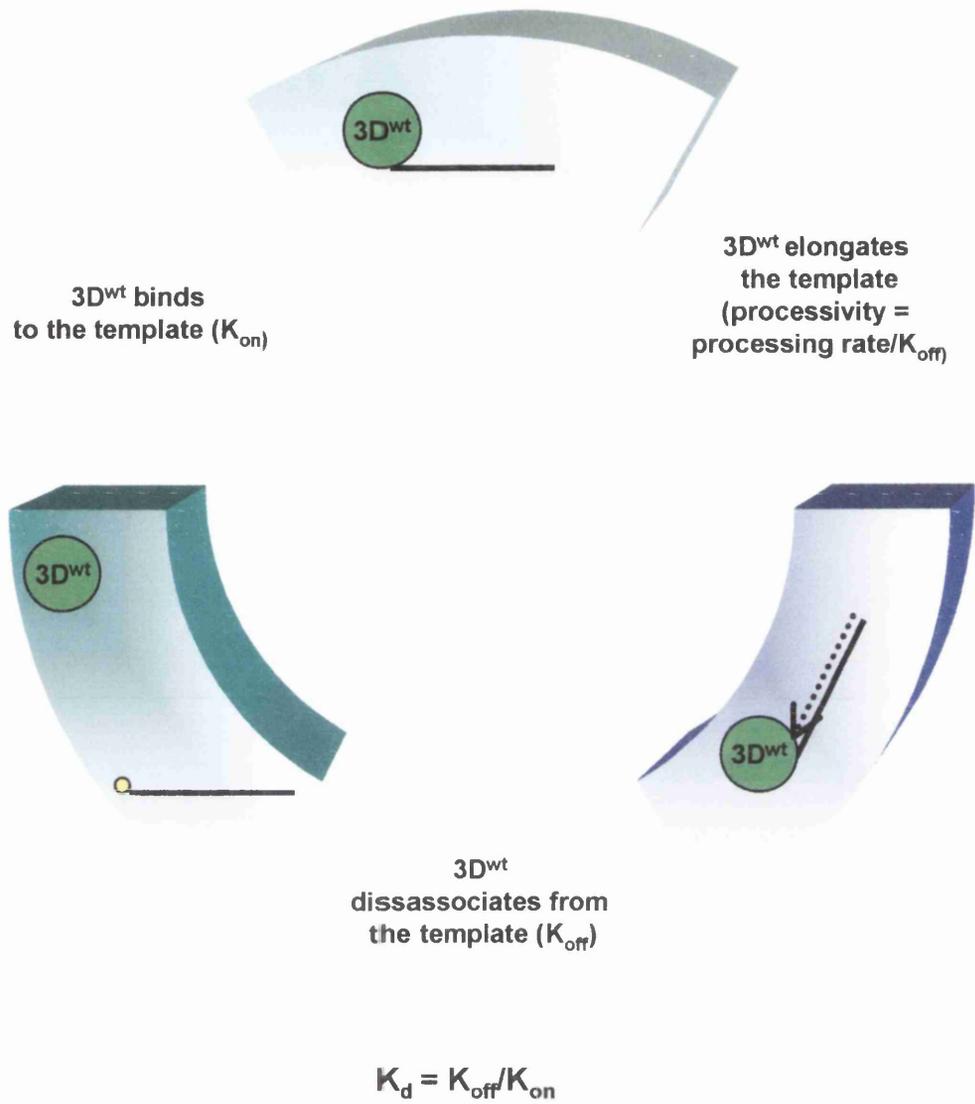


Figure 4.25 Illustration of the process of recycling of 3D^{wt}

If the lattice model proposed by Lyle *et al* is an accurate representation of the replication complex then these factors suggest that the k_{on} rate *in vivo* would be higher than the calculated rate *in vitro*. The increase in the k_{on} value would lead to an increase in the rate of polymerase recycling by enhancing the ability of the polymerase following disassociation for the template. Certainly the ability of 3D^{wt} to recycle efficiently *in vivo* is suggested by the rate at which replicative recombination occurs (Jarvis & Kirkegaard, 1992).

4.8.5 RNA binding

Biochemical analysis has shown that the stimulation of the polymerase upon binding of 3AB to 3D^{wt} is associated with an increase in binding of the 3AB/3D^{wt} complex to the primer-template (Rodriguez-Wells *et al.*, 2001). *In vitro* both 3AB and 3D^{wt} show non-specific RNA binding activity. It is of interest to deduce the mechanism by which the 3AB/3D^{wt} super-polymerase *in vivo* specifically recognises the 3' termini of the poliovirus genome and anti-genome. Three possible explanations for this have been put forward. The first is the existence of an unidentified cellular factor that promotes the interaction of the 3AB/3D^{wt} complex with a specific structure within the template. A second explanation of the specific-recognition of the poliovirus genome as a template for RNA synthesis would be the exposure of a motif within the polymerase, upon binding of 3AB that specifically recognises a specific structure at the 3' termini of the picornavirus genome and anti-genome. Lastly, it is equally possible that the polymerase, though showing a relatively low affinity for nucleic acid binding *in vitro* (Arnold & Cameron, 2000), does have a higher affinity for the virus genome. This increase in affinity for genomic template could be mediated by a direct interaction between 3D^{wt} and a RNA secondary structure present within the genome. A slight increase in the binding affinity of the polymerase for the genomic template in the context of the replication complex, which contains a high localised concentration of virus factors and nucleic acids, could be enough to confer specificity to the process of RNA synthesis.

Support for the idea that N18 may be involved in RNA-binding has been provided by research on RT and 3D^{wt} that implicates the thumb domain that is adjacent to N18 in the crystal structure of 3D^{wt}, in binding the RNA template. Additional supporting evidence can be drawn from structural alignments of HCV NS5B and 3D^{wt}. Evidence from the structural alignments (Bressanelli *et al.*, 1999) indicates that N18 is located in an area of the polymerase analogous to a region of NS5B that has been shown to be in direct contact with

the template strand (O'Farrell et al., 2003). Given that an increase in the affinity of the “superpol” for its template correlates to a stimulation of 3D^{wt} polymerisation activity one explanation for the phenotype, of PV3.HRV14 3'UTR.3D^{N18Y}, observed by Meredith *et al* (Figure 4.13) is that the N18Y mutation positively influences the k_{on} rate of the RdRp. An increase in k_{on} rate of 3D^{wt} could be by enhancing the inherent ability of the polymerase to bind the 3'UTR. It is also possible that the N18Y mutation increases the rate at which the initial isomerisation reaction to form the active catalytic complex occurs. A positive influence on the k_{on} rate through either of the above mechanisms should be reflected in a higher affinity for the primer-template i.e. a reduction in the K_d value of the polymerase should be obtained.

In HCV residues 14, 93, 95, 97, 98, 139, 141 and 160 directly contact the RNA (O'Farrell et al., 2003). By directly contacting the RNA the polymerase can position the template so that it is positioned correctly with respect to the primer and the incoming nucleotide. The data obtained by Meredith *et al* suggested that the N18Y repaired a direct-interaction between the 3'UTR and polymerase that had been disrupted in viruses with a mut4 3'UTR (Meredith et al., 1999). By drawing an analogy with the HCV-liganded structure two models can be put forward to describe how an effect on RNA-binding could explain the phenotype observed by Meredith *et al*. The mutations introduced by Meredith *et al* into the stem of the HRV14 3'UTR to create mut 4 disrupted the interaction between N18 and the RNA. Based on the HCV-liganded structure it would be expected that multiple contacts between the polymerase and the RNA exist. As a consequence the disruption of the N18/3'UTR interaction would not be predicted to completely abolish binding. The phenotype of virus with a mut4 3'UTR, in the absence of the N18Y mutation, can be explained by the following scenario. Disruption of the N18/3'UTR enables the template to “slip” within the polymerase. In the HCV-liganded structure the template and incoming nucleotide fit tightly in RNA-binding groove and NTP tunnel respectively (O'Farrell et al., 2003). If the genome with the mut4 3'UTR was to “slip” within the polymerase molecule it is likely that it would either slip into the binding site of the incoming nucleotide or towards the thumb domain. Both of these scenarios would result in a reduction in the k_{on} value, by decreasing the rate at which a productive catalytic complex would be formed. It is also possible that slipping of the template might increase the k_{off} rate. A decrease in the affinity of the polymerase for the genomic RNA *in vivo*, either by increasing the k_{off} rate or decreasing the k_{on} rate would reduce the efficiency of polymerase recycling. As a

consequence the titre of virus progeny produced from a virus with a mut4 3'UTR would be expected to be decreased relative to the parental virus.

It is clear that the N18Y mutation must reform the disrupted interaction between the polymerase and the virus 3'UTR in the case of PV3.mut4 3'UTR.3D^{N18Y}. However, what is the possible explanation for the effect exerted on the growth of PV3.HRV14 3'UTR by the presence of 3D^{N18Y} (Meredith et al., 1999)? One possibility is that the N18Y coding change results in a subtle narrowing of the RNA-binding groove around residue 18. This subtle narrowing may be enough to prevent the RNA with a mut4 3'UTR from slipping out of place. In virus with a parental 3'UTR this subtle narrowing may increase the stability of the 3D^{N18Y}/template complex. The effect of this may be an increase in the rate at which a productive catalytic complex is formed. It is also possible that a subtle narrowing of the RNA-binding groove may decrease the k_{off} rate of the polymerase.

Alternatively, the presence of a tyrosine at residue 18 may increase the inherent affinity of the polymerase for the genomic template. If the presence of a mut 4 3'UTR on a genomic RNA resulted in a higher rate of disassociation of the polymerase for the template this could be compensated by increasing the rate at which the RdRp binds the template. An example of this type of compensation is provided by RT which shows characteristic low processivity and high k_{on} values (DeStefano et al., 1992, Huber et al., 1989, Reardon, 1993). An increase in the inherent k_{on} value of 3D^{N18Y}, in comparison with 3D^{wt}, would result in an increase in fitness of the virus by enhancing the rate of polymerase recycling.

Equations exist that enable the association (k_{on}), the disassociation (k_{off}) and affinity (K_d) constants of polymerases for template *in vitro* to be calculated. If N18Y compensates the mut 4 3'UTR through a non-specific increase in the inherent k_{on} value then this should be detectable *in vitro* using sym/sub. In addition the enhancement in the recycling rate that accompanies an increase in k_{on} value should result in an increased percentage of primer-template being extended by 3D^{N18Y}, compared to 3D^{wt}, over the course of a standard reaction *in vitro*. If the effects of N18Y are affected by a subtle narrowing of the RNA-binding groove it may be that differences exist in the k_{on} and k_{off} values of 3D^{wt} and 3D^{N18Y}. However as the 3'UTR is structured it may be that to fully characterise any differences in the k_{on} and k_{off} values between 3D^{wt} and 3D^{N18Y} a template will have to be designed that more fully mimics the mut 4 and parental HRV14 3'UTR tailed genomes.

4.8.6 Conclusion

During the course of the chapter the expression and purification of functional 3D^{wt} and 3D^{N18Y} has been described. The preliminary characterisation of polymerisation activity appeared to show that 3D^{N18Y} was catalytically more active than 3D^{wt}, though recent characterisation has not been able to confirm this. Further characterisation of polymerase oligomerisation and of the polymerase-3'UTR interaction using purified 3D^{wt} and 3D^{N18Y} polymerases will hopefully enable the biochemical basis of the phenotype observed by Meredith *et al* to be elucidated.

However in characterising purified 3D^{N18Y} it has to be remembered that the *in vitro* techniques used to analyse the purified polymerases does not completely reconstitute the *in vivo* situation where 3D^{wt} appears to be found in large lattice structures, built on a membrane scaffold. It therefore may be that subtle variation in an activity, for example terminal uridylyl transferase or RNA binding, between the purified polymerases *in vitro* may be reflected *in vivo* by a virus that can replicate more efficiently in the presence or absence of the mut 4 3'UTR.

5 Discussion of results

Since the worldwide vaccination campaign against poliomyelitis was initiated remarkable advances in eradicating the disease have been made. According to World Health Organisation (WHO) statistics the number of cases of poliomyelitis has decreased from 350,000 in 1988 to less than a 784 in 2003 (http://www.who.int/vaccines/case_count.cfm). Problems encountered during the eradication campaign include the low symptomatic: asymptomatic infection ratio, political interference and the number of viruses, amongst them the coxsackie A viruses, which result in similar clinical outcomes (flaccid paralysis) to symptomatic poliovirus infections.

Within the picornavirus family nine distinct genera have so far been defined. Within certain genera further distinctions can be made. Phylogenetic analysis of the enterovirus genus has shown that individual enteroviruses cluster into the same four species regardless of whether the phylogeny is based on the structural or non-structural genes (Poyry et al., 1996). Phylogenetic analysis has shown that members of the HEV-C species cluster, which includes both poliovirus and some coxsackie A viruses, can only be distinguished from each other by receptor usage (Brown et al., 2003). Evidence from outbreaks of vaccine-associated disease, like the recent case in Hispaniola in which the entire non-structural region of the isolate's genome had been derived from HEV-C viruses (Figure 5.1), suggest that all HEV-C viruses capable of using PVR can cause paralytic poliomyelitis (Kew et al., 2002).

Oberste and others have demonstrated using phylogenetic analysis that whilst intra-species recombination occurs frequently within the non-structural region of the genome the structural proteins (VP1-4) are always inherited as a single unit. Oberste *et al* therefore suggested that enteroviruses could be viewed either as replicons in search of a vehicle that allows them to access the host machinery required for virus replication or as capsid proteins looking for the non-structural proteins of the highest replication fitness (Oberste et al., 2004). Several independent studies have reported that while intra-species recombination is highly prevalent amongst circulating enteroviruses no inter-species recombination can be detected (Lukashev et al., 2003, Oberste et al., 2004). Thus, while enteroviruses can be viewed as replicons in search of capsid sequences it would appear that recombination is restricted to species level. There are a number of possible explanations

for why recombinants between more distantly related enteroviruses are not observed in circulating viruses.

An absolute requirement is that two enteroviruses genomes must be within the same cell. This is defined by receptor usage. At a phylogenetic level this has been used as the explanation as to why distantly related serotypes are significantly less likely to recombine with each other than more closely related serotypes (Brown et al., 2003, Poyry et al., 1996, Pulli et al., 1995). However as the receptors used by enteroviruses (Table 1.4) are fairly ubiquitous in their expression on the cell surface it is likely that co-infection of cells is more common than the recombination rate suggests.

A second explanation for the apparent species-block is incompatibility between the viral proteases that process the viral polyprotein and the heterologous polyprotein resulting from recombination between enterovirus species. A compatibility requirement such as this has been observed between enterovirus species *in vitro* where it has been demonstrated that 3C^{pro} of CVB3, in the context of the full-length genome, was incapable of processing the structural proteins of PV1 (Dewalt et al., 1989, Lawson et al., 1990). Correct processing of the PV1 non-structural protein (regions P2-P3) by the CVB3 3C^{pro} *in vitro* could be demonstrated to occur in the same system (Dewalt et al., 1989, Lawson et al., 1990). The enterovirus species-block on the processing of the P1 region observed *in vitro* is surprising given the conserved nature of the protein between picornaviruses, and the ability of 3C^{pro} to cleave a wide variety of cellular proteins. Therefore it may be that the apparent species-block on processing of the P1 region is an artefact of the *in vitro* translation system used to investigate the processing of the virus in these studies. Certainly, the available evidence suggests that protease/processing incompatibility between enterovirus species is restricted to the processing of the capsid proteins. Within the literature it has been documented that exchanging the CB4 2A^{pro} gene for the PV1 2A^{pro} gene does not affect virus viability (Lu et al., 1995). It has also been reported that the CB3 3B gene can be swapped with that of PV3 3B without affecting the growth of the virus (Barclay et al., 1998) and that viable virus can be generated with CB3/PV1 2B hybrid proteins (van Kuppeveld et al., 1997c).

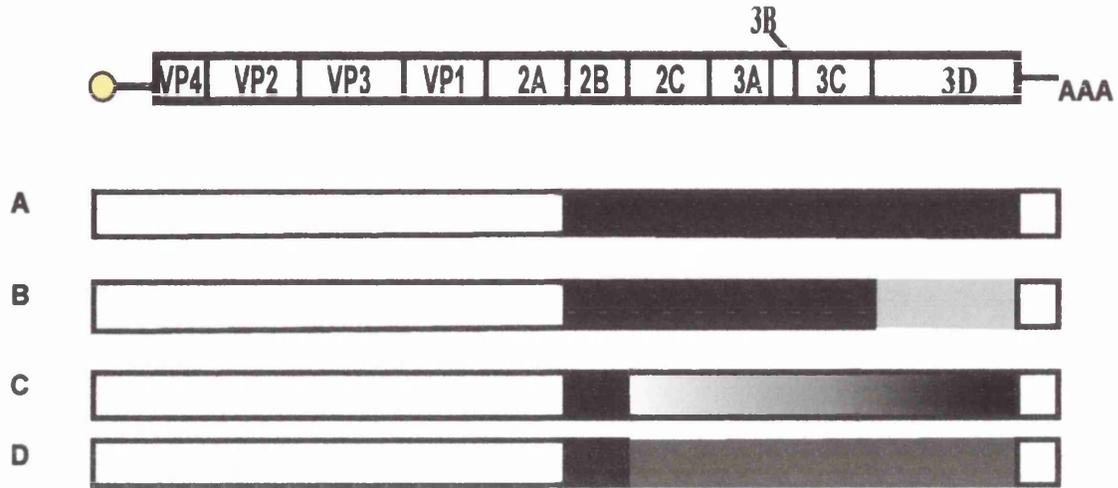


Figure 5.1 Location of recombination breakpoints of the four different classes found among Hispaniola Vaccine-derived poliovirus isolates. In the schematic of the poliovirus genome the single ORF is indicated by a single rectangle, flanked by the 5' and 3' UTR. VPg is shown as a yellow circle. In rectangles A to E Sabin-1 derived sequences are indicated by a white fill. Sequences derived from enteroviruses other than Sabin OPV strains are indicated by coloured fills. The rectangles symbolise isolates (A) D0R00-013 (B) D0R00-041C1 (C) HAI00-003 and (D)HAI01-007. Figure was taken from Kew et al, 2002

The probability of inter-species recombination occurring could be reduced further if differences between the species existed at the level of replication or packaging. Both these processes involve protein-RNA interactions. Therefore a detailed analysis of RNA may highlight differences between enterovirus species. The use of computational algorithms to assess similarities and differences between aligned related genomes provides a useful method of identifying genera and species-specific motifs that may exist within virus families. Previous computational analysis of aligned picornavirus genomes identified unique RNA structural motifs within each genus (Witwer et al., 2001). In contrast, a computational algorithm, that analyses aligned genomes for regions of SSSV, was used in this study to identify sequences that were subject to constraint and therefore of putative functional importance. The identification of a functional role for one of these regions in the replication of poliovirus was followed by its subsequent identification as an HEV-C species-specific sequence. Within poliovirus the introduction of mutations within this region caused a consistent small-plaque phenotype in HeLa cells that has since been confirmed independently by Dr.H.Harvala in RD and VERO cell-lines. Though it was not tested in this study, the expectation would be that introduction of mutations into the analogous region of any of the non-poliovirus HEV-C viruses would result in the same small-plaque phenotype. The identification of a defined replication phenotype within a standard cell line argues that the RNA secondary structure influences a fundamental part of the replication of the HEV-C viruses. Within the literature no evidence has been presented so far that suggests that the mechanics of RNA synthesis differ significantly between members of the picornavirus family at the species-level. One area where differences might exist between the HEV-B species and HEV-C species is in the compartmentalisation of the virus replication complexes within the infected cell.

Recent characterisation using FISH analysis has shown that upon uncoating the poliovirus genome migrates to a precise region of the perinuclear region. Egger and Bienz were able to demonstrate that sequential co-infection of a cell with two different serotypes of poliovirus resulted in ~80% of the replication complexes formed in the cell containing both genomes (Egger & Bienz, 2002). Variation in the localisation of the replication complex, between species, within the cell would significantly reduce the probability that any two enterovirus genomes would be found in close enough proximity to allow template-switching to occur. From the published material indirect evidence can be provided to support and refute this view.

Research using brefeldin A (BFA), an inhibitor of COP I dependent stages in the ER-Golgi transport system has established that differences exist between picornavirus genera in the cellular membranes that are utilised to form the replication complex (Gazina et al., 2002). However the data obtained has shown that the replication of both poliovirus (Cuconati et al., 1998a, Irurzun et al., 1992, Maynell et al., 1992) and the HEV-B echovirus 11 (Gazina et al., 2002) were inhibited in the presence of BFA. This suggests that the replication complexes of all enteroviruses are derived from the COP I pathway. On the basis of this observation the balance of probability argues that the genomes of all enteroviruses should localise to the same region of the perinuclear region defined by Egger and Bienz for poliovirus.

In contrast, data obtained by studies investigating HEV-C encapsidation has suggested that variation in the precise location of the replication complex amongst HEV may exist. Within picornaviruses it is known that the encapsidation of the genome is coupled to replication (Nugent et al., 1999). A consequence of this is that any difference in the cellular localisation of the replication complex between enterovirus species would be expected to influence the profile not only of recombination but also of *trans*-encapsidation. Research carried out by Barclay *et al* showed that by using an M.O.I of 1000 a 10-fold increase in the *trans*-encapsidation of PV3 by CB4 could be observed (Barclay et al., 1998). At the time Barclay *et al* discussed the possibility that the increase in M.O.I led to an increase in the local concentration of replication complexes and as a direct consequence this increased the probability of RNA and the capsid proteins from the different enterovirus replication complexes coming into contact. The observations, made by Barclay *et al* therefore provide indirect support for a difference in compartmentalisation between enterovirus species.

5.1 A membrane-targeting hypothesis

If the RNA secondary structure did function as a membrane-targeting signal three questions are raised. Firstly, what are the possible implications for this with respect to enterovirus replication? Secondly how does such a model correlate with the results that have been obtained using the subgenomic replicon and virus systems? And lastly, how could these possible factors impact on future research in this area?

A model can be put forward where the uncoated genome, via the presence of an RNA sequence or structure, is recognised and bound by a cellular transport protein (or complex).

The cellular protein would transport the viral genome, perhaps using the microtubule system, to a specific region of the ER-Golgi system thereby restricting translation to the membrane bound ribosomes, found on the rough endoplasmic reticulum. Following translation, the membrane associated non-structural proteins, would be inserted directly into the ER-Golgi, thus coupling translation to the formation of the replication complexes. Differences that are known to exist in the steps of the ER-Golgi secretory system affected by the formation of these replication complexes would arise from differences in the exact location at which the proteins are co-translationally inserted into the membrane. In other words, compartmentalism at the genera and species level might exist because the viruses have evolved to interact with cellular proteins that result in the formation of the replication complexes at distinct locations within the ER-Golgi system.

To my mind the hijacking of a cellular transport system to target the genome to the perinuclear region offers the virus two possible advantages. Firstly, the binding of a cellular protein, or the coating of the genome by a cellular transport complex could reduce the possibility of the genome being degraded by cellular RNAases. Secondly, by targeting the genome to a region of the cell where the concentrations of components required for translation and replication to occur are optimal, the virus ensures that the genome is given the best opportunity to replicate successfully.

So with this in mind what would the consequences of introducing mutations into the RNA structure or sequence be? In terms of the above model the effects of introduction of mutations into the RNA sequence or structural recognition motif could range from the subtle to the severe depending on how great the effect on the k_{on} value of the protein for the RNA genome is.

5.1.1 Possible effect of disrupting the membrane-targeting of the genome: Scenario 1

If binding of the cellular factor was completely abolished then one possible outcome is that the genome would be retained near the site of un-coating. It is known from studies of other viruses that the translation of structural proteins occurs at the perimeter of the cell e.g. the known site of translation of HIV-1 *gag* and *vpr* mRNA is at the plus end of microtubules, near the plasma membrane (Mouland et al., 2001). It is therefore possible that translation of the virus genome at the site of un-coating could occur. The implications for the

replication phenotype observed *in vitro* in this case are hard to quantify, though based on EM analysis that suggests only low-amounts of “free” ribosomes are found at the perimeter of the cell, it is probable that the efficiency of replication would be greatly reduced, as a direct consequence of translation being down-regulated.

A number of techniques have been described that would enable this hypothesis to be tested. Firstly, in an area of the cell where the concentrations of ribosomes are low it is likely that the number of polysomes traversing an RNA molecule will be reduced. The number of polysomes traversing the P3/leon/37 and P3/leon/37^{Nhel} following uncoating could be analysed using EM analysis. Secondly, under normal circumstances the poliovirus replication complexes are derived from the COP I membranes of the ER-Golgi. In the situation outlined above, where the genome was retained at the periphery of the cell the replication complexes would instead be derived from the plasma membrane. Immuno-gold labelling of markers of either the ER-Golgi or plasma membrane would enable the origin of the membrane, from which the replication complex is derived, to be determined using EM. Alternatively, the origin of the membrane could also be identified by assessing whether the replicating RNA co-localises with markers of either the ER-Golgi or plasma membrane, using confocal microscopy.

Lastly, the location of the RNA within the infected cell could be identified. The most common way to investigate the location of RNA within a cell is to use FISH (Egger & Bienz, 2002, Troxler et al., 1992b) however it has been recently reported that Ribogreen, a fluorescent dye used to measure RNA concentration, can be used to label genomic RNA inside intact HRV capsids (Kremser et al., 2004). Kremser *et al* also reported that the fluorescently labelled genome could be detected upon release from the capsid in a heat-induced manner (Kremser et al., 2004). Given the structural similarities between the capsids of HRV and poliovirus it may be possible to label the poliovirus genome using the method described by Kremser *et al*. If the poliovirus genome could be successful labelled in this manner this would provide a method of tracing the genome following uncoating of the virus capsid using live-cell microscopy. If scenario 1 is correct then the P3/leon/37^{Nhel} genome should be detected at the periphery of the cell using either Ribogreen-labelling or FISH techniques.

5.1.2 Possible effect of disrupting the membrane-targeting of the genome: Scenario 2

As an alternative, the genome may not be restricted to the periphery of the cell. In this case the genome should be translated upon “free” ribosomes within the cytoplasm. The proteins will be co-translationally processed and the proteins like 2BC and 3AB that contain hydrophobic domains will then be inserted into the nearest membrane. As a consequence of this instead of the discrete sites of replication usually observed in poliovirus-infected cells using FISH analysis numerous areas of the cell would be likely to fluoresce, as multiple distinct replication sites were created. The formation of the replication complexes in a poliovirus infection is characterised by the high localised concentrations of virus proteins. Without membrane targeting the concentration of the P2 proteins and 3AB, through the process of random insertion into the membranes, may be diluted over a wide area. One possible side effect of this is that the onset of RNA synthesis might be delayed until such a time when the required concentration of factors is reached.

5.1.3 Possible effect of disrupting the membrane-targeting of the genome: Scenario 3

The last possible scenario that can be foreseen is one in which the introduction of the mutations does not prevent the formation of the protein/RNA complex but does significantly destabilise it. If this was the case then the significant destabilisation might result in a high percentage of the complexes disassociating before reaching the ER-Golgi system. Although a reduction in virus titre would be expected the effect of such a defect on the replication phenotype of the virus would depend very much on the probability of a productive wild-type infection occurring. Though such a scenario does imply the growth defect will be more severe at lower M.O.I.

The results presented shown in figure 3.20 have shown that at the non-permissive temperature the replication phenotype is not more severe at a lower M.O.I. If the RNA sequence is identified as a membrane-target signal this would argue against scenario 3 being relevant. Interestingly, the results from the analysis of replication over time suggested that the RNA sequence affected an early stage in the replication of the virus. This would be consistent with the possible situation described in scenario 2.

5.2 Motif X as an antagonist of the anti-viral response

The presence of double-stranded RNA, within a cell induces a number of anti-viral responses, including dsRNA-dependent protein kinase R (PKR) and 2'-5' oligoadenylate synthetase, that result in the inhibition of translation, induction of apoptosis and the stimulation of cytokine and chemokine production (reviewed in Goodbourn et al., 2000). To replicate efficiently, viruses must evade stimulating these responses. Within the literature, numerous virus countermeasures have been documented. These range from the production of proteins that, through binding, inhibit transcription from promoters of interferon (human papillomavirus 16 and human herpesvirus-8; Ronco et al., 1998; Zimring et al., 1998) to the production of short RNA molecules that bind to but do not activate PKR (adenovirus and HIV-1; (reviewed in Gunnery et al., 1990, Mathews, 1995). Recently, it was identified that Epstein-Barr virus produced microRNAs, to regulate virus and cellular gene expression (Pfeffer et al., 2004). Amongst the cellular genes predicted to be targets of these microRNAs were chemokines, cytokines and components of the signal transduction pathway (Pfeffer et al., 2004). One possibility, that can not be discounted, is that RNA transcripts may be derived from the motif X region that function as antagonists of anti-viral responses.

5.3 Future work

The investigation, detailed within has led to the identification of a species-specific replication determinant. Aside from the availability of a reverse genetic system and a small-animal model of disease a number of techniques have been developed that enable aspects of RNA virus replication to be analysed. These areas include assays to analyse the effect of mutations on the catalytic functions of the polymerase, the yeast three-hybrid system, to detect virus and cellular protein-RNA interactions and microscopy techniques to assess the migration of the RNA genome to its location within the replication complex.

As a first step in identifying the function of this RNA structure the selection of revertant viruses at the non-permissive temperature must be a priority. There are four possible outcomes, each of which have implications regarding the avenue future research will take (summarised in figure 5.2). The four outcomes are as follows:

- No revertants can be selected

- Non-synonymous changes are selected
 - a) Structural proteins
 - b) Non-structural proteins

- Direct revertants are selected
 - a) Sequence
 - b) Structural

- Synonymous second site changes

5.3.1.1 No revertants are recovered

From the study of viral escape mutants it has been shown that the introduction of a single nucleotide change into a sequence occurs with a frequency of 1×10^{-3} . If two nucleotide changes are required to cause a reversion, a revertant virus would be expected to arise from a virus population with a frequency of 1×10^{-6} . If no revertants are obtained from the initial screen then one possibility that must be considered is that the number of mutations required to recover the wild-type replication phenotype is greater than 2 nucleotides. In the published research on the CRE Goodfellow *et al* reported that revertants could only be selected once the number of mutations introduced into the CRE was reduced from 8 to 3 (Goodfellow *et al.*, 2000b). If no compensating mutations are selected from blind passages of P3/Leon^{Nhe I} then it may be that to recover compensating mutations, viruses with a reduced number of mutations introduced into structure I must be constructed. It may then be possible, through blind passage of the tissue culture supernatant, to select revertants.

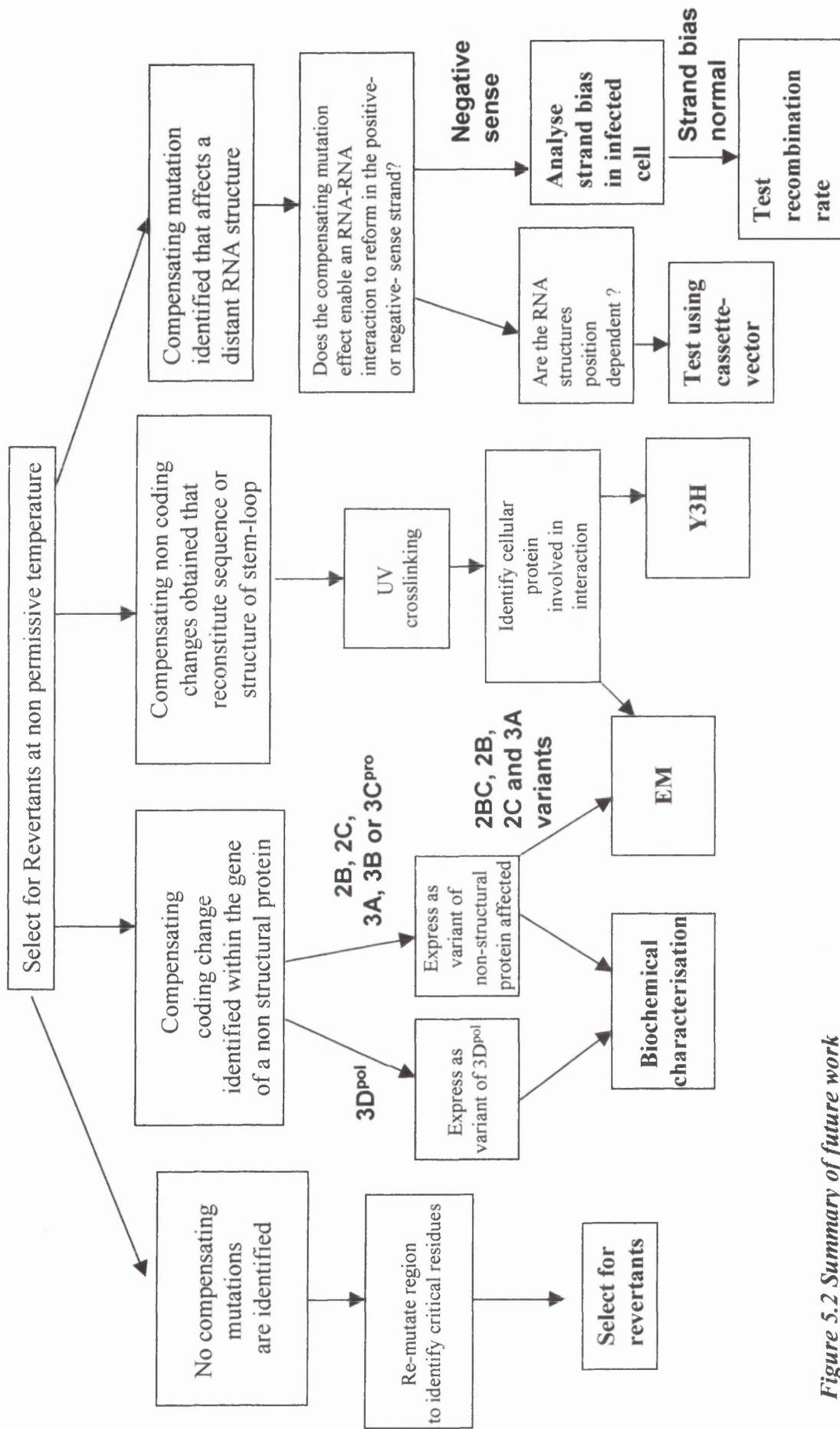


Figure 5.2 Summary of future work

5.3.1.2 Non-synonymous changes to the viral proteins are selected

Given that enterovirus non-structural proteins have not been detected within capsids, the identification of compensating coding changes that map to these same proteins would be at odds with a role for the RNA sequence as a membrane-targeting signal. However identification of coding changes that compensate for the mutated RNA sequence would provide insight into the possible function of the sequence and the best approach to be used to investigate the function of this RNA-protein interaction further. The observation that introduction of destabilising mutations into structure I generates a replication defect means it is possible that any revertants selected may have coding changes that map to the viral polymerase. In chapter 4 the expression and purification of 3D^{pol} and 3D^{N18Y} using a bacterial expression system was described. This successful purification of functional 3D^{pol} and 3D^{N18Y} will enable a full biochemical characterisation of the protein, including K_{on} , K_{off} constants and the ability of the protein to form oligomers. However, as reported in section 4.8.1, discrepancies have been reported between protein purified using this system and polymerase purified using the ubiquitin-tagged system described by Gohara *et al* (Gohara et al., 1999). The advantage of the ubiquitin-tagged system is that it enables the generation of large quantities of polymerase with a fully authentic N-terminus. The data from *in vitro* biochemical characterisation obtained using the latter should provide a more accurate picture of the activities of the 3D^{pol} *in vivo*, than that obtained using His-tagged 3D^{pol} or 3D^{pol} that has an N-formyl methionine as the first codon. Given that the biochemical differences between variants of 3D^{pol} might be quite subtle, this would argue strongly that all future research should be carried out using the ubiquitin-tagged system described by Gohara *et al* (Gohara et al., 1999).

A number of functions have been described for the final cleavage products (2B, 2C, 3A, 3B or 3C^{pro}) and the intermediate proteins (2BCP3, 3AB and 3CD^{pro} for example) of the P2 and P3 regions of the polyprotein (see introduction section 1.8.2). Distinguishing whether the coding change introduced affects the function of the final cleavage product or the intermediate protein would require biochemical analysis. Subsequent to this, the yeast three-hybrid system could be used to investigate potential virus and cellular proteins that may interact with the RNA/protein complex.

5.3.1.3 Direct revertants are selected

The identification of compensating mutations that reconstitute either the structure or sequence of structure I but which does not introduce non-synonymous changes into the non-structural proteins would be consistent with a function dependent on an interaction with a cellular protein or RNA structure. In such a scenario the compensating mutation (s) that are retrieved provide information regarding whether putative interactions are dependent on the sequence or the structure of the motif.

It will be important to determine whether the RNA binds a cellular protein. Although the yeast three-hybrid system could be used for this purpose, problems such as the sensitivity of the screen, described in chapter 3, argue against its use at this stage of the process. Instead a preliminary test, using UV-crosslinking, could be carried out to identify whether the un-mutated structure binds a cellular protein. Further confirmation that this interaction is required for efficient virus replication would be obtained if binding could be demonstrated to be abolished in RNA with a disrupted structure and recovered following introduction of the identified compensating mutations.

If an RNA-protein interaction can be demonstrated to occur then establishing the identity of the protein becomes a priority. This can be achieved through the use of matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF) mass spectrometry analysis. If the membrane-targeting hypothesis is correct, it is likely that the cellular protein identified will belong to a class of transport proteins, for example a protein like hnRNPA2. If the function of the protein involved in the interaction is involved in a cellular process other than intracellular transport, then identification of the protein involved may offer valuable insight into the advantage this interaction imparts to the virus. The nature of the interaction can be further investigated by quantifying the strength of the interaction by surface plasmon resonance (SPR) or filter binding analysis.

5.3.1.4 Synonymous second-site changes are selected

The last possible outcome from the selection of revertants would be the identification of a synonymous second-site mutations. This would imply that structure I was involved in an RNA-RNA interaction. RNA-RNA interactions have previously been reported to control

aspects of the replication of bacteriophage Q β (Klovins et al., 1998, Klovins & van Duin, 1999) and tombusviruses (Pogany et al., 2003).

The identification of an RNA-RNA interaction required for efficient replication of the HEV-C viruses would not rule out the possibility that the function of the structure formed was a recognisable cellular membrane targeting signal. It is possible that the membrane-targeting signal is non-linear in sequence. In such a scenario the cellular transport protein would recognise the tertiary structure formed as a result of the RNA-RNA interaction rather than the sequence itself. However the identification of an RNA-RNA interaction would also raise the prospect that the resultant structure was functionally required for a different aspect of HEV-C replication such as down-regulating the initiation of negative-strand synthesis or the translocation of VPg-pUpU from the CRE to the 3' end of the negative-sense strand.

The compensating mutation(s) identified provides a valuable source of information about the RNA-RNA interaction. Firstly, the position of the second site mutation within the genome provides information about whether the RNA is involved in a long or short-range interaction and secondly, the synonymous mutation may enable the determination of which strand (genome or anti-genome) the interaction occurs. Although, it is not clear whether there would be any anti-genome sequence available to participate in an RNA-RNA interaction. The detection of which strand the RNA-RNA interaction is formed has consequences for research aimed at deducing its role in the replication of the HEV-C. For example, an RNA-RNA interaction occurring in the positive-sense strand would argue against a role in recombination but would not exclude the possibility of the interaction functioning as a membrane-targeting signal. The converse is true of an RNA-RNA interaction that occurs within the anti-genome.

Independently of whether the RNA-RNA interaction occurs in the genome or anti-genome, the following question needs answered: Does the function of the interaction depend on its location within the genome? This question can be answered by analysing the mutations using an *in vitro* cassette-vector similar to that described by Goodfellow *et al* (Goodfellow et al., 2003a).

5.4 Conclusion

One of the most pressing questions to be answered by the WHO concerning the global eradication of poliomyelitis is whether the niche left vacant by poliovirus could be filled by a related enterovirus. The view, put forward from phylogenetic analysis, is that enteroviruses are replicons in search of a capsid sequence, though the evidence does suggest this is restricted to species level. Within this thesis evidence is presented of the identification of a highly conserved determinant of replication common only to viruses of HEV-C species. Characterising the exact function of the species-specific motif in terms of virus replication will provide further insight into the replication of a group of medically important viruses and may help further our understanding as to why inter-species recombinants are not observed. Taken as a whole the accumulating evidence suggests that viruses classified within the individual enterovirus clusters (HEV A-D) have evolved so far from the progenitor enterovirus and hence away from each other that recombination between species is not feasible on a number of different levels. In terms of the global poliovirus eradication program this strongly suggests that if a virus were to evolve to re-occupy the niche vacated by poliovirus, such a virus would only be likely to originate from the HEV-C species and not from one of the other enterovirus clusters.

Aside from the concerns raised regarding the poliomyelitis eradication campaign, poliovirus has, in recent years, provided a valuable model for understanding the replication of less amenable viruses, such as the human caliciviruses and HCV. The identification of a further replication determinant in the poliovirus genome implies that picornaviruses still have much to teach us in terms of understanding the interactions of positive-strand viruses with the host cell.

6 Materials and Methods

6.1 Computer programs

All SSSV analysis, described within this thesis, was carried out by Prof. P. Simmonds using the Simmonic computer package. SSSV analysis was carried out using a CLUSTAL alignment of enterovirus sequences that differed from each other by >1%. The Genbank accession numbers of all HEV-C and HEV-B sequences used in the alignments are shown in appendix 1 and 2, respectively.

Thermodynamic predictions were carried out using the MFOLD server (Zuker, 1989, Zuker, 2003) available online at <http://www.bioinfo.rpi.edu/mfoldserver>.

Sequencing was analysed by Bioedit, while plasmid maps were generated using Vector NTI. Figures were generated using Microsoft PowerPoint, Adobe Illustrator and Adobe Photoshop packages.

6.2 Solutions and Chemical suppliers

All chemicals were supplied by Sigma-Aldrich or BDH

6.2.1 Standard solutions

Phenol: chloroform

25 parts phenol: 25 parts chloroform: 1 part isomyl alcohol.

Two microspatulas of 8-hydroxyquinoline were added prior to use.

10 X TE

100 mM Tris-HCl pH 7.5

10 mM EDTA pH 8.0

10X Lithium Acetate

1 M Lithium acetate pH 7.5

Z buffer

60 mM Na₂HPO₄

40 mM NaH₂PO₄

10 mM KCl

1 mM MgSO₄·7H₂O

Cracking buffer stock solution

8 M Urea
5% (w/v) SDS
40 mM Tris-HCl pH 6.8
0.1 mM EDTA

For Complete Cracking Buffer

add 1% (v/v) β -mercaptoethanol to cracking buffer stock solution

Resuspension buffer I

50 mM Tris-HCl pH 8.0
0.5 mM EDTA
1.0 mM DTT
0.1 % (w/v) Nonidet-P40

Wash buffer I

50 mM Tris-HCl pH 8.0
0.5 mM EDTA
1.0 mM DTT

Elution buffer I-IV

50 mM Tris-HCl pH 8.0
0.5 mM EDTA
1.0 mM DTT

In addition to the above recipe elution buffers I to IV contained NaCl at final concentrations of:

- + 50 mM NaCl (**Elution buffer I**),
- + 100 mM NaCl (**Elution buffer II**),
- + 200 mM NaCl (**Elution buffer III**),
- + 500 mM NaCl (**Elution buffer IV**),

Polymerase storage buffer

50 mM Tris-HCl pH 8.0
50 mM KCl
20 % (v/v) Glycerol
2 mM DTT

6x native loading dye

30% (v/v) glycerol
0.25% (w/v) bromophenol blue
made up in TAE

1x TAE

40 mM Tris-acetate
2 mM EDTA

10x TBE (1 litre)

890 mM Tris-HCl
890 mM Boric Acid
20 mM EDTA pH 8.0

10x Tris-Glycine buffer (1 litre)

250 mM Tris-HCl

2.5 M Glycine

1% (w/v) SDS

Destain

10% (v/v) methanol,

10% (v/v) acetic acid

80% (v/v) dH₂O**Coomassie Brilliant Blue stain**

20% (v/v) methanol,

20% (v/v) glacial acetic acid,

0.2% (w/v) coomassie brilliant blue

2x HBS

50mM HEPES

280 mM NaCl

1.5mM Na₂HPO₄

pH to 7.1

10X glucose

10g/ litre

Crystal violet (1 litre)

0.5g crystal violet dissolved in 20mls ethanol,

0.9g NaCl,

100ml 40% formaldehyde

make up to 1 litre with dH₂O**6.3 Antibodies**

Antibody	Manufacturer	Type of antibody	Dilution used at
α-HA	Sigma	Mouse Monoclonal	1:5000
α-cMyc	Cell signalling technologies	Mouse Monoclonal	1:3000
α-SV5	n/a	Rabbit Polyclonal	1:1000
α-mouse HRPO conjugate	Pierce	Goat monoclonal	1:10,000
α-rabbit HRPO conjugate	Pierce	Goat monoclonal	1:10,000

Table 6.1 Antibodies used

6.4 Nucleic acid manipulations

6.4.1 Nucleotide suppliers

Unlabelled nucleotides (rNTPs and dNTPs) were supplied by Promega. Radiolabelled nucleotides [^{14}C] ATP (57 mCi/mmol), [^{35}S] methionine (1,175 Ci/mmol) and [α - ^{32}P] UTP (3000 Ci/mmol) were supplied by Amersham Pharmacia.

6.4.2 Oligonucleotides

Oligonucleotides were initially supplied by MWG-Biotech and Cruachem. Latterly oligonucleotides were supplied by TAGN. Table 6.2 shows all the oligonucleotide primers used during the course of this work.

6.4.3 Polymerase chain reaction (PCR)

PCR was used to amplify regions of DNA for cloning and to introduce specific mutations. A typical reaction contained 20 ng of template DNA, 100 pmol of each primer, 25 mM of each dNTP, 2 mM MgSO_4 , 1x pfx reaction buffer and 1x pfx enhancer and 10 units of pfx^{platinum} polymerase (Invitrogen) in a total volume of 100 μl .

Amplification of the DNA was carried out using a Techne touchgene gradient thermal cycler using the following parameters: 94°C for 1 minute, to denature the DNA; 50°C for 1 minute, to allow primers to anneal to the template DNA; 72°C for 1 minute per thousand base pairs to be amplified. The amplification was performed for 30 cycles, finally the reaction was held at 72°C for 10 minutes to ensure that the majority of final product was full-length double-stranded DNA. The annealing temperature of the reaction was varied according to the particular base composition of the primers involved.

6.4.4 Agarose gel electrophoresis

DNA was run on agarose gels containing 0.5 µg/ml ethidium bromide in 1x TAE buffer. Samples were loaded into gels using 6x native loading dye. DNA was visualised using Biorad's gel trans-illuminator system. The size of DNA fragment that was obtained was estimated using 1kb ladder (Invitrogen).

6.4.5 Restriction digestion of DNA

Enzymes were supplied by New England Biolab (NEB), Roche and Promega. Restriction digests were carried in buffers recommended by the manufacturers for use with the appropriate enzyme. Enzymes were used at a concentration of 1 unit (U) per µg of DNA in a final reaction volume of 50µl for preparative digests and 15µl for analytical digests. Restriction digests were carried out at 37° C unless otherwise specified by the manufacturer, in which case that temperature was used.

6.4.6 DNA extractions

DNA that was to be extracted was visualised on a longwave UV transilluminator, following gel electrophoresis, carried out using low melting point agarose (Roche). The DNA was excised from the agarose with a scalpel. The DNA was extracted from the agarose gel slice using Promega's Wizard extraction kit.

6.4.7 SAP treatment of DNA

DNA vector that was to be used in ligations was treated with Shrimp Alkaline phosphatase. A standard reaction contained 1x reaction buffer (Roche), 5 U of Shrimp alkaline phosphatase and 2 µg purified linearised DNA in a final volume of 50 µl. The reaction was incubated at 30° C for 30 minutes. The treated vector was purified using standard phenol-chloroform extraction and DNA precipitation techniques.

Name of oligonucleotide	Sequence 5' → 3'	Comments
O-AB-2I	TTAAAACAGCTCTGGGGT	Cloverleaf sense
O-AB-3I	GTTGCGGGGGAGGGAGTA	Cloverleaf antisense
O-AB-4I	TGGGTGTCCGTGTTGGT	IRES sense
O-AB-5I	TTCCTGGATGGCCAATC	IRES antisense
O-AB-6I	GGCCTCGTGGCGCACATG	MS2 leader sense
O-AB-7I	GGAAGTTGGATATGGGGG	MS2 terminator antisense
O-AB-8I	GAGAGATATCCTAAACAAGCAAACC	3D sequencing primer sense
O-AB-9I	GGTGTTAGAGAAAATTGGTTT	3D sequencing primer sense
IG10	GGAGTGTGCCAATCGGAAGAGC	3' end sense primer binds nucleotides 7288-7308
IG11	CCCCTACAACAGTATGAC	3'end antisense primer binds nucleotides 7392-7409
IG25	GCGCGGATCCCTAAAAATGAGTCAAGCCA	3D antisense primer
IG26	GCGATGCCATGGGGCTGGGTTGACTATG	3C sense primer
IG50	GTAATACGACTCACTATAGGGCGA	T7 sequencing primer
O-BSmut3Dpol	GGGGACGATGTTATTGCTAGCTATCCC	Mutates the base stem of the stem loop 1 identified in region 6773-7173 of PV3. Contains an <i>Nhe I</i> site. Sense primer
O-BSmut3Dpolrev	ATGGGGATAGCTAGCAATAACATCGTCCCCATAGG CAATCAT	Mutates the base stem of the stem loop 1 identified in region 6773-7173 of PV3. Contains an <i>Nhe I</i> site. Antisense primer
IG24	GCGATGCCATGGGTGAAATCCAGTGGATG	3D sense primer
2CSS	TTTAAAACGCGTCATTAATAATTACATACAG	2C LOOP sense (Goodfellow et al., 2000a)
2CSA	TTTAAAGCGCGCACTAACAAACATACAGG	2C LOOP antisense (Goodfellow et al., 2000a)
Rep3Rev	TTGAAGGCTCTCAAGGGCAT	Binds just downstream of <i>Sal I</i> site in Rep3
34-0042	ATGAGACCATCAAAGGAGGC	(Meredith et al., 1999) PV3 sense primer to amplify 3D region
IG48	GCCATTTTACCAACTGCAGCCTCACCTGGTGAG	Mutagenic primer. Binds to region in 3C
O-AC-8C	CAGCAGGAGCTCGGTGATAGTTGGTTGAAA	2C sense primer contains <i>Sac I</i> site.
O-AC-9C	CAGCAGGAGCTCGGACCACTCCAGTACAAA	3A sense primer contains <i>Sac I</i> site.
O-AC-1D	CAGCAGGAGCTCGGGCCTGGGTTTACTATGCA	3C sense primer contains <i>Sac I</i> site.

O-AC-2D	CAGCAGGAGCTCGGTGAAATCCAGTGGATG	3D sense primer contains <i>SacI</i> site.
O-AC-3D	CAGCAGGTCGACCTAAAAATGAGTCAAGCCAAC	3D antisense primer contains <i>SalI</i> site. Stop sign just before cloning site
O-AC-4D	CTGCTGCTCGAGCTATTGGAACAAAAGCCTCCAT	2C antisense primer contains <i>XhoI</i> . Allows cloning into Y2H
O-AC-5D	CTGCTGCTCGAGCTATTGCAGTTTTGCTGC	3B antisense primer contains <i>XhoI</i> site.
O-AD-1E	CGTCACCATAAGCAATCATTTTTAAGTGGTCCAGA TCTATTCCCTTGTATGTTTTTCAG	Contains <i>BglII</i> site. Antisense mutagenesis primer
O-AD-8C	GCTTACGGAATCAACCTAACC	Sense sequencing primer. Binds nt 6418-6437 PV3.
O-AD-3E	GACTATGGACTAACCTTGACTAAGACAAAATAAATT TACCACTTTTGAG	Sense mutagenesis primer. Contains <i>ApoI</i> site
O-AD-4E	CCAATCAGGAGAAGACTATGGACTGACCTTGATTG CGGCAGATAAATCTGCCGCTTTTGAGA	Sense mutagenesis primer
O-AD-4F	CCTATCCCCACGAGGTCGATGCATCACTACTAGCC CAATC	Contains <i>NsiI</i> site. Introduces mutations designed to disrupt structure II
O-AD-5F	CTAAAAATGATTGCATATGGAGATGATGTTATAG CTTCC	Contains <i>NdeI</i> site. Introduces mutations designed to disrupt structure I

Table 6.2 Oligonucleotides

6.4.8 Ligations

Ligations were carried out in a final reaction volume of 10-15 μ l. Each reaction contained 1x ligase buffer; a molar ratio of 1:3 of vector:insert and 10U of T4 DNA Ligase (Invitrogen). Ligations were left overnight at 16 $^{\circ}$ C. All ligations were cleaned up by ethanol precipitation and resuspended in 5 μ l of SDW prior to transformation.

6.4.9 Strains and genotypes of *E. coli*

Strain of <i>E. coli</i>	Genotype of <i>E. coli</i> strain	Supplier of <i>E. coli</i> Strains
ER2738	<i>F'</i> proA ⁺ B ⁺ lac ^f Δ(lacZ)M15zzf::miniTn10 (KanR) /Δ(argFlacZ)U169glnV44e 4-(McrA-) <i>rfbD1?recA1relA1?endA1spoT1?thi-1Δ(mcrCmrr)114::is10</i>	NEB
JM109	<i>F'</i> traD36proA ⁺ B ⁺ lac ^f Δ(lacZ)M15/Δ(lac-proAB) glnV44014 <i>gyrA96recA1relA1end A1thihsdR17</i>	NEB
DH5α	<i>F'</i> endA1hsdR17(r _k m _k ⁺)glnV44thi1recA1gyrA(relA1Δ(lacIZYAargF)U16 9deoR(phage80dlacΔ(lacZ/M15):	
BL21DE3(pLysS)	<i>F'</i> ompThsdS _B (r _B -m _B -)galdcn (DE3)pLysS (CamR)	

Table 6.3 Genotypes of bacterial Strains used

Lauria-sulphate Broth

10 g/litre Bacto tryptone

5 g/litre Bacto Yeast extract

5 g/litre NaCl

6.4.9.1 Preparation of Electrocompetant *E. coli*

A single colony of *E. coli* was used to inoculate 5 ml of LB media. The inoculated culture was then incubated at 37° C in an orbital shaker overnight. 2.5 ml of the overnight culture was used to inoculate 500 ml of LB media. The freshly inoculated culture was incubated at 37° C, with shaking, until the OD600 was between 0.5-0.6 at which point the bacterial culture was transferred into 2 x 500 ml centrifuge bottles that had been pre-chilled at 4° C. The cells were pelleted by centrifugation at 4,200 rpm, for 20 minutes, using a Multifuge 3R benchtop-centrifuge (Heraeus). The supernatant was removed from the cells and the pellet were washed once in 500mls ice cold dH2O (w/v) and once with 40 ml ice-cold 10% glycerol (v/v). The cells were pelleted by centrifugation for 15 minutes at 2000 rpm, using a Multifuge 3R benchtop-centrifuge (Heraeus). The pellet was resuspended in an equal volume of ice-cold 10% glycerol and aliquoted into pre-chilled cryovials. The aliquots of competent cells were frozen on dry ice and stored at -70° C.

6.4.9.2 Preparation of Calcium chloride competent cells

A single colony of *E. coli* was used to inoculate 10 ml of LB media. The inoculated culture was then incubated at 37° C in an orbital shaker overnight. The cells were pelleted at 3000 rpm using a Multifuge 3S-R benchtop centrifuge (Haereus). The supernatant was removed and the pellet was resuspended in 10 mls of ice-cold 0.1 M CaCl₂. The cells were re-pelleted at 3000 rpm using a Multifuge 3S-R benchtop centrifuge (Haereus) after 30 minutes during which time the sample was incubated on ice. Following centrifugation the pellet was resuspended in 1 ml of ice-cold 0.1 M CaCl₂.

6.4.10 Bacterial cell transformations

Two methods of transforming bacterial cells with plasmid DNA were used. For retransformations calcium chloride treatment was the preferred option; for the transformation of ligations electroporation was the preferred method.

6.4.10.1 Transformations using Calcium chloride

For calcium chloride transformations the cells were prepared as described in section 6.4.9.2. 150µl of calcium chloride competent cells were incubated on ice for 30 minutes in the presence of 2µl of DNA. The cells were heated shocked for 90 seconds at 42° C and allowed to recover for 15 minutes on ice prior to plating out on Lauria agar containing the required concentration of antibiotic.

6.4.10.2 Transformation using electroporation

Electroporation was carried out using BTX ECM 100 electroporator set at 1600v, 100Ω and 25 µF. Following transformation 1 ml of LB was added and the cells were allowed to grow at 37° C for 1 hour prior to plating out on Lauria agar containing the required antibiotic. All agar plates were inverted and incubated at 37° C for 24 hours.

6.5 Antibiotic selection

Construct derivative of plasmid	Company supplier	Antibiotic selection	Concentration of Antibiotic selection used
pIIIA/MS2-1	Clontech	Ampicillin	50 µg/ml
pT7FLC	n/a	Ampicillin	50 µg/ml
pT7FLC/Rep3	n/a	Ampicillin	50 µg/ml
pGBKT7	Clontech	Ampicillin	50 µg/ml
pACTII	Clontech	Ampicillin	50 µg/ml
pHybLex/Zeo	Invitrogen	Zeocin	25 µg/ml - <i>E.coli</i> ER2738 200 µg/ml - <i>S.cerevisiae</i> EGY48/pSH18-34
<i>pYESTrp2</i>	<i>Invitrogen</i>	<i>Kanamycin</i>	25 µg/ml

Table 6.4 Antibiotic selection used with each plasmid

6.6 Preparation of plasmid DNA

Overnight cultures of 10 or 500ml were set up in LB in the presence of antibiotic as required. Small scale preparation of plasmid DNA was carried out by using the alkaline lysis method (Sambrook et al., 1989). Large scale preparation of plasmid DNA was carried out using a Qiagen maxi-prep kit.

6.7 Sequencing

All sequencing was carried out at the Glasgow University's inhouse sequencing centre (Molecular biology sequencing unit;MBSU), using a Perkin-Elmer ABI Prism™ 377 DNA sequencer using multicolour ABI dRhodamine 'BigDye' terminators. Prior to sequencing PCR templates were purified using Qiagen's Qiaquick PCR purification kit. Sequencing was also carried out using a plasmid (cDNA) template that had been prepared using Qiagen's Midi-preparation kit.

6.8 T7 RNA transcriptions

6.8.1 Preparation of linearised DNA template

Plasmid DNA was linearised using the restriction endonuclease *Sal I*. The linearised template was extracted using phenol chloroform and precipitated using ethanol and sodium

acetate. The template was washed in 70% ethanol and airdried before being resuspended in 15 μ l dH₂O.

6.8.2 Preparation of rU15 /poly (rA)₄₀₀ primer/template (x20 stock)

40 μ M oligo U₁₅ and 3 μ M poly(rA)₄₀₀ were mixed and heated to 85° C in 50 mM HEPES pH 7 and 10 mM KCl. The solution was then allowed to cool slowly to room temperature.

6.8.3 T7 RNA transcription reactions

2 μ g of linearised template was incubated in 1 x transcription buffer, in the presence of 2.5 mM rNTPs, 50 units of RNAase OUT (NEB) ribonuclease inhibitor and 125 U of T7 RNA polymerase in a final reaction volume of 50 μ l. The reaction was incubated at 37° C for 2 hours. An aliquot was then visualised on an agarose gel electrophoresis using Biorads gel documentation system.

In order to generate enough good quality RNA to carry out a translation assay Promega's RIBOMAX™ transcription kit was used. Briefly, 5 μ g of linearised DNA was incubated for 2.5 hours at 37° C in the presence of 1x transcription buffer, 6 mM of each rNTP and 3 μ l of T7 enzyme mix (supplied with the kit). Transcribed RNA was purified using lithium chloride after being treated with 1 unit DNAase RQ1 for 30 minutes at 37° C. The RNA was washed in 70% ethanol and dried prior to resuspension in 20 μ g nuclease free H₂O.

6.9 Protein analysis

6.9.1 Preparation of dialysis tubing.

Tubing cut to the desired length was boiled for 10 minutes in 500mls of 2% (w/v) sodiumbicarbonate and 1 mM EDTA pH 8.0. The dialysis tubing was washed thoroughly 5 to 10 times following boiling in dH₂O. Following washing the tubing was boiled for a further 10 minutes in 500 ml of 1 mM EDTA pH 8.0. Prepared tubing was stored at 4° C in 500 ml 1 mM EDTA till required. The dialysis tubing was washed and allowed to equilibrate in dialysis buffer, prior to use.

6.9.2 Protein concentration determination

The concentration of protein was assayed using a Bradford protein assay kit supplied by Biorad. The dye reagent was prepared by diluting 1 part of the dye reagent concentrate with 5 parts of dH₂O. A series of dilutions of the protein standard (bovine serum albumin) were prepared. 20µl of each standard and sample solution were assayed in 200µl of diluted dye reagent. The samples were thoroughly mixed and incubated at room temperature for 5 minutes. The absorbance at 630nm was measured using a plate reader (Dynex), and the protein concentration of the samples was calculated from the standard curve.

6.9.3 SDS-PAGE electrophoresis

To ensure full denaturation samples were heated to 95° C for 5 minutes in 2X SDS-sample buffer. Full reduction of the protein samples was achieved by the addition of 0.1% β-mercaptoethanol to the 2x SDS-sample buffer (40% glycerol, 6% SDS 125mM Tris pH 6.8, 0.25% bromophenol blue in distilled water).

SDS-PAGE electrophoresis was carried out using BioRad's Protean II or III electrophoresis tanks. Acrylamide was supplied by National diagnostics. Proteins were separated by SDS-PAGE, using 5% stacking gels and 12.5% separating gels unless otherwise specified. Electrophoresis of SDS-PAGE gels was carried out in 1x Tris-glycine buffer.

6.9.4 Tris-Tricine SDS polyacrylamide gel electrophoresis

Gels contained 15% acrylamide 0.4% bis-acrylamide. The cathode buffer contained 0.1 M Tris, 0.1 M Tricine and 0.1% (w/v) SDS. The anode buffer contained 0.2 M Tris-Cl pH 8.9. Electrophoresis of Tris-tricine SDS polyacrylamide gels was carried out at 50 volts until dye had run-off the base of the gel.

6.9.5 Urea-Acrylamide gel

Polyacrylamide (19:1 acrylamide: bis-acrylamide mix) was purchased from Biorad. Urea-Acrylamide electrophoresis was carried out using BioRad's Protean II or III

electrophoresis tanks. Electrophoresis of Urea-Acrylamide gels was carried out in 0.5x TBE buffer.

6.9.6 Coomassie brilliant blue staining

Proteins separated by denaturing SDS-PAGE were visualized by staining with Coomassie brilliant blue stain for 30 minutes at room temperature with rocking. The staining solution was removed and the gel was destained. The Destain (for recipe see 6.3.1.) was changed frequently until the background was clear.

6.9.7 Western blotting

Proteins separated by SDS-PAGE were electrophoretically transferred onto Hybond™ECL™ nitrocellulose membrane (Amersham Pharmacia), using a Biorad semi-dry blot apparatus. Briefly the procedure was as follows: the gel was washed in transfer buffer (25 mM Tris, 200 mM Glycine, 20% [v/v] methanol) and was then placed on top of three layers of filter paper and nitrocellulose membrane that had been presoaked in transfer buffer. A further three presoaked layers of filter paper were placed on top carefully to ensure all air bubbles were removed. The blot was placed in the semidry blotting apparatus and an appropriate current ($0.5\text{mA}/\text{cm}^2$) was applied for 60 minutes.

The membrane was incubated in blocking solution (PBS with 5% non-fat dry milk and 0.1% Tween 20{Sigma}). After blocking the nitrocellulose membrane was incubated with the primary anti-sera at room temperature, with rocking, for 1-2 hours. Where more sensitive detection was required the nitrocellulose membrane was blocked for 1 hour and then incubated on a rocker, at 4°C, in the presence of the primary sera overnight. Following washing with PBS/0.1% Tween 20, the blot was incubated with HRPO conjugated IgG for 1 hour. The membrane was visualised after further washing using the Supersignal West chemiluminescent detection kit using the instructions provided by the manufacturer (Pierce).

The blot was exposed to X-ray film (Kodak) and developed using an automatic developer (Konica SRX- 101A).

6.9.8 Stripping Antibody from proteins immobilised on nitrocellulose membrane

The nitrocellulose membrane was incubated at 80°C for 40 minutes in 100 mM β -mercaptoethanol, 2% SDS and 62.5 mM Tris-HCl pH 6.7. The nitrocellulose membrane was then thoroughly washed in PBS prior to re-probing with primary antibody.

6.9.9 Standard VPg Uridylylation assay

Reaction mixtures contained 0.5 μ g 3CD, 0.5 μ g CRE template, and 2 μ g VPg in reaction buffer (50mM HEPES pH 7.5, 10% glycerol, 5mM magnesium acetate, 60 μ M Zinc Chloride, 10 mM β -mercaptoethanol, 10 μ M UTP and 0.004 μ M [α -³²P] UTP). The reactions were initiated by the addition of 3D^{pol} to the reaction before which all reactants were mixed and placed on ice. Reaction mixtures were incubated at 34 ° C for 30 minutes. The reaction was halted by the addition of an equal volume of loading buffer (100mM EDTA in 90% formamide containing 0.05% bromophenol blue and xylene cyanol dyes). VPg uridylylation was visualised by Tris-tricine SDS polyacrylamide gel electrophoresis.

6.9.10 Poly (rU) polymerase assay

Reactions contained 3Dpol and 2 μ M rU15 /0.15 μ M poly (rA)₄₀₀ primer/template in reaction buffer (50 mM HEPES pH 7.5, 10 mM β -mercaptoethanol, 5 mM MgCl₂, 500 μ M UTP and 0.4 μ Ci/ μ l [α -³²P] UTP). Reactions were initiated by the addition of 3D^{pol} and were carried out in a total volume of 25 μ l. The addition of a final concentration of 50 mM EDTA was used to stop the reaction after incubation at 34° C for 30 minutes. 5 μ l of each reaction was spotted on to a DE81 filter disc and allowed to air-dry completely. The dried DE81 were washed in 5 % dibasic sodium phosphate solution. Bound radioactivity was quantified by liquid scintillation counting in 3 ml of Ecoscint scintillation fluid (National Diagnostics) using a Beckman LS5000 CE scintillation counter.

6.9.11 Terminal uridyl transferase assay

Reactions contained 250 ng of RNA in reaction buffer (50 mM HEPES pH 7.5, 10 mM β -mercaptoethanol, 5 mM MgCl₂, 8 U RNAase inhibitor, 150 μ M UTP and 0.4 μ Ci/ μ l

[α - 32 P] UTP). Reactions were initiated by the addition of 3D^{pol}, diluted to the required concentration, and were carried out in a total volume of 20 μ l. The reaction was stopped by the addition of an equal volume of RNA loading buffer (8 M urea, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol), after the samples had been incubated for 1 hour at 30 °C. 20 μ l of the sample was loaded onto a 12.5% urea-acrylamide gel which had been pre-run for 30 minutes at 150 V.

6.9.12 Luciferase assay

Transfected cells were pelleted at 13,000 rpm in a benchtop centrifuge. The transfected cell pellets were washed once in PBS and resuspended in 100 μ l 1x cell culture lysis reagent (Promega). After incubating the resuspended pellets at room temperature for 10 minutes, the sample was centrifuged at 13,000rpm for 2 minutes in a Biofuge pico (Haereus) the supernatant was collected. Luciferase activity was detected from 10 μ l of the supernatant, using a TD20/20 luminometer (Promega), after the addition of 100 μ l of luciferase reagent (Promega). Prior to mixing, both supernatant and luciferase reagent were allowed to equilibrate at room temperature for 1 hour.

6.9.13 *In Vitro* translation and processing assay

Each translation reaction contained 2mM guanidine hydrochloride, 1x transcription buffer (containing amino acids-methionine) and 50 μ Ci of 35 S-labelled methionine per reaction. Rabbit reticulocyte lysate (Promega) and HeLa S10 extracts (Barton & Flanagan, 1993, Molla et al., 1991) made up 10% (v/v) and 60% (v/v) of the final reaction volume of 50 μ l respectively. The samples were incubated at 30° C for five hours. Laemmli loading buffer was added to 1x and equal volumes of the sample were loaded onto a 12.5% SDS-PAGE gel. The gel was dried using a vacuum gel dryer (Biorad) and exposed to an autorad. The image was visualised using a phosphoimager and analysed using the Personal FX computer program (BioRad).

6.10 Yeast Techniques

6.10.1 Strains of *S.cerevisiae* used

Strain of <i>S.Cerevisiae</i>	Genotype	Supplier
AH109	MATa, trp1-90, leu2-3,112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL1 _{UAS} -GAL1 _{TATA} -HIS3,GAL2 _{UAS} -GAL2 _{TATA} -ADE2,URA3::MEL1 _{UAS} -MEL1 _{TATA} -lacZ	Clontech
Y187	MATa,ura3-52,his3-200,ade2-101,trp1901,leu2-3,112,gal4Δ,met,gal80Δ,URA3:: GAL1 _{UAS} -GAL1 _{TATA} -lacZ, MEL1	Clontech
EGY48/pSH18-34	MATa ura3 trp1 his3 6 lexAop-leu2 URA3:(8LEXAop-lacZ) expressed from plasmid pSH18-34 pretransformed into EGY48 strain	Invitrogen
R40coat	MAYa, his3Δ200, trp1-901 leu2-3, 112 ade2 LYS2:::(4lexAophis3) URA3:::(8lexAop-lacZ)gal4	

Table 6.5 Genotypes of yeast strains used

6.10.2 Yeast media and solutions

Y2H and Y3H solutions were made up according to Clontech Yeast Protocols handbook.

Dropout media and supplements were supplied by Clontech initially and later by Anachem.

Yeast Media

YPD 20g/litre Difco Peptone
10g/litre Yeast Extract
add dH₂O to 950ml. Autoclave then add Glucose to 2% (filtersterilised)

YPDA as above but add adenine hemisulphate solution to final concentration of 0.003%.

Synthetically defined medium

26.7g Yeast Nitrogen Base
required amount of drop out supplement
dH₂O to 1 litre

Agar for culturing yeast on was made as above except Bactoagar was added at a concentration of 20g per litre of media prior to autoclaving.

6.10.3 Yeast transformation**6.10.3.1 Preparation of yeast cells competent for transformation**

Single colonies of *S. cerevisiae* were grown overnight, at 30° C in an orbital shaker, in 10 ml of YPD or the required selective media. 50 ml (small-scale transformation) or 300 ml (library scale transformation) of YPD were inoculated with the overnight culture to an OD₆₀₀ of 0.2-0.3. The freshly inoculated cultures were grown at 30 °C in an orbital shaker till the OD₆₀₀ was between 0.4-0.6. The yeast cells were centrifuged at 2000 rpm for 10 minutes in a Kendro multifuge 3S-R. The supernatant was removed and the pelleted yeast cells were resuspended in 50 ml of 1x TE. The yeast cells were pelleted at 2000 rpm for 10 minutes in a Kendro multifuge 3S-R and after removal of the supernatant the cells were resuspended in 1x LiAc/1x TE to a final volume of 2 ml.

6.10.3.2 Transformation of DNA

For each transformation 1µg of plasmid DNA and 100µg denatured sheared salmon sperm was incubated with 100µl of the prepared competent cells. Each sample was thoroughly mixed by inverting the tube and 600µl of 1x LiAc/40%PEG/1x TE solution was added. Following the addition of the 1x LiAc/40%PEG/1x TE solution each transformation was incubated at 30 °C for 30 minutes with shaking. Following incubation DMSO was added to 10% (v/v) and inverted to mix. The sample was heat-shocked for 10 minutes at 42 °C. The heat-shocked yeast cells were pelleted at 2000 rpm using a biofuge pico centrifuge (Haereus). The yeast cells were resuspended in 1x TE buffer prior to plating out on the required selective agar. The agar plates were inverted and incubated at 30 °C.

6.10.4 Extraction of protein samples from yeast

Single colonies of *S. cerevisiae* (containing the expression plasmid to be tested) were grown overnight, at 30° C in an orbital shaker, in 5ml of media selective for the protein coding plasmid (Leucine deficient media). The entire overnight culture was used as inoculum in a final volume of 50 ml of selective media as the expression culture. The 50 ml culture was grown at 30°c in an orbital shaker until the OD₆₀₀ of the culture was between 0.4-0.8.

To pellet the yeast cells, the entire 50 ml culture was centrifuged at 2000 rpm for 10 minutes in a Kendro multifuge 3S-R. The supernatant was removed and the pelleted yeast cells were washed once in 50 ml of dH₂O. Following washing the pelleted cells were frozen at -20° C. The frozen cells were thawed in complete cracking buffer that had been prewarmed to 60° C. Complete cracking buffer was used at a volume of 100µl per 7.5 total OD₆₀₀ units:

$$\text{Total OD}_{600} \text{ units} = \text{OD}_{600} \text{ units in 1 ml} \times \text{volume of yeast culture}$$

The resuspended cells were transferred to a 1.5ml screw-cap vial containing approximately 100-150 µl glass beads. The sample was then vortexed vigorously for 1 minute. Afterwards the sample was boiled at 100 °C for 10 minutes. The sample was centrifuged briefly in a Beckman biofuge pico to pellet the glass beads. The supernatant was transferred to a fresh 1.5ml vial. The sample was then loaded onto an SDS-PAGE electrophoresis gel. The remainder of the sample was stored at -20 °C.

6.10.5 β-Galactosidase activity

To confirm that the interaction between the protein and RNA was a true positive the yeast were assayed for β-galactosidase activity. A quantitative and a qualitative method were used:

6.10.5.1 β-galactosidase colony filter lifts.

After 3 days of growth on the mating plates (media lacking leucine and uracil), filters (Whatman 5) were lifted from the plates. The yeast on the filters were lysed using repeated

freeze-thawing before being incubated at 30° C in the presence of X-gal (10 mg/ml) in Z buffer. Where the protein interacts with the RNA plasmid the colony will be positive (blue colour).

6.10.5.2 Liquid O-nitrophenyl-β-D-galactopyranoside (ONPG) assay

For verification of the relative affinity of the RNA-protein interaction a liquid ONPG assay was used to determine β-galactosidase activity. Single colonies of the strains to be tested, were grown overnight at 30°c in an orbital shaker, in 5ml of media selective for both the protein coding plasmid (Leucine+ gene) and the RNA containing plasmid (uracil+ gene). 2mls of the overnight culture was used to inoculate 8mls of selective media and was grown with shaking to an OD₆₀₀ of between 0.2-0.3. The cells were grown at 30° C in an orbital shaker to an OD₆₀₀ of between 0.4-0.8. To enable calculation of the β-galactosidase units produced the exact OD₆₀₀, when the cells harvested, was recorded. For each colony of interest the assay for β-galactosidase activity was carried out in triplicate. 1.5 ml of the culture, per sample, was transferred to a microcentrifuge tube. The yeast cells were pelleted at 13,000 rpm for 30 secs using a Biofuge pico centrifuge (Haereus). The pelleted cells were washed once in 1 ml of Z buffer. Following the single wash step the cells were resuspended in 300µl of Z buffer. The yeast samples were lysed using repeated freeze-thawing. 100µl of sample was incubated at 30° C in the presence of 0.6 mg ONPG and 700 µl Z buffer/β-mercaptoethanol (370:1 v/v ratio) Reactions were stopped by the addition of Na₂CO₃ to a concentration of 0.5 mM. The cellular debris was removed by centrifugation, and the supernatant was transferred to a cuvette. The optical density of the supernatant was measured at OD₄₂₀. The relative strength of the interactions was determined by calculating, using the Miller equation (Miller, 1972), the β-galactosidase units produced.

The Miller equation

$$\text{where } \frac{1000 \times \text{OD}_{420}}{T \times V \times \text{OD}_{600}}$$

T = Incubation time in minutes
V = 0.1 x Concentration factor

The OD₆₀₀ was determined from 1 ml of culture. If 1.5 ml of culture is pelleted and resuspended in 0.3 ml Z buffer, this corresponds to a concentration factor of 5.

6.11 Mammalian cell techniques

6.11.1 Cell culture

HeLa T4 cells were cultured in Dulbeccos modified eagles medium (DMEM) containing 10% FCS and L-glutamine and incubated at 37° C /5% CO₂.

6.11.2 Cell techniques

6.11.2.1 Transfection using electroporation

Cells were trypsinised and washed in PBS. Cells were resuspended in PBS to give 4x 10⁶ cells per 500 µl transfection. Electroporation cuvettes (0.4cm) were chilled prior to the addition of the RNA/cell mix. Cells were electroporated at 250V, 250µF and 400Ω. Following electroporation cells were plated out and incubated at 37° C/5%CO₂ in the presence of DMEM/10% FCS.

6.11.2.2 Transfection using DEAE–dextran

RNA to be transfected was made up to 100µl in transfection buffer (1x HBS, 1x glucose 50µg DEAE-dextran) and pre-incubated on ice for 30 minutes. Cells to be used in the transfection were prepared as previously described. Cells were incubated for 30 minutes at room temperature, following addition of the transfection mix, with occasional shaking to ensure complete coverage of the cells.

6.12 Virological techniques

6.12.1 Passage of virus

Virus was used to infect cells that were 80% confluent at an M.O.I. of approximately 10 overnight in serum-free DMEM. The supernatant was freeze thawed three times and the cellular debris was removed by centrifugation at 1500 rpm in a Multifuge 3S-R (Haereus). The virus supernatant was aliquoted and frozen at -20° C.

6.12.2 Plaque assays

Cells were prepared to 80% confluency in DMEM/10% FCS (v/v) and L-glutamine in 6 well dishes. Serum was removed by washing in PBS. Virus was incubated for 30 minutes at room temperature, in the absence of media, with occasional shaking. Overlay media was used in a 67:30 ratio to 2 % bactoagar. 4mls of overlay agar was added to each well and incubated at room temperature until the agar was set. Plates were subsequently inverted and incubated at 37° C/ 5% CO₂ for 2-3 days and plaques were visualised following staining with crystal violet.

7 Bibliography

- Agirre, A., Barco, A., Carrasco, L. & Nieva, J. L. (2002). Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *Journal of Biological Chemistry* **277**, 40434-40441.
- Ahlquist, P. (2002). RNA-dependent RNA polymerases, viruses and RNA silencing. *Science* **296**, 1270-1273.
- Ainger, K., Avossa, D., Diana, A. S., Barry, C., Barbarese, E. & Carson, J. H. (1997). Transport and Localization elements in Myelin Basic Protein mRNA. *The journal of Cell Biology* **138**, 1077-1087.
- Ainger, K., Avossa, D., Morgan, F., Hill, S. J., Barry, C., Barbarese, E. & Carson, J. H. (1993). Transport and localisation of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. *Journal of Cell Biology* **123**, 431-441.
- Aldabe, R. & Carrasco, L. (1995). Induction of membrane proliferation by poliovirus proteins 2C and 2BC. *Biochem Biophys Res Commun* **206**, 64-76.
- Alexander, L., Lu, H. H. & Wimmer, E. (1994). Polioviruses containing picornavirus type-1 and/or type-2 internal ribosomal entry site elements - genetic hybrids and the expression of a foreign gene. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1406-1410.
- Aminev, A. G., Amineva, S. P. & Palmenberg, A. C. (2003). Encephalomyocarditis viral protein 2A localises to nucleoli and inhibits cap-dependent mRNA translation. *Virus research* **95**, 45-57.
- Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. (1993). Poliovirus RNA Synthesis Utilizes an RNP Complex Formed Around the 5'-End of Viral RNA. *EMBO Journal* **12**, 3587-3598.
- Andino, R., Rieckhof, G. E. & Baltimore, D. (1990a). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* **63**, 369-380.
- Andino, R., Rieckhof, G. E. & Baltimore, D. (1990b). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* **63**, 369-380.
- Andino, R., Rieckhof, G. E., Trono, D. & Baltimore, D. (1990c). Substitutions in the protease (3C^{pro}) gene of poliovirus can suppress a mutation in the 5' non-coding region. *Journal of Virology* **64**, 607-612.
- Antao, V. P., Lai, S. Y. & Tinoco, I., Jr. (1991). A Thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic acids research* **19**, 5901-5905.
- Antao, V. P. & Tinoco, I., Jr. (1992). Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic acids research* **20**, 819- 824.

- Arnold, J. J. & Cameron, C. E. (1999). Poliovirus RNA-dependent RNA polymerase (3Dpol) is sufficient for template switching in vitro. *J Biol Chem* **274**, 2706-16.
- Arnold, J. J. & Cameron, C. E. (2000). Poliovirus RNA-dependent RNA polymerase (3D(pol)). Assembly of stable, elongation-competent complexes by using a symmetrical primer-template substrate (sym/sub). *J Biol Chem* **275**, 5329-36.
- Arnold, J. J., Ghosh, S. K. & Cameron, C. E. (1999). Poliovirus RNA-dependent RNA polymerase (3D(pol)). Divalent cation modulation of primer, template, and nucleotide selection. *J Biol Chem* **274**, 37060-9.
- Bacharach, E. & Goff, S. P. (1998). Binding of the human immunodeficiency virus type 1 Gag protein to the viral RNA encapsidation signal in the yeast three-hybrid system. *Journal of Virology* **72**, 6944-9.
- Back, S. H., Kim, Y. K., Kim, W. J., Cho, S., Oh, H. R., Kim, J. E. & Jang, S. K. (2002). Translation of polioviral mRNA is inhibited by cleavage of polypyrimidine tract-binding proteins executed by polioviral 3C(pro). *J Virol* **76**, 2529-42.
- Badorff, C., Berkely, N., Mehrotra, S., Talhouk, J. W., Rhoads, R. E. & Knowlton, K. U. (2000). Enteroviral protease 2A directly cleaves dystrophin and is inhibited by a dystrophin-based substrate analogue. *Journal of Biological Chemistry* **275**, 11191-7.
- Baltimore, D., Franklin, R., Eggers, H. J. & Tamm, I. (1963). Poliovirus-induced RNA polymerase and the effects of virus-specific inhibitors on its production. *Proceedings of the National Academy USA* **49**, 843-849.
- Banerjee, R. & Dasgupta, A. (2001). Interaction of picornavirus 2C polypeptide with the viral negative-strand RNA. *J Gen Virol* **82**, 2621-7.
- Banerjee, R., Tsai, W., Kim, W. & Dasgupta, A. (2001). Interaction of poliovirus-encoded 2C/2BC polypeptides with the 3' terminus negative-strand cloverleaf requires an intact stem-loop b. *Virology* **280**, 41-51.
- Barclay, W., Li, Q., Hutchinson, G., Moon, D., Richardson, A., Percy, N., Almond, J. & Evans, D. (1998). Encapsidation studies of poliovirus subgenomic replicons. *Journal of General Virology* **79**, 1725-1734.
- Barco, A. & Carrasco, L. (1995). Cloning and inducible synthesis of poliovirus non-structural proteins in *Saccharomyces cerevisiae*. *Gene* **156**, 19-25.
- Barco, A., Ventoso, I. & Carrasco, L. (1997). The yeast *Saccharomyces cerevisiae* as a genetic system for obtaining variants of poliovirus protease 2A. *J Biol Chem* **272**, 12683-91.
- Barton, D. J. & Flanagan, J. B. (1993). Coupled translation and replication of poliovirus RNA in vitro: synthesis of functional 3D polymerase and infectious virus. *J Virol* **67**, 822-31.

- Barton, D. J., Morasco, B. J. & Flanagan, J. B. (1999). Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *Journal of Virology* **73**, 10104-12.
- Beales, L., Holzenburg, A. & Rowlands, D. J. (2003). Viral internal ribosome entry site structures segregate into two distinct morphologies. *Journal of Virology* **77**, 6574-6579.
- Belnap, D. M., Filman, D. J., Trus, B. L., Cheng, N., Booy, F. P., Conway, J. F., Curry, S., Hiremath, C. N., Tsang, S. K., Steven, A. C. & Hogle, J. M. (2000). Molecular tectonic model of virus structural transitions: the putative cell entry states of poliovirus. *Journal of Virology* **74**, 1342-54.
- Belsham, G. J. & Sonenberg, N. (2000). Picornavirus RNA translation: roles for cellular proteins. *Trends in Microbiology* **8**, 330-335.
- Ben-Basset, A., Bauer, K., Chang, S. Y., Myambo, K., Boosman, A. & Chang, S. (1987). Processing of the initiation methionine from proteins: Properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *Journal of Bacteriology* **169**, 751-757.
- Bergelson, J. M., Mohanty, J. G., Crowell, R. L., St. John, N. F., Lublin, D. M. & Finberg, R. W. (1995). Coxsackievirus B3 adapted to growth in RD cells binds to decay-accelerating factor (CD55). *J Virol* **69**, 1903-6.
- Bernstein, D. S., Buter, N., Stumpf, C. & Wickens, M. (2002). Analyzing mRNA-protein complexes using a yeast-three-hybrid system. *Methods* **26**, 123-141.
- Bieniasz, P. D., Grdina, T. A., Bogard, H. P. & Cullen, B. R. (1999). Analysis of the Effect of Natural Sequence Variation in Tat and in Cyclin T on the formation and RNA Binding Properties of Tat-Cyclin T complexes. *Journal of Virology* **73**, 5777-5786.
- Bieniasz, P. D., Grdina, T. A., Bogerd, H. P. & Cullen, B. R. (1998). Recruitment of a protein complex containing Tat and cyclin T1 to TAR governs the species specificity of HIV-1 Tat. *EMBO* **17**, 7056-7065.
- Bienz, K., Egger, D. & Pasamontes, L. (1987). Association of polioviral 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**, 220-6.
- Bienz, K., Egger, D. & Pfister, T. (1994). Characteristics of the poliovirus replication complex. *Arch Virol Suppl* **9**, 147-57.
- Bienz, K., Egger, D., Pfister, T. & Troxler, M. (1992). Structural and Functional Characterization of the Poliovirus Replication Complex. *Journal of Virology* **66**, 2740-2747.
- Bienz, K., Egger, D., Troxler, M. & Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *Journal of Virology* **64**, 1156-1163.

- Blomqvist, S., Bruu, A.-L., Stenvik, M. & Hovi, T. (2003). Characterization of a recombinant type 3/type 2 poliovirus isolated from a healthy vaccinee and containing a chimeric capsid protein VP1. *J Gen Virol* **84**, 573-580.
- Blyn, L. B., Towner, J. S., Semler, B. L. & Ehrenfeld, E. (1997). Requirement of Poly(rC) binding protein 2 for translation of poliovirus RNA. *Journal of Virology* **71**, 6243-6246.
- Bouffard, P., Barbar, E., Briere, F. & Boire, G. (2000). Interaction, cloning and characterization of RoBPI, a novel protein binding to human Ro ribonucleoprotein. *RNA* **6**, 66-78.
- Brent, R. & Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729-36.
- Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R. & Rey, F. A. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Science* **96**, 13034-13039.
- Brown, B., Oberste, M. S., Maher, K. & Pallansch, M. (2003). Complete genomic sequencing shows that poliovirus and members of human enterovirus species C are closely related in the noncapsid coding region. *Journal of Virology* **77**, 8973-8984.
- Cassiday, L. & Maher III, J. (2001). In Vivo Recognition of an RNA Aptamer by Its Transcription factor Target. *Biochemistry* **40**, 2433-2438.
- Chapman, N. M., Ragland, A., Leser, J. S., Hofling, K., Willian, S., Semler, B. L. & Tracy, S. (2000). A group B coxsackievirus/poliovirus 5' nontranslated region chimera can act as an attenuated vaccine strain in mice. *J Virol* **74**, 4047-56.
- Cheetham, G. M. & Steitz, T. A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science* **286**, 2305-2309.
- Cho, M. W., Teterina, N., Egger, D., Bienz, K. & Ehrenfeld, E. (1994). Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. *Virology* **202**, 129-145.
- Chow, M., Newman, J. F., Filman, D., Hogle, J. M., Rowlands, D. J. & Brown, F. (1987). Myristylation of picornavirus capsid protein VP4 and its structural significance. *Nature* **327**, 482-486.
- Clark, M. E., Lieberman, P. M., Berk, A. J. & Dasgupta, A. (1993). Direct cleavage of human TATA-binding protein by poliovirus protease 3C in vivo and in vitro. *Mol Cell Biol* **13**, 1232-7.
- Cole, C. N. & Baltimore, D. (1973). Defective interfering particles of poliovirus: IV. Mechanisms of enrichment. *Journal of Virology* **12 No. 6**, 1414-1426.

- Crary, S., Towner, J., Honig, J., Shoemaker, T. & Nichol, S. (2003). Analysis of the role of predicted RNA secondary structures in Ebola virus replication. *Virology* **306**, 210-218.
- Crotty, S., Gohara, D. W., Gilligan, D. K., Karelsky, S., Cameron, C. & Andino, R. (2003). Manganese-Dependent Poliovirus Caused by Mutations Within the Viral Polymerase. *Journal of Virology* **77**, 5378-5388.
- Cuconati, A., Molla, A. & Wimmer, E. (1998a). Brefeldin A inhibits cell-free, de novo synthesis of poliovirus. *J Virol* **72**, 6456-64.
- Cuconati, A., Xiang, W., Lahser, F., Pfister, T. & Wimmer, E. (1998b). A protein linkage map of the P2 nonstructural proteins of poliovirus. *J Virol* **72**, 1297-307.
- Cui, T. & Porter, A. G. (1995). Localization of binding site for encephalomyocarditis virus RNA polymerase in the 3'-noncoding region of the viral RNA. *Nucleic Acids Res* **23**, 377-82.
- Culley, A. I., Lang, A. S. & Suttle, C. A. (2003). High diversity of unknown picorna-like viruses in the sea. *Nature* **424**, 1054-1057.
- Daneholt, B. (1999). Pre-mRNP particles: from gene to nuclear pore. *Current Biology* **9**, R412-R415.
- Danthi, P., Toteson, M., Li, Q. & Chow, M. (2003). Genome Delivery and Ion Channel properties are altered in VP4 Mutants of Poliovirus. *Journal of Virology* **77**, 5266-5274.
- Datta, U. & Dasgupta, A. (1994). Expression and subcellular localization of poliovirus VPg-precursor protein 3AB in eukaryotic cells: evidence for glycosylation in vitro. *J Virol* **68**, 4468-77.
- Davies, M. V., Pelletier, J., Meerovitch, K., Sonenberg, N. & Kaufman, R. J. (1991). The Effect of Poliovirus Proteinase 2Apro Expression on Cellular Metabolism - Inhibition of DNA Replication, RNA Polymerase-II Transcription, and Translation. *Journal of Biological Chemistry* **266**, 14714-14720.
- de Jong, A. S., Schrama, I. W., Willems, P. H., Galama, J. M., Melchers, W. J. & van Kuppeveld, F. J. (2002). Multimerization reactions of coxsackievirus proteins 2B, 2C and 2BC: a mammalian two-hybrid analysis. *J Gen Virol* **83**, 783-93.
- Deitz, S., Dodd, D., Cooper, S., Parham, P. & Kirkegaard, K. (2000). MHC-I dependent antigen presentation is inhibited by poliovirus protein 3A. *Proceedings of the National Academy of Science USA* **97**, 13790-13795.
- DeStefano, J. J., Buiser, R. G., Mallaber, L. M., Fay, P. J. & Bambara, R. A. (1992). Parameters that influence processive synthesis and site-specific termination by human immunodeficiency virus reverse transcriptase on RNA and DNA templates. *Biochim. Biophys. Acta* **1131**, 270-280.

- Dewalt, P. G., Lawson, M. A., Colonno, R. J. & Semler, B. L. (1989). Chimeric picornavirus polyproteins demonstrate a common 3C proteinase substrate specificity. *Journal of Virology* **63** No. **8**, 3444-3452.
- Diamond, S. E. & Kirkegaard, K. (1994). Clustered charged-to-alanine mutagenesis of poliovirus RNA-dependent RNA polymerase yields multiple temperature-sensitive mutants defective in RNA synthesis. *J Virol* **68**, 863-76.
- Dildine, S. L. & Semler, B. L. (1989). The deletion of 41 proximal nucleotides reverts a poliovirus mutant containing a temperature-sensitive lesion in the 5' noncoding region of genomic RNA. *Journal of Virology* **63**, 847-862.
- Dildine, S. L., Stark, K. R., Haller, A. A. & Semler, B. L. (1991). Poliovirus Translation Initiation - Differential Effects of Directed and Selected Mutations in the 5' Noncoding Region of Viral RNAs. *Virology* **182**, 742-752.
- Dodd, D., Giddings Jr, T. & Kirkegaard, K. (2001). Poliovirus 3A protein Limits Interleukin-6(IL-6),IL-8 and beta interferon secretion during viral infection. *Journal of Virology* **75**, 8158-8165.
- Doedens, J., Giddings Jr, T. & Kirkegaard, K. (1997). Inhibition of ER-golgi traffic by poliovirus protein 3A: genetic and ultrastructural analysis. *Journal of virology* **71**, 9054-9064.
- Doedens, J. R. & Kirkegaard, K. (1995). Inhibition of Cellular Protein Secretion By Poliovirus Proteins 2B and 3A. *EMBO Journal* **14**, 894-907.
- Donnelly, M. L. L., Hughes, L. E., Luke, G., Mendoza, H., ten Dam, E., Gani, D. & Ryan, M. (2001a). The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *Journal of General Virology* **82**, 1027-1041.
- Donnelly, M. L. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D. & Ryan, M. (2001b). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *Journal of General Virology* **82**, 1013-1025.
- Duggal, R., Cuconati, A., Gromeier, M. & Wimmer, E. (1997). Genetic recombination of poliovirus in a cell-free system. *Proceedings of the National Academy of Science* **94**, 13786-13791.
- Edwards, T. A., Pyle, S. E., Wharton, R. P. & Aggarwal, A. K. (2001). Structure of Pumilo reveals similarities between RNA and peptide binding motifs. *Cell* **105**, 281-289.
- Egger, D. & Bienz, K. (2002). Recombination of Poliovirus RNA Proceeds in Mixed Replication Complexes Originating from Distinct Replication Start Sites. *J. Virol.* **76**, 10960-10971.
- Egger, D., Pasamontes, L., Bolten, R., Boyko, V. & Bienz, K. (1996). Reversible Dissociation of the Poliovirus Replication Complex - Functions and Interactions of Its Components in Viral-Rna Synthesis. *Journal of Virology* **70**, 8675-8683.

- Egger, D., Teterina, N., Ehrenfeld, E. & Bienz, K. (2000). Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. *Journal of Virology* **74**, 6570-80.
- Ellington, A. D. & Szostak, J. W. (1990). In vitro selection of RNA molecules that specifically bind specific ligands. *Nature* **346**, 818-822.
- Estojak, J., Brent, R. & Golemis, E. A. (1995). Correlation of two-hybrid affinity data with in vitro measurements. *Mol Cell Biol* **15**, 5820-9.
- Evans, D. & Almond, J. (1998). Cell receptors for picornaviruses as determinants of cell tropism and pathogenesis. *Trends in Microbiology* **6**, 198-202.
- Evans, D. M. A., Dunn, G., Minor, P. D., Schild, G. C., Cann, A. J., Stanway, G., Almond, J. W., Currey, K. & Maizel, J. V. (1985). Increased neurovirulence associated with a single nucleotide change in a noncoding region of the sabin type-3 poliovaccine genome. *Nature* **314**, 548-550.
- Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-6.
- Fosmire, J. A., Hwang, K. & Makino, S. (1992). Identification and Characterization of a Coronavirus Packaging Signal. *Journal of Virology* **66**, 3522-3530.
- Freistadt, M. S. & Eberle, K. E. (1996). Correlation between poliovirus type 1 Mahoney replication in blood cells and neurovirulence. *Journal of Virology* **70**, 6486-6492.
- Fujiwara, T., Oda, S., Yokota, A. & Ikehara, Y. (1988). Brefeldin A causes disassembly of the Golgi complex and accumulation of secretory proteins in the endoplasmic reticulum. *Journal of Biological Chemistry* **263**, 18545-18552.
- Gamarnik, A. V. & Andino, R. (1996). Replication of poliovirus in *Xenopus* oocytes requires two human factors. *Embo J* **15**, 5988-98.
- Gamarnik, A. V. & Andino, R. (1998). Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes and Development* **12**, 2293-2304.
- Gao, G., Olova, M., Georgiadis, M., Hendrickson, W. A. & Goff, S. P. (1997). Conferring RNA polymerase activity to a DNA polymerase: A single residue in reverse transcriptase controls substrate selection. *Proceedings of the National Academy of Science USA* **94**, 407-411.
- Gazina, E. V., Mackenzie, J. M., Gorrell, R. J. & Anderson, D. A. (2002). Differential requirements for COPI coats in formation of replication complexes among three genera of Picornaviridae. *J Virol* **76**, 11113-22.
- Gerber, K., Wimmer, E. & Paul, A. V. (2001). Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a cis-replicating element in the coding sequence of 2A(pro). *J Virol* **75**, 10979-90.

- Gingras, A. C., Raught, B. & Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annual Review of Biochemistry* **68**, 913-963.
- Girard, M. & Baltimore, D. (1967). The poliovirus replication complex: Site for synthesis of poliovirus RNA. *Journal of Molecular Biology* **24**, 59-74.
- Gmyl, A. P., Belousov, E. V., Maslova, S. V., Khitrina, E. V., Chetverin, A. B. & Agol, V. I. (1999). Nonreplicative RNA recombination in poliovirus. *J Virol* **73**, 8958-65.
- Gohara, D. W., Ha, C. S., Ghosh, S. K. B., Arnold, J. J., Wisniewski, T. J. & Cameron, C. (1999). Production of "Authentic" Poliovirus RNA-dependent RNA polymerase (3Dpol) by Ubiquitin-Protease-Mediated Cleavage in *Escherichia coli*. *Protein Expression and Purification* **17**, 128-138.
- Goodbourn, S., Didcock, L. & Randell, R. E. (2000). Interferons: cell signalling, immune modulation, antiviral responses and virus countermeasures. *Journal of General Virology* **81**, 2341-2364.
- Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W. & Evans, D. J. (2000a). Identification of a cis-acting replication element within the poliovirus coding region. *J Virol* **74**, 4590-600.
- Goodfellow, I., Kerrigan, D. & Evans, D. J. (2003a). Structure and Function analysis of the poliovirus cis-acting replication element (CRE). *RNA* **9**, 124-137.
- Goodfellow, I. G., Chaudhry, Y., Richardson, A., Meredith, J. M., Almond, J. W., Barclay, W. S. & Evans, D. J. (2000b). Identification of a cis-acting replication element (CRE) within the poliovirus coding region. *Journal of Virology* **74**, 4590-4600.
- Goodfellow, I. G., Polacek, C., Andino, R. & Evans, D. J. (2003b). The poliovirus 2C cis-acting replication element-mediated uridylylation of VPg is not required for synthesis of negative-sense genomes. *Journal of General Virology* **84**, 2359-2363.
- Gorbalenya, A. E., Koonin, E. V. & Wolf, Y. I. (1990). A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. *FEBS Letters* **262**, 145-148.
- Gorbalenya, A. E., Pringle, F. M., Zeddari, J. L., Luke, B. T., Cameron, C., Kalkmakoff, J., Hanzlik, Gordon, K. H. J. & Ward, V. K. (2002). The Palm Subdomain-based Active Site is Internally Permuted in Viral RNA-dependent RNA polymerases of an Ancient Lineage. *Journal of Molecular Biology* **324**, 47-62.
- Gradi, A., Svitkin, Y. V., Imataka, H. & Sonenberg, N. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc Natl Acad Sci U S A* **95**, 11089-94.
- Graff, J., Cha, J., Blyn, L. B. & Ehrenfeld, E. (1998). Interaction of poly(rC) binding protein 2 with the 5' noncoding region of hepatitis A virus RNA and its effects on translation. *J Virol* **72**, 9668-75.

- Gromeier, M., Alexander, L. & Wimmer, E. (1996). Internal ribosomal entry site substitution eliminates neurovirulence in intergeneric poliovirus recombinants. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2370-2375.
- Gromeier, M., Bossert, B., Arita, M., Nomoto, A. & Wimmer, E. (1999). Dual stem loops within the poliovirus internal ribosomal entry site control neurovirulence. *J Virol* **73**, 958-64.
- Gromeier, M. & Wimmer, E. (1998). Mechanisms of injury-provoked poliomyelitis. *Journal of Virology* **72**, 5056-5060.
- Gunnery, S., Rice, A. P., Robertson, H. D. & Mathews, M. B. (1990). Tat-responsive region RNA of human immunodeficiency virus 1 can prevent activation of the double-stranded RNA-activated protein kinase. *Proceedings of the National Academy of Sciences* **87**, 8687-8691.
- Gustin, K. E. (2003). Inhibition of nucleo-cytoplasmic trafficking by RNA viruses: targeting the nuclear pore complex. *Virus Research* **95**, 35-44.
- Gustin, K. E. & Sarnow, P. (2001). Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. *EMBO* **20**, 240-249.
- Gutierrez, A. L., Denova, O., Campo, M., Racaniello, V. R. & delAngel, R. M. (1997). Attenuating mutations in the poliovirus 5' untranslated region alter its interaction with polypyrimidine tract-binding protein. *Journal of Virology* **71**, 3826-3833.
- Hagino Yamagishi, K. & Nomoto, A. (1989). In vitro construction of poliovirus defective interfering particles. *Journal of Virology* **63**, 5386-5392.
- Haller, A. A. & Semler, B. L. (1992). Linker Scanning Mutagenesis of the Internal Ribosome Entry Site of Poliovirus RNA. *Journal of Virology* **66**, 5075-5086.
- Hamada, S., Ishiyama, K., Sakulsingharoj, C., Choi, S., Wu, Y., Wang, C., Singh, S., Kawai, N., Messing, J. & Okita, T. W. (2003). Dual regulated RNA transport pathways to the cortical region in developing rice endosperm. *The plant cell* **15**, 2265-2272.
- Hammerle, T., Hellen, C. U. & Wimmer, E. (1991). Site-directed mutagenesis of the putative catalytic triad of poliovirus 3C proteinase. *Journal of Biological Chemistry* **266**, 5412-5416.
- Hanecak, R., Semler, B. L., Anderson, C. W. & Wimmer, E. (1982). Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 3973-3977.
- Hansen, J. L., Long, A. M. & Schultz, S. C. (1997). Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* **5**, 1109-22.

- Harris, K. S., Xiang, W., Alexander, L., Lane, W. S., Paul, A. V. & Wimmer, E. (1994). Interaction of poliovirus polypeptide 3CDpro with the 5' and 3' termini of the poliovirus genome. Identification of viral and cellular cofactors needed for efficient binding. *Journal of Biological Chemistry* **269**, 27004-27014.
- Hellen, C. U. T., Witherell, G. W., Schmid, M., Shin, S. H., Pestova, T. V., Gil, A. & Wimmer, E. (1993). A cytoplasmic 57-kda protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 7642-7646.
- Herold, J. & Andino, R. (2001). Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol Cell* **7**, 581-91.
- Hobson, S. D., Rosenblum, E. S., Richards, O. C., Richmond, K., Kirkegaard, K. & Schultz, S. C. (2001). Oligomeric structures of poliovirus polymerase are important for function. *Embo J* **20**, 1153-63.
- Hogle, J. M. (2002). Poliovirus Cell entry: Common structural themes in viral cell entry pathways. *Annual Review of Microbiology* **56**, 677-702.
- Hope, D. A., Diamond, S. E. & Kirkegaard, K. (1997). Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. *J Virol* **71**, 9490-8.
- Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C. (1998). Structure of a Covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669-1675.
- Huber, H. E., McCoy, J. M., Sehra, J. S. & Richardson, C. C. (1989). Human immunodeficiency virus 1 reverse transcriptase template binding, processivity, strand displacement synthesis and template switching. *Journal of Biological Chemistry* **264**, 4669-4678.
- Hunt, S. L., Hsuan, J. J., Totty, N. & Jackson, R. J. (1999). unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA. *Genes & Development* **13**, 437-448.
- Hyypia, T., Hovi, T., Knowles, N. J. & Stanway, G. (1997). Classification of enteroviruses based on molecular and biological properties. *Journal of General Virology* **78**, 1-11.
- Irurzun, A., Perez, L. & Carrasco, L. (1992). Involvement of membrane traffic in the replication of poliovirus genomes - effects of brefeldin-A. *Virology* **191**, 166-175.
- Jacobson, A. B. & Zuker, M. (1993). Structural analysis by energy dot plot of a large mRNA. *J Mol Biol* **233**, 261-9.
- Jang, S. K. & Wimmer, E. (1990). Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes & Development* **4**, 1560-1572.

- Jarvis, T. C. & Kirkegaard, K. (1992). Poliovirus RNA recombination - mechanistic studies in the absence of selection. *EMBO Journal* **11**, 3135-3145.
- Jia, X. Y., Van Eden, M., Busch, M. G., Ehrenfeld, E. & Summers, D. F. (1998). trans-encapsidation of a poliovirus replicon by different picornavirus capsid proteins. *J Virol* **72**, 7972-7.
- Joachims, M., Van Breugel, P. C. & Lloyd, R. E. (1999). Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol* **73**, 718-27.
- Jore, J., De Geus, B., Jackson, R. J., Pouwels, P. H. & Enger Valk, B. E. (1988). Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. *Journal of General Virology* **69**, 1627-1636.
- Jurgens, C. & Flanagan, J. B. (2003). Initiation of poliovirus negative-strand RNA synthesis requires precursor forms of p2 proteins. *J Virol* **77**, 1075-83.
- Kaarianinen, L. & Ahola, T. (2002). Functions of alphavirus nonstructural proteins in RNA replication. *Prog.Nucleic.Acid.Res. Mol.Biol* **71**, 187-222.
- Kajigaya, S., Arakawa, H., Kuge, S., Koi, T., Imura, N. & Nomoto, A. (1985). Isolation and characterization of defective-interfering particles of poliovirus Sabin 1 strain. *Virology* **142**, 307-316.
- Kaminski, A., Hunt, S. L., Patton, J. G. & Jackson, R. J. (1995). Direct evidence that polypyrimidine tract binding-protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus-RNA. *RNA* **1**, 924-938.
- Kang, Y., Bogerd, H. P., Yang, J. & Cullen, B. R. (1999). Analysis of the RNA binding specificity of the human tap protein, a constitutive transport element- specific nuclear RNA export factor. *Virology* **262**, 200-209.
- Kao, C. C., Singh, P. & Ecker, D. J. (2001). De Novo Initiation of Viral RNA-dependent RNA synthesis. *Virology* **287**, 251-260.
- Kaplan, G. & Racaniello, V. R. (1988). Construction and characterization of poliovirus subgenomic replicons. *Journal of Virology* **62**, 1687-1696.
- Karnauchow, T. M., Tolson, D. L., Harrison, B. A., Altman, E., Lublin, D. M. & Dimock, K. (1996). The HeLa cell receptor for enterovirus 70 is decay-accelerating factor (CD55). *J Virol* **70**, 5143-52.
- Kean, K. M. (2003). The role of mRNA 5' non-coding and 3'end sequences on 40S ribosomal subunit recruitment, and how RNA viruses successfully compete with cellular mRNAs to ensure their own protein synthesis. *Biology of the Cell* **95**, 129-139.

- Kerekatte, V., Keiper, B. D., Badorff, C., Cai, A., Knowlton, K. U. & Rhoads, R. E. (1999). Cleavage of Poly(A)-binding protein by coxsackievirus 2A protease in vitro and in vivo: another mechanism for host protein synthesis shutoff? *Journal of Virology* **73**, 709-17.
- Kew, O., Morris-Glasgow, V., Landaverde, M., Burns, C., Shaw, J., Garib, Z., Andre, J., Blackman, E., Freeman, C. J., Jorba, J., Sutter, R., Tambini, G., Venczel, L., Pedreira, C., Laender, F., Shimizu, H., Yoneyama, T., Miyamura, T., van Der Avoort, H., Oberste, M. S., Kilpatrick, D., Cochi, S., Pallansch, M. & de Quadros, C. (2002). Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. *Science* **296**, 356-9.
- King, A. M. (1988). Preferred sites of recombination in poliovirus RNA: an analysis of 40 intertypic cross-over sequences. *Nucleic Acids Research* **16**, 11705-11723.
- Kirkegaard, K. & Baltimore, D. (1986). The mechanism of RNA recombination in poliovirus. *Cell* **47**, 433-443.
- Klovins, J., Berzins, V. & van Duin, J. (1998). A long-range interaction in Qbeta RNA that bridges the thousand nucleotides between the M-site and the 3' end is required for replication. *Rna* **4**, 948-57.
- Klovins, J. & van Duin, J. (1999). A long-range pseudoknot in Qbeta RNA is essential for replication. *J Mol Biol* **294**, 875-84.
- Koonin, E. V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *Journal of General Virology* **72**, 2197-2206.
- Krausslich, H. G., Nicklin, M. J., Toyoda, H., Etchison, D. & Wimmer, E. (1987). Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. *Journal of Virology* **61**, 2711-2718.
- Kremser, L., Okun, V. M., Nicodemou, A., Blaas, D. & Kenndler, E. (2004). Binding of Fluorescent Dye to Genomic RNA Inside Intact Human Rhinovirus after Viral Capsid Penetration by Capillary Electrophoresis. *Analytical Chemistry* **76**, 882-887.
- Kuge, S. & Nomoto, A. (1987). Construction of viable deletion and insertion mutants of the Sabin strain of type 1 poliovirus: function of the 5' noncoding sequence in viral replication. *Journal of Virology* **61**, 1478-1487.
- Kuge, S., Saito, I. & Nomoto, A. (1986). Primary structure of poliovirus defective-interfering particle genomes and possible generation mechanisms of the particles. *Journal of Molecular Biology* **192**, 473-487.
- Kusov, Y., Weitz, M., Dollenmeier, G., Gaussmuller, V. & Siegl, G. (1996). Rna-protein interactions at the 3'-end of the hepatitis-a virus- rna. *Journal of Virology* **70**, 1890-1897.
- Kusov, Y. Y. & GaussMuller, V. (1997). In vitro RNA binding of the hepatitis A virus proteinase 3C (HAV 3C(pro)) to secondary structure elements within the 5' terminus of the HAV genome. *Rna-a Publication of the Rna Society* **3**, 291-302.

- Laird-Offringa, I. A. & Belasco, J. G. (1995). Analysis of RNA-binding proteins by *in vitro* genetic selection: Identification of an amino acid residue important for locking U1A onto its RNA target. *Proceedings of the National Academy of Science USA* **92**, 11859-11863.
- Lall, S., Francis-Lang, H., Flament, A., Norvell, A., Schupbach, T. & Ish-Horowicz, D. (1999). Squid hnRNP protein promotes apical cytoplasmic transport and localization of *Drosophila* pair-rule transcripts. *Cell* **98**, 171-180.
- Lama, J. & Carrasco, L. (1992). Expression of Poliovirus Nonstructural Proteins in Escherichia- Coli Cells - Modification of Membrane Permeability Induced by 2B and 3A. *Journal of Biological Chemistry* **267**, 15932-15937.
- Lama, J., Paul, A. V., Harris, K. S. & Wimmer, E. (1994). Properties of purified recombinant poliovirus protein 3AB as substrate for viral proteinases and as co-factor for RNA polymerase 3Dpol. *J Biol Chem* **269**, 66-70.
- Lamphear, B. J., Kirkweger, R., Skern, T. & Rhoads, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. *Journal of Biological Chemistry* **270**, 21975-21983.
- Lamphear, B. J., Yan, R., Yang, F., Waters, D., Liebig, H.-D., Klump, H., Kuechler, E., Skern, T. & Rhoads, R. E. (1993). Mapping the cleavage site in protein synthesis initiation factor eIF-4G of the 2A proteases from human coxsackievirus and rhinovirus. *Journal of Biological Chemistry* **268**, 19200-19203.
- Lawson, M. A., Dasmahapatra, B. & Semler, B. L. (1990). Species-specific substrate interaction of picornavirus 3C proteinase suballelic exchange mutants. *Journal of Biological Chemistry* **265**, 15920-15931.
- Le, S. Y., Chen, J. H., Sonenberg, N. & Maizel, J. V. (1992). Conserved tertiary structure elements in the 5' untranslated region of human enteroviruses and rhinoviruses. *Virology* **191**, 858-866.
- Lee, C. K. & 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.
- Lee, E.-G., Yeo, A., Kraemer, B., Wickens, M. & Linial, M. (1999). The gag domain required for avian retroviral RNA encapsidation determined by using two independent assays. *Journal of Virology* **73**, 6282-6292.
- Lee, Y. F., Nomoto, A., Detjen, B. M. & Wimmer, E. (1977). The genome-linked protein of picornaviruses I. A protein covalently linked to poliovirus genome RNA. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 59-63.
- Li, J. P. & Baltimore, D. (1990). An intragenic revertant of a poliovirus 2C mutant has an uncoating defect. *Journal of Virology* **64**, 1102-1107.

- Liljestrom, P., Lusa, S., Huylebroeck, D. & Garoff, H. (1991). In vitro mutagenesis of full-length cDNA clone of Semlike Forest Virus: the small 6,000- molecular weight membrane protein modulates virus release. *Journal of Virology* **65**, 4107-4113.
- Lindberg, A. M., Andersson, P., Savolainen, C., Mulders, M. N. & Hovi, T. (2003). Evolution of the genome of Human Enterovirus B: incongruence between phylogenies of the VP1 and 3CD regions indicates frequent recombination within the species. *Journal of General Virology* **84**, 1223-1235.
- Lippencott-Schwartz, J., Yuan, L., Bonifacino, J. S. & Klausner, R. D. (1989). Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from the Golgi to ER. *Cell* **56**, 801-813.
- Lippencott-Schwartz, J., Yuan, L., Tipper, C., Amherdt, M., Orci, L. & Klausner, R. D. (1991). Brefeldin A's effects on endosomes, lysosomes, and the TGN suggest a general mechanism for regulating organelle structure and membrane traffic. *Cell* **67**, 601-616.
- Liu, H., Zheng, D., Zhang, L., Oberste, M. S., Kew, O. M. & Pallensch, M. A. (2003). Serial recombination during circulation of Type 1 Wild vaccine Recombinant Poliovirus in China. *Journal of Virology* **77**, 10994-11005.
- Lobert, P. E., Escriou, N., Ruelle, J. & Michiels, T. (1999). A coding RNA sequence acts as a replication signal in cardioviruses. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11560-11565.
- Lohmann, V., Kroner, F., Herian, U. & Bartenschlager, R. (1997). Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *Journal of Virology* **71**, 8416-8428.
- Lommel, S. A., Morris, T. J. & Pinnock, D. E. (1985). Characterization of nucleic acids associated with Arkansas bee virus. *Intervirology* **23**, 199-207.
- Long, A. C., Orr, D. C., Cameron, J. M., Dunn, B. M. & Kay, J. (1989). A consensus sequence for substrate hydrolysis by rhinovirus 3C proteinase. *FEBS Letters* **258**, 75-78.
- Long, R. M., Gu, W., Lorimer, E., Singer, R. H. & Chartrand, P. (2000). She2p is a novel RNA-binding protein that recruits the Myo4p-She2p complex to ASH1 mRNA. *EMBO journal* **19**, 6592-6601.
- Lowary, P. T. & Uhlenbeck, O. C. (1987). An RNA mutation that increases the affinity of an RNA-protein interaction. *Nucleic Acid Research* **15**, 10483-10493.
- Lu, H. H., Li, X. Y., Cuconati, A. & Wimmer, E. (1995). Analysis of Picornavirus 2A(Pro) Proteins - Separation of Proteinase From Translation and Replication Functions. *Journal of Virology* **69**, 7445-7452.

- Lukashev, A. N., Lashkevich, V. A., Ivanova, O. E., Koroleva, G. A., Hinkkanen, A. E. & Ilonen, J. (2003). Recombination in circulating enteroviruses. *Journal of Virology* **77**, 10423-10431.
- Lyle, J. M., Bullitt, E., Bienz, K. & Kirkegaard, K. (2002a). Visualization and functional analysis of RNA-dependent RNA polymerase lattices. *Science* **296**, 2218-22.
- Lyle, J. M., Clewell, A., Richmond, K., Richards, O. C., Hope, D. A., Schultz, S. C. & Kirkegaard, K. (2002b). Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. *J Biol Chem* **277**, 16324-31.
- Ma, J. & Ptashne, M. (1988). Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55**, 443-6.
- Martin, F., Michel, F., Zenkluson, D., Muller, B. & Schumperli, D. (2000). Positive and Negative Selection in the human histone-binding protein using the yeast three-hybrid system. *Nucleic Acid Research* **28**, 1594-1603.
- Masson, L., Tabashnik, E., Liu, Y. B., Brousseau, R. & Schwartz, J. L. (1999). Helix 4 of the Baccillus thuringiensis CryIAa Toxin Lines the Lumen of the Ion Channel. *Journal of Biological Chemistry* **274**, 31996-32000.
- Mathews, D. H., Sabina, J., Zuker M & Turner, D. H. (1999). Expanded Sequence Dependence of Thermodynamic Parameters improves Prediction of RNA secondary structure. *Journal of Molecular Biology* **288**, 911-940.
- Mathews, M. B. (1995). Structure, function and evolution of adenovirus virus-associated RNAs. *Current Topics in Microbiology and Immunology* **199**, 173-187.
- Mayeda, A., Munroe, S. H., Caceres, J. F. & Krainer, A. R. (1994). Function of conserved domains of hnRNP A1 and other hnRNP A/B. *EMBO* **13**, 5483-5495.
- Maynell, L. A., Kirkegaard, K. & Klymkowsky, M. W. (1992). Inhibition of Poliovirus RNA Synthesis by Brefeldin-A. *Journal of Virology* **66**, 1985-1994.
- McBride, A. E., Schlegel, A. & Kirkegaard, K. (1996). Human Protein Sam68 Relocalization and Interaction With Poliovirus RNA-Polymerase in Infected-Cells. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2296-2301.
- McInerney, G. M., King, A. M., Ross_Smith, N. & Belsham, G. J. (2000). Replication-competent foot-and-mouth disease virus RNAs lacking capsid coding sequences. *Journal of General Virology* **81 Pt 7**, 1699-702.
- McKnight, K. L. & Lemon, S. M. (1996). Capsid Coding Sequence Is Required For Efficient Replication of Human Rhinovirus-14 RNA. *Journal of Virology* **70**, 1941-1952.

- McKnight, K. L. & Lemon, S. M. (1998). The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. *RNA* **4**, 1569-84.
- Meerovitch, K., Pelletier, J. & Sonenberg, N. (1989). A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implications for internal translation initiation. *Genes & Development* **3**, 1026-1034.
- Meerovitch, K. & Sonenberg, N. (1993). Internal Initiation of Picornavirus RNA Translation. *Semin. Virol.* **4**, 217-227.
- Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowitz, F., Kenan, D. J., Chan, E. K., Agol, V. I., Keene, J. D. & Sonenberg, N. (1993). La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J Virol* **67**, 3798-807.
- Melchers, W. J., Hoenderop, J. G., Bruins Slot, H. J., Pleij, C. W., Pilipenko, E. V., Agol, V. I. & Galama, J. M. (1997). Kissing of the two predominant hairpin loops in the coxsackie B virus 3' untranslated region is the essential structural feature of the origin of replication required for negative-strand RNA synthesis. *J Virol* **71**, 686-96.
- Melnick, J. L. (1996a). Current Status of Poliovirus Infections. *Clinical Microbiology Reviews* **9**, 293-300.
- Melnick, J. L. (1996b). Enteroviruses : Polioviruses, Coxsackieviruses, Echoviruses, and Newer Enteroviruses. In *Fields Virology*, pp. 655-712. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott-Raven.
- Meredith, J. M., Rohll, J. B., Almond, J. W. & Evans, D. J. (1999). Similar Interactions of the Poliovirus and Rhinovirus 3D polymerase with the 3' Untranslated Region of Rhinovirus 14. *Journal of Virology* **73**, 9952-9958.
- Merkle, I., van Ooij, M. J., van Kuppeveld, F. J., Glaudemans, D. H., Galama, J. M., Henke, A., Zell, R. & Melchers, W. J. (2002). Biological significance of a human enterovirus B-specific RNA element in the 3' nontranslated region. *J Virol* **76**, 9900-9.
- Miller, D. J., Schwartz, M. D. & Ahlquist, P. (2001). Flock House virus RNA replicates on outer mitochondrial membranes in *Drosophila* cells. *Journal of Virology* **75**, 11664-11676.
- Miller, D. J., Schwartz, M. D., Dye, B. T. & Ahlquist, P. (2003). Engineered Retargeting of viral RNA Replication Complexes to an Alternative Intracellular Membrane. *Journal of Virology* **77**, 12193-12202.
- Miller, J. (1972). *Experiments in Molecular Genetics*: Harbor Laboratory Press, Cold Spring Harbor, NY.

- Mirmomeni, M. H., Hughes, P. J. & Stanway, G. (1997). An RNA tertiary structure in the 3' untranslated region of enteroviruses is necessary for efficient replication. *Journal of Virology* **71**, 2363-2370.
- Mirzayan, C. & Wimmer, E. (1994a). Biochemical studies on poliovirus polypeptide 2C: evidence for ATPase activity. *Virology* **199**, 176-87.
- Mirzayan, C. & Wimmer, E. (1994b). Biochemical-studies on poliovirus polypeptide 2c - evidence for atpase activity. *Virology* **199**, 176-187.
- Molla, A., Harris, K. S., Paul, A. V., Shin, S. H., Mugavero, J. & Wimmer, E. (1994). Stimulation of poliovirus proteinase 3Cpro-related proteolysis by the genome-linked protein VPg and its precursor 3AB. *J Biol Chem* **269**, 27015-20.
- Molla, A., Paul, A. V. & Wimmer, E. (1991). Cell-free, *de novo* synthesis of poliovirus. *Science* **254**, 1647-1651.
- Molla, A., Paul, A. V. & Wimmer, E. (1993). Effects of temperature and lipophilic agents on poliovirus formation and RNA synthesis in a cell-free system. *J Virol* **67**, 5932-8.
- Morasco, B. J., Sharma, N., Parilla, J. & Flanagan, J. B. (2003). Poliovirus *cre*(2C)-Dependent Synthesis of VPgpUpU is Required for Positive- but not Negative-Strand RNA synthesis. *Journal of Virology* **77**, 5136-5144.
- Morrow, C. D., Warren, B. & Lentz, M. R. (1987). Expression of enzymatically active poliovirus RNA-dependent RNA polymerase in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 6050-6054.
- Mouland, A. J., Xu, H., Cui, H., Krueger, W., Munro, T. P., Prasol, M., Mercier, J., Rekosh, D., Smith, R., Barbarese, E., Cohen, E. A. & Carson, J. H. (2001). RNA trafficking signals in human immunodeficiency virus type I. *Molecular and cellular biology* **21**, 2133-2143.
- Mueller, S., Cao, X., Welker, R. & Wimmer, E. (2002). Interaction of the Poliovirus Receptor CD155 with the Dynein Light Chain Tctex-1 and Its Implication for Poliovirus Pathogenesis. *The Joournal of Biological Chemistry* **77**, 7897-7904.
- Murray, K. E. & Barton, D. J. (2003). Poliovirus CRE-Dependent VPg Uridylylation Is Required for Positive-Strand RNA Synthesis but Not for Negative-Strand RNA Synthesis. *J. Virol.* **77**, 4739-4750.
- Muscio, O. A., La Torre, J. L. & Scodeller, E. A. (1988). Characterization of Triatoma virus, a picorna-like virus isolated from the triatomine bug Triatoma infestans. *Journal of General Virology* **69**, 2929-2934.
- Neufeld, K. L., Galarza, J. M., Richards, O. C., Summers, D. F. & Ehrenfeld, E. (1994). Identification of terminal adenylyl transferase activity of the poliovirus polymerase 3Dpol. *J Virol* **68**, 5811-8.

- Neznanov, N., Kondratova, A., Chumakov, K. M., Angres, B., Zhumabayeva, B., Agol, V. I. & Gudkov, A. V. (2001). Poliovirus protein 3a inhibits tumor necrosis factor (tnf)-induced apoptosis by eliminating the tnf receptor from the cell surface. *Journal of Virology* **75**, 10409-20.
- Ng, K. K. S., Cherney, M. M., Vazquez, A. L., Machin, A., Alonso, J. M. M., Parra, F. & James, M. N. G. (2002). Crystal structures of active and inactive conformations of a caliciviral RNA-dependent RNA polymerase. *Journal of Biological Chemistry* **277**, 1381-1387.
- Nicholson, R., Pelletier, J., Le, S. Y. & Sonenberg, N. (1991). Structural and functional analysis of the ribosome landing pad of poliovirus type-2 - in vivo translation studies. *Journal of Virology* **65**, 5886-5894.
- Nicklin, M. J., Harris, K. S., Pallai, P. V. & Wimmer, E. (1988). Poliovirus proteinase 3C: large-scale expression, purification, and specific cleavage activity on natural and synthetic substrates in vitro. *Journal of Virology* **62**, 4586-4593.
- Nieva, J. L., Agirre, A., Nir, S. & Carrasco, L. (2003). Mechanisms of membrane permeabilization by picornavirus 2B viroporin. *FEBS letters* **552**, 68-73.
- Novak, J. E. & Kirkegaard, K. (1991). Improved Method for Detecting Poliovirus Negative Strands Used to Demonstrate Specificity of Positive-Strand Encapsidation and the Ratio of Positive to Negative Strands in Infected Cells. *Journal of Virology* **65**, 3384-3387.
- Nugent, C. I., Johnson, K. L., Sarnow, P. & Kirkegaard, K. (1999). Functional coupling between replication and packaging of poliovirus replicon RNA. *J Virol* **73**, 427-35.
- Nurani, G., Lindqvist, B. & Casasnovas (2003). Receptor Priming of Major Group Rhinoviruses for uncoating and entry at Mild low-pH environments. *Journal of Virology* **77**, 11985-11991.
- Oberste, M. S., Maher, K. & Pallensch, M. A. (2004). Evidence for frequent recombination within species *human enterovirus B* based on complete genomic of all thirty-seven serotypes. *Journal of Virology* **78**, 855-867.
- Ochs, K., Saleh, L., Bassili, G., Sonntag, V. H., Zeller, A. & Niepmann, M. (2002). Interaction of translation initiation factor eIF4B with the poliovirus internal ribosome entry site. *J Virol* **76**, 2113-22.
- Ochs, K., Zeller, A., Saleh, L., Bassili, G., Song, Y., Sonntag, A. & Niepmann, M. (2003). Impaired binding of standard initiation factors mediates poliovirus translation attenuation. *J Virol* **77**, 115-22.
- O'Farrell, D., Trowbridge, R., Rowlands, D. J. & Jager, J. (2003). Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): Structural evidence for nucleotide import and *de novo* initiation. *Journal of Molecular Biology* **326**, 1025-1035.

- Ohka, S., Yang, W. X., Terada, E., Iwasaki, K. & Nomoto, A. (1998). Retrograde transport of intact poliovirus through the axon via the fast transport system. *Virology* **250**, 67-75.
- Okita, T. W. & Choi, S. (2002). mRNA localization in plants: targeting to the cell's cortical region and beyond. *Current opinion in Plant biology* **5**, 553-559.
- Oleynikov, Y. & Singer, R. H. (1998). RNA localization: different zipcodes, same postman? *Trends in cell biology* **8**, 381-383.
- Omata, T., Horie, H., Kuge, S., Imura, N. & Nomoto, A. (1986). Mapping and sequencing of RNAs without recourse to molecular cloning: application to RNAs of the Sabin 1 strain of poliovirus and its defective interfering particles. *J. Biochem. (Tokyo)* **99**, 207-217.
- Ouzilou, L., Caliot, E., Pelletier, I., Prevost, M. C., Pringault, E. & Colbere-Garapin, F. (2002). Poliovirus transcytosis through M-like cells. *J Gen Virol* **83**, 2177-82.
- Pacheco, J. M., Henry, T. M., O'Donnell, V. K., Gregory, J. B. & Mason, P. W. (2003). Role of Nonstructural proteins 3A and 3B in host range and pathogenicity of foot and mouth disease virus. *Journal of Virology* **77**, 13017-13027.
- Pallansch, M., Kew, O., Semler, B., Omilianowski, D., Anderson, C., Wimmer, E. & Ruekert, R. (1984). Protein processing map of poliovirus type 1. *Journal of Virology* **49**, 873-880.
- Pallansch, M. & Roos, R. P. (2001). Enteroviruses: Polioviruses, Coxsackieviruses, Echoviruses and Newer Enteroviruses. In *Fields Virology*, Fourth edn, pp. 723-777. Edited by D. M. Knipe, P. M. Howley & D. E. Griffin: Lippincott-Williams and Williams.
- Park, J., Desvoyes, B. & Scholthof, H. B. (2002). *Tomato Bushy stunt virus* Genomic RNA Accumulation is regulated by interdependent *cis*-acting Elements within the Movement Protein open reading frame. *Journal of virology* **76**, 12747-12757.
- Parsley, T. B., Towner, J. S., Blyn, L. B., Ehrenfeld, E. & Semler, B. L. (1997). Poly (rC) binding protein 2 forms a ternary complex with the 5'- terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**, 1124-34.
- Pata, J. D., Schultz, S. C. & Kirkegaard, K. (1995). Functional oligomerization of poliovirus RNA-dependent RNA polymerase. *RNA* **1**, 466-77.
- Pathak, H. B., Ghosh, S. K., Roberts, A. W., Sharma, S. D., Yoder, J. D., Arnold, J. J., Gohara, D. W., Barton, D. J., Paul, A. V. & Cameron, C. E. (2002). Structure-function relationships of the RNA-dependent RNA polymerase from poliovirus (3Dpol). A surface of the primary oligomerization domain functions in capsid precursor processing and VPg uridylylation. *J Biol Chem* **277**, 31551-62.
- Paul, A. V., Cao, X., Harris, K. S., Lama, J. & Wimmer, E. (1994a). Studies with poliovirus polymerase 3Dpol. Stimulation of poly(U) synthesis in vitro by purified poliovirus protein 3AB. *J Biol Chem* **269**, 29173-81.

- Paul, A. V., Molla, A. & Wimmer, E. (1994b). Studies of a putative amphipathic helix in the N-terminus of poliovirus protein 2C. *Virology* **199**, 188-99.
- Paul, A. V., Peters, J., Mugavero, J., Yin, J., Van Boom, J. H. & Wimmer, E. (2003). Biochemical and Genetic Studies of the VPg Uridylylation Reaction Catalyzed by the RNA Polymerase of Poliovirus. *J Virol* **77**, 891-904.
- Paul, A. V., Rieder, E., Kim, D. W., van Boom, J. H. & Wimmer, E. (2000). Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J Virol* **74**, 10359-70.
- Paul, A. V., Schultz, A., Pincus, S. E., Oroszlan, S. & Wimmer, E. (1987). Capsid protein VP4 of poliovirus is N-myristoylated. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7827-7831.
- Pelletier, J., Kaplan, G., Racaniello, V. R. & Sonenberg, N. (1988). Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. *Molecular and Cellular Biology* **8**, 1103-1112.
- Pelletier, J. & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320-325.
- Percy, N., Barclay, W. S., Sullivan, M. & Almond, J. W. (1992). A Poliovirus Replicon Containing the Chloramphenicol Acetyltransferase Gene Can Be Used to Study the Replication and Encapsidation of Poliovirus RNA. *Journal of Virology* **66**, 5040-5046.
- Perez, L. & Carrasco, L. (1993). Entry of poliovirus into cells does not require a low-pH step. *Journal of Virology* **67**, 4543-4548.
- Pestova, T. V., Shatsky, I. N. & Hellen, C. U. T. (1996). Functional dissection of eukaryotic initiation-factor 4F - the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43s preinitiation complexes. *Molecular and Cellular Biology* **16**, 6870-6878.
- Petersen, J. F., Cherney, M. M., Liebig, H. D., Skern, T., Kuechler, E. & James, M. N. (1999). The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. *Embo Journal* **18**, 5463-75.
- Pfeffer, S., Zavolan, M., Grasser, F. A., Chein, M., Russo, J. J., Ju, J., John, B., Enright, A. J., Marks, D., Sander, C. & Tuschl, T. (2004). Identification of Virus-Encoded MicroRNAs. *Science* **304**, 734-736.
- Pfister, T. & Wimmer, E. (1999). Characterization of the nucleoside triphosphate activity of poliovirus protein 2C reveals a mechanism by which Guanidine inhibits poliovirus replication. *The Journal of Biological Chemistry* **274**, 6992-7001.

- Pierangeli, A., Bucci, M., Pagnotti, P., Degener, A. M. & Perez Bercoff, R. (1995). Mutational analysis of the 3'-terminal extra-cistronic region of poliovirus RNA: secondary structure is not the only requirement for minus strand RNA replication. *FEBS Lett* **374**, 327-32.
- Pilipenko, E. V., Blinov, V. M., Chernov, B. K., Dmitrieva, T. M. & Agol, V. I. (1989a). Conservation of the secondary structure elements of the 5'- untranslated region of cardio and aphthovirus RNAs. *Nucleic Acids Research* **17** No. **14**, 5701-5711.
- Pilipenko, E. V., Blinov, V. M., Romanova, L. I., Sinyakov, A. N., Maslova, S. V. & Agol, V. I. (1989b). Conserved structural domains in the 5'-untranslated region of picornaviral genomes: an analysis of the segment controlling translation and neurovirulence. *Virology* **168**, 201-209.
- Pilipenko, E. V., Gmyl, A. P., Maslova, S. V., Svitkin, Y. V., Sinyakov, A. N. & Agol, V. I. (1992a). Prokaryotic-Like Cis Elements in the Cap-Independent Internal Initiation of Translation on Picornavirus RNA. *Cell* **68**, 119-131.
- Pilipenko, E. V., Maslova, S. V., Sinyakov, A. N. & Agol, V. I. (1992b). 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. *Nucleic Acids Res* **20**, 1739-45.
- Pilipenko, E. V., Poperechny, K. V., Maslova, S. V., Melchers, W. J., Slot, H. J. & Agol, V. I. (1996). Cis-element, oriR, involved in the initiation of (-) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions. *EMBO J* **15**, 5428-36.
- Plotch, S. J. & Palant, O. (1995). Poliovirus protein 3AB forms a complex with and stimulates the activity of the viral RNA polymerase, 3Dpol. *J Virol* **69**, 7169-79.
- Plotch, S. J., Palant, O. & Gluzman, Y. (1989). Purification and properties of poliovirus RNA polymerase expressed in Escherichia coli. *Journal of Virology* **63**, 216-225.
- Poch, O., Sauvaget, I., Delarue, M. & Torodo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO journal* **12**, 3867-3874.
- Pogany, J., Fabian, M., White, K. & Nagy, P. (2003). A replication silencer element in a plus-strand RNA virus. *EMBO* **22**, 5602-5611.
- Porter, D. C., Ansardi, D. C., Wang, J., McPherson, S., Moldoveanu, Z. & Morrow, C. D. (1998). Demonstration of the specificity of poliovirus encapsidation using a novel replicon which encodes enzymatically active firefly luciferase. *Virology* **243**, 1-11.
- Powell, R. M., Ward, T., Evans, D. J. & Almond, J. W. (1997). Interaction between echovirus 7 and its receptor, decay-accelerating factor (CD55): evidence for a secondary cellular factor in A-particle formation. *J Virol* **71**, 9306-12.

- Poyry, T., Kinnunen, L., Hyypia, T., Brown, B., Horsnell, C., Hovi, T. & Stanway, G. (1996). Genetic and phylogenetic clustering of enteroviruses. *Journal of General Virology* **77**, 1699-1717.
- Pulli, T., Koskimies, P. & Hyypia, T. (1995). Molecular comparison of coxsackie A virus serotypes. *Virology* **212**, 30-8.
- Rajagoplan, P. T., Datta, A. & Pei, D. (1997). Purification, characterization, and inhibition of peptide deformylase from *Escherichia coli*. *Biochemistry* **36**, 13910-13918.
- Ray, P. S. & Das, S. (2002). La autoantigen is required for the internal ribosome entry site-mediated translation of Coxsackievirus B3 RNA. *Nucleic Acids Res* **30**, 4500-8.
- Reardon, J. E. (1993). Human immunodeficiency virus reverse transcriptase A kinetic analysis of RNA-dependent and DNA-dependent DNA polymerisation. *Journal of Biological Chemistry* **140**, 1-22.
- Reilly, E. K. & Kao, C. C. (1998). Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structures. *Virology* **252**, 287-303.
- Ren, R., Costantini, F., Gorgacz, E. J., Lee, J. J. & Racaniello, V. R. (1990). Transgenic mice expressing a human poliovirus receptor: a new model for poliomyelitis. *Cell* **63**, 353-362.
- Repass, J. F. & Makino, S. (1998). Importance of the positive-strand RNA secondary structure of a murine coronavirus defective interfering RNA internal replication signal in positive-strand RNA synthesis. *Journal of Virology* **72**, 7926-7933.
- Rho, S. B. & Martinis, S. A. (2000). The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. *RNA* **6**, 1882-1894.
- Richards, O. C., Yu, P., Neufeld, K. L. & Ehrenfeld, E. (1992). Nucleotide Binding by the Poliovirus RNA Polymerase. *Journal of Biological Chemistry* **267**, 17141-17146.
- Rieder, E., Paul, A. V., Kim, D. W., van Boom, J. H. & Wimmer, E. (2000). Genetic and biochemical studies of poliovirus cis-acting replication element cre in relation to VPg uridylation. *J Virol* **74**, 10371-80.
- Rodriguez, P. L. & Carrasco, L. (1993). Poliovirus protein-2c has atpase and gtpase activities. *Journal of Biological Chemistry* **268**, 8105-8110.
- Rodriguez, P. L. & Carrasco, L. (1995). Poliovirus protein 2c contains 2 regions involved in rna-binding activity. *Journal of Biological Chemistry* **270**, 10105-10112.
- Rodriguez-Wells, V., Plotch, S. J. & DeStefano, J. J. (2001). Primer-dependent synthesis by poliovirus RNA-dependent RNA polymerase (3D(pol)). *Nucleic Acids Res* **29**, 2715-24.

- Roehl, H. H. & Semler, B. L. (1995). Poliovirus infection enhances the formation of two ribonucleoprotein complexes at the 3' end of viral negative-strand RNA. *J Virol* **69**, 2954-61.
- Rohll, J. B., Moon, D. H., Evans, D. J. & Almond, J. W. (1995). The 3'-untranslated region of picornavirus RNA - features required for efficient genome replication. *Journal of Virology* **69**, 7835-7844.
- Romanova, L. I., Blinov, V. M., Tolskaya, E. A., Kolesnikova, M. S., Viktorova, E. G., Guseva, E. A. & Agol, V. I. (1986). The primary structure of crossover regions of intertypic poliovirus recombinants: A model of recombination between RNA genomes. *Virology* **155**, 202-213.
- Ronco, L. V., Karpova, A. Y., Vidal, M. & Howley, P. M. (1998). Human papillomavirus 16 E6 oncoprotein binds to and inhibits its transcriptional activity. *Genes and Development* **12**, 2061-2072.
- Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. (1978a). The genome-linked protein of picornaviruses V. O4-(5'- Uridyl)- tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4868-4872.
- Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. (1978b). O4-(5'- uridyl)tyrosine is the bond between the genome linked protein & the RNA of poliovirus. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4868-4872.
- Rueckert, R. R. (1996). *Picornaviridae: the viruses and their replication*. In *Fields Virology*, 3 edn, pp. 609-654. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott-Raven.
- Ruigrok, R. W. H., Hirst, E. M. A. & Hay, A. J. (1991). The specific inhibition of influenza A virus maturation by amantadine: an electron microscopic examination. *Journal of General Virology* **72**, 191-194.
- Rust, R. C., Landmann, L., Gosert, R., Tang, B. L., Hong, W., Hauri, H. P., Egger, D. & Bienz, K. (2001). Cellular COPII proteins are involved in production of the vesicles that form the poliovirus replication complex. *J Virol* **75**, 9808-18.
- Sakoda, Y., Ross-Smith, N., Inoue, T. & Belsham, G. J. (2001). An attenuating mutation in the 2A protease of swine vesicular disease virus, a picornavirus, regulates cap- and internal ribosome entry site-dependent protein synthesis. *J Virol* **75**, 10643-50.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A laboratory Manual*: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sambrook, J. F. (1990). The involvement of calcium in transport of secretory proteins from the endoplasmic reticulum. *Cell* **61**, 197-199.
- Sandman, K., Grayling, R. A. & Reeve, J. N. (1995). Improved amino-processing of recombinant proteins synthesized in *Escherichia coli*. *Biotechnology* **13**, 504-506.

- Sandoval, I. V. & Carrasco, L. (1997). Poliovirus infection and expression of the poliovirus protein 2B provoke the disassembly of the Golgi complex, the organelle target for the antipoliovirus drug Ro-090179. *J Virol* **71**, 4679-93.
- Sarnow, P. (1989). Role of 3'-end sequences in infectivity of poliovirus transcripts made in vitro. *Journal of Virology* **63**, 467-470.
- Sasaki, J., Kusuhara, Y., Maeno, Y., Kobayashi, N., Yamashita, T., Sakae, K., Takeda, N. & Taniguchi, H. (2001). Construction of an infectious cDNA clone of the Aichi virus (a new member of the family *Picornaviridae*) and mutational analysis of a stem-loop structure at the 5' end of the genome. *Journal of Virology* **75**, 8021-8030.
- Sasaki, J. & Taniguchi, K. (2003). The 5'-End Sequence of the Genome of Aichi Virus, a Picornavirus, Contains an Element Critical for Viral RNA Encapsidation. *J. Virol.* **77**, 3542-3548.
- Schlesinger, S., Makino, S. & Linial, M. L. (1994). Cis-Acting Genomic Elements and Trans-Acting Proteins Involved in the Assembly of Rna Viruses. *Seminars in Virology* **5**, 39-49.
- Schober, D., Kronenberger, P., Prchla, E., Blaas, D. & Fuchs, R. (1998). Major and minor receptor group human rhinoviruses penetrate from endosomes by different mechanisms. *J Virol* **72**, 1354-64.
- Semler, B. L., Anderson, C. W., Hanecak, R., Dorner, L. F. & Wimmer, E. (1982). A membrane-associated precursor to poliovirus VPg identified by immunoprecipitation with antibodies directed against a synthetic heptapeptide. *Cell* **28**, 405-412.
- Sengupta, D. J., Wickens, M. & Fields, S. (1999). Identification of RNAs that bind to a specific protein using the yeast three-hybrid system. *Rna* **5**, 596-601.
- SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. & Wickens, M. (1996). A three-hybrid system to detect RNA-protein interactions in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 8496-501.
- Shafren, D. R., Dorahy, D. J., Ingham, R. A., Burns, G. F. & Barry, R. D. (1997). Coxsackievirus A21 binds to decay-accelerating factor but requires intercellular adhesion molecule 1 for cell entry. *Journal of Virology* **71**, 4736-4743.
- Shan, J., Munro, T. P., Barbarese, E., Carson, J. H. & Smith, R. (2003). A Molecular mechanism for mRNA trafficking in neuronal dendrites. *The journal of neuroscience* **26**, 8859-8866.
- Shieh, J. T. & Bergelson, J. M. (2002). Interaction with decay-accelerating factor facilitates coxsackievirus B infection of polarized epithelial cells. *J Virol* **76**, 9474-80.
- Shiroki, K., Isoyama, T., Kuge, S., Ishii, T., Ohmi, S., Hata, S., Suzuki, K., Takasaki, Y. & Nomoto, A. (1999). Intracellular redistribution of truncated La protein produced by poliovirus3Cpro-mediated cleavage. *Journal of Virology* **73**, 2193-2200.

- Simmonds, P. & Smith, D. B. (1999). Structural constraints on RNA virus evolution. *J Virol* **73**, 5787-94.
- Smith, D. B. & Simmonds, P. (1997). Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both end of the genome. *Journal of Molecular Evolution* **45**, 238-246.
- Sokolowski, M., Scott, J. E., Heaney, R. P., Patel, A. H. & Clements, J. B. (2003). Identification of Herpes simplex virus RNAs that interact specifically with regulatory protein ICP27 *in vivo*. *The journal of Biological chemistry* **278**, 33540-33549
- Sonoda, J. & Wharton, R. P. (1999). Recruitment of Nanos to hunchback mRNA by Pumilio. *Genes and Development* **13**, 2704-2712.
- Sonoda, J. & Wharton, R. P. (2001). Drosophila Brain Tumor is a translational repressor. *Genes and Development* **15**, 762-773.
- Suhy, D. A., Giddings, T. H., Jr. & Kirkegaard, K. (2000). Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. *J Virol* **74**, 8953-65.
- Svitkin, Y. V., Maslova, S. V. & Agol, V. I. (1985). The genomes of attenuated and virulent poliovirus strains differ in their *in vitro* translation efficiencies. *Virology* **147**, 243-252.
- Tershak, D. R. (1982). Inhibition of poliovirus polymerase by guanidine *In Vitro*. *Journal of Virology* **41 no. 1**, 313-318.
- Teterina, N., Rinaudo, M. S. & Ehrenfeld, E. (2003). Strand-specific RNA synthesis Defects in a poliovirus with a mutation in protein 3A. *Journal of Virology* **77**, 12679-12691.
- Teterina, N. L., Gorbalenya, A. E., Egger, D., Bienz, K. & Ehrenfeld, E. (1997). Poliovirus 2C protein determinants of membrane binding and rearrangements in mammalian cells. *J Virol* **71**, 8962-72.
- Teterina, N. L., Kean, K. M., Gorbalenya, A. E., Agol, V. I. & Girard, M. (1992). Analysis of the Functional Significance of Amino Acid Residues in the Putative NTP-Binding Pattern of the Poliovirus-2C Protein. *Journal of General Virology* **73**, 1977-1986.
- Tinoco, I., Jr., & Bustamante, C. (1999). How RNA Folds. *Journal of Molecular Biology* **293**, 271-281.
- Todd, S., Towner, J. S., Brown, D. M. & Semler, B. L. (1997). Replication-competent picornaviruses with complete genomic RNA 3' noncoding region deletions. *Journal of Virology* **71**, 8868-8874.

- Tolskaya, E. A., Romanova, L. I., Blinov, V. M., Viktorova, E. G., Sinyakov, A. N., Kolesnikova, M. S. & Agol, V. I. (1987). Studies on the recombination between RNA genomes of poliovirus: the primary structure and nonrandom distribution of crossover regions in the genomes of intertypic poliovirus recombinants. *Virology* **161**, 54-61.
- Toriyama, S., Guy, P. L., Fuji, S. & Takahashi, M. (1992). Characterization of a new picorna-like virus, himetobi P virus, in planthoppers. *Journal of General Virology* **73**, 1021-1023.
- Towner, J. S., Ho, T. V. & Semler, B. L. (1996). Determinants of membrane association for poliovirus protein 3AB. *J Biol Chem* **271**, 26810-8.
- Toyoda, H., Nicklin, M. J., Murray, M. G., Anderson, C. W., Dunn, J. J. & Wimmer, E. (1986). A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* **45**, 761-770.
- Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. (1988). Functional dissection of VP16 the trans-activator of herpes simplex virus immediate early gene expression. *Genes and Development* **2**, 718-729.
- Troxler, M., Egger, D., Pfister, T. & Bienz, K. (1992a). Intracellular localization of poliovirus RNA by in situ hybridization at the ultrastructural level using single-stranded riboprobes. *Virology* **191**, 687-97.
- Troxler, M., Egger, D., Pfister, T. & Bienz, K. (1992b). Intracellular-localization of poliovirus rna by insitu hybridization at the ultrastructural level using single-stranded riboprobes. *Virology* **191**, 687-697.
- Tsai, C.-H., Cheng, C.-W., Peng, C.-W., Lin, B.-Y., Lin, N. S. & Hsu, Y. H. (1999). Sufficient length of a poly(A) tail for the formation of a potential pseudokont is required for efficient replication of bamboo mosaic potexvirus RNA. *Journal of Virology* **73**, 2703-2709.
- Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510.
- Tuplin, A., Wood, J., Evans, D. J., AH, P. & Simmonds, P. (2002). Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus. *RNA* **8**, 824-841.
- Van Dijk, A. A., Mayekev, E. V. & Bamford, D. H. (2004). Initiation of viral RNA-dependent RNA polymerization. *Journal of General Virology* **85**, 1077-1093.
- van Kuppeveld, F. J., Galema, J. M. D., Zoll, J., Van den Hurk, P. J. & Melchers, W. J. (1996a). Coxsackie B3 virus 2B contains a cationic amphipathic helix that is required for viral RNA replication. *Journal of Virology* **70**, 3876-3886.

- van Kuppeveld, F. J., Hoenderop, J. G., Smeets, R. L. L., Willems, P. H. M., Dijkman, H. B. P. M., Galama, J. M. D. & Melchers, W. J. (1997a). Coxsackievirus protein 2B modifies the endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO* **16**, 3519-3532.
- van Kuppeveld, F. J., van den Hurk, P. J., Zoll, J., Galama, J. & Melchers, W. J. (1996b). Mutagenesis of the Coxsackie B3 virus protein 2B/2c cleavage site: determinants of processing efficiency and effects on viral replication. *Journal of Virology* **70**, 7632-7640.
- van Kuppeveld, F. J. M., Melchers, W. J. G., Kirkegaard, K. & Doedens, J. R. (1997b). Structure-function analysis of coxsackie B3 virus protein 2B. *Virology* **227**, 111-118.
- van Kuppeveld, F. J. M., van den Hurk, P., van der Vliet, W., Galama, J. M. D. & Melchers, W. J. G. (1997c). Chimeric coxsackie B3 virus genomes that express hybrid coxsackievirus-poliovirus 2B proteins: functional dissection of structural domains involved in RNA replication. *Journal of General Virology* **78**, 1833-1840.
- Vance, L. M., Moscufo, N., Chow, M. & Heinz, B. A. (1997). Poliovirus 2C region functions during encapsidation of viral RNA. *J Virol* **71**, 8759-65.
- Vance, V. & Vaucheret, H. (2001). RNA silencing in plants - defence and counterdefence. *Science* **292**, 2277-2280.
- Vankuppeveld, F. J. M., Galama, J. M. D., Zoll, J. & Melchers, W. J. G. (1995). Genetic-Analysis of a Hydrophobic Domain of Coxsackie B3 Virus Protein 2b - a Moderate Degree of Hydrophobicity Is Required For a Cis-Acting Function in Viral-Rna Synthesis. *Journal of Virology* **69**, 7782-7790.
- Varani, G., Cheong, C. & Tinoco, I., Jr. (1991). Structure of an unusually stable RNA hairpin. *Biochemistry* **30**, 3280-3289.
- Ventoso, I., Macmillan, S. E., Hershey, J. W. & Carrasco, L. (1998). Poliovirus 2A^{pro} cleaves directly the eIF4G subunit eIF4F. *FEBS letters* **435**, 79-83.
- Verlinden, Y., Cuconati, A., Wimmer, E. & Rombaut, B. (2000). Cell-free synthesis of poliovirus: 14S subunits are the key intermediates in the encapsidation of poliovirus RNA. *J Gen Virol* **81**, 2751-4.
- Waggoner, S. & Sarnow, P. (1998). Viral ribonucleoprotein complex formation and nucleolar-cytoplasmic relocalization of nucleolin in poliovirus-infected cells. *J Virol* **72**, 6699-709.
- Walewski, J. L., Keller, T. R., Stump, D. D. & Branch, A. D. (2001). Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA* **7**, 710-721.

- Wang, Q. M., Hockman, M. A., Staschke, K., Johnson, R. B., Case, K. A., Lu, J., Parsons, S., Zhang, F. & Rathnachalam, R. (2002). Oligomerization and Cooperative RNA synthesis Activity of Hepatitis C Virus RNA-dependent RNA polymerase. *Journal of Virology* **76**, 3865-3872.
- Waterhouse, P., Wang, M. B. & Lough, T. (2001). Gene silencing as an adaptive defence against viruses. *Nature* **411**, 824-842.
- Weidman, M., Yalamanchili, P., Ng, B., Tsai, W. & Dasgupta, A. (2001). Poliovirus 3C Protease- Mediated Degradation of Transcriptional Activator p53 Requires a cellular activity. *Virology* **291**, 260-271.
- Williams, D. T., Chaudhry, Y., Goodfellow, I. G., Lea, S. & Evans, D. J. (2003). Interaction of Decay Accelerating Factor (DAF) with haemagglutinating human enteroviruses: Utilising variation in primate DAF to map virus binding sites. *Journal of General Virology* **85**, 731-738.
- Wimmer, E. (1982). Genome linked proteins of viruses. *Cell* **28**, 199-201.
- Wimmer, E., Hellen, C. U. T. & Cao, X. M. (1993). Genetics of Poliovirus. *Annual Review of Genetics* **27**, 353-436.
- Witwer, C., Rauscher, S., Hofacker, I. L. & Stadler, P. F. (2001). Conserved RNA secondary structures in Picornaviridae genomes. *Nucleic Acids Res* **29**, 5079-89.
- Wood, J., Frederickson, R. M., Fields, S. & Patel, A. H. (2001). Hepatitis C virus 3'X Region interacts with Human Ribosomal proteins. *Journal of Virology* **75**, 1348-1358.
- Xia, T., McDowell, J. A. & Turner, D. H. (1997). Thermodynamics of nonsymmetric tandem mismatches adjacent to G.C base pairs in RNA. *Biochemistry* **36**, 12486-12487.
- Xiang, W., Cuconati, A., Hope, D., Kirkegaard, K. & Wimmer, E. (1998). Complete protein linkage map of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. *J Virol* **72**, 6732-41.
- Xiang, W., Cuconati, A., Paul, A. V., Cao, X. & Wimmer, E. (1995a). Molecular dissection of the multifunctional poliovirus RNA-binding protein 3AB. *Rna* **1**, 892-904.
- Xiang, W., Harris, K. S., Alexander, L. & Wimmer, E. (1995b). Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. *Journal of Virology* **69**, 3658-3667.
- Yalamanchili, P., Datta, U. & Dasgupta, A. (1997a). Inhibition of host cell transcription by poliovirus: Cleavage of transcription factor CREB by poliovirus-encoded protease 3C(pro). *Journal of Virology* **71**, 1220-1226.

- Yalamanchili, P., Harris, K., Wimmer, E. & Dasgupta, A. (1996). Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. *J Virol* **70**, 2922-9.
- Yalamanchili, P., Weidman, M. K. & Dasgupta, A. (1997b). Cleavage of transcriptional activator Oct-1 by poliovirus encoded protease 3Cpro. *Virology* **239**, 176-185.
- Yang, W. X., Terasaki, T., Shiroki, K., Ohka, S., Aohi, J., Tanabe, S., Nomura, T., Terada, E. & Nomoto, A. (1997). Efficient delivery of circulating poliovirus to the central nervous system independently of poliovirus receptor. *Virology* **229**, 421-428.
- Yang, Y., Rijnbrand, R., McKnight, K. L., Wimmer, E., Paul, A., Martin, A. & Lemon, S. M. (2002). Sequence requirements for viral RNA replication and VPg uridylylation directed by the internal cis-acting replication element (cre) of human rhinovirus type 14. *J Virol* **76**, 7485-94.
- Yi, M. & Lemon, S. M. (2002). Replication of subgenomic hepatitis A virus RNAs expressing firefly luciferase is enhanced by mutations associated with adaptation of virus to growth in cultured cells. *J Virol* **76**, 1171-80.
- You, S., Stump, D., Branch, A. & Rice, C. M. (2004). A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for Hepatitis C virus RNA replication. *Journal of Virology* **78**, 1352-1366.
- Ypma Wong, M. F., Dewalt, P. G., Johnson, V. H., Lamb, J. G. & Semler, B. L. (1988). Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology* **166**, 265-270.
- Yu, S. Y. F. & Lloyd, R. E. (1992). Characterization of the Roles of Conserved Cysteine and Histidine Residues in Poliovirus 2A-Protease. *Virology* **186**, 725-735.
- Zhang, G., Haydon, D. T., Knowles, N. J. & McCauley, J. W. (1999). Molecular evolution of swine vesicular disease virus. *Journal of General Virology* **80**, 639-651.
- Zhong, W., Ferrari, E., Lesburg, C. A., Maag, D., Ghosh, S. K. B., Cameron, C., Lau, J. Y. N. & Hong, Z. (2000). Template/Primer Requirements and Single Nucleotide Incorporation by Hepatitis C Virus Nonstructural Protein 5B Polymerase. *Journal of Virology* **74**, 9134-9143.
- Zimring, J. C., Goodbourn, S. & Offermann, M. K. (1998). Human herpesvirus 8 encodes an interferon regulatory factor (IRF) homolog that represses IRF-1 mediated transcription. *Journal of Virology* **72**, 701-707.
- Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* **244**, 48-52.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acid research* **31**, 3406-3415.

Appendix 1

Species C sequences (n = 50)

Genbank accession number	Comments
AF405669	Human poliovirus 1 isolate HAI00003, complete genome.
AF405682	Human poliovirus 1 isolate DOR00041C1, complete genome.
AF405690	Human poliovirus 1 isolate DOR00013, complete genome.
AF462418	Human poliovirus 1 99/056-252-14, complete genome.
AF462419	Human poliovirus 1 RUS-1161-96-001, complete genome.
AY056701	Human poliovirus 1 strain 3788ALB96 polyprotein gene, partial cds.
AY056702	Human poliovirus 1 strain 3914ALB96 polyprotein gene, partial cds.
AY056703	Human poliovirus 1 strain 4019ALB96 polyprotein gene, partial cds.
AY184219	Human poliovirus 1 strain Sabin 1, complete genome.
AY278553	Human poliovirus 1 isolate PIW/Bar65 (19276), complete genome.
HPO132960/ AJ132960	Human poliovirus 1, complete genome. Isolated from faeces of an immunodeficient patient
HPO132961/	Human poliovirus 1, complete genome. Isolated from an immunodeficient patient
HPO416942/ AJ416942	Human poliovirus 1 genomic RNA for polyprotein, strain CHAT 10A-11.
POLIO1A/ V01148	Genome of Human poliovirus type 1 (Mahoney strain). (One of two versions.).
NC_002058	Human poliovirus 1, complete genome

AJ430385	Human poliovirus 1 genomic RNA for polyprotein gene, strain Cox.
POLIOS1 V01150 J02282 J02285 J02286 V01133	Human poliovirus strain Sabin 1 complete genome
AF448782	Human poliovirus 2 strain EGY88-074, complete genome.
AF448783	Human poliovirus 2 strain EGY93-034, complete genome.
AY177685	Human poliovirus 2, complete genome.
AY184220	Human poliovirus 2 strain Sabin 2, complete genome.
AF405666	Human poliovirus 1 isolate HAI01007, complete genome.
AY238473	Human poliovirus 2 strain MEF-1 polyprotein gene, complete cds.
AY278549	Human poliovirus 2 isolate P2S/Mog65-3 (20120), complete genome.
AY278550	Human poliovirus 2 isolate P2S/Mog65-1 (20003), complete genome.
AY278551	Human poliovirus 2 isolate P2S/Mog66-4 (21043), complete genome.
AY278552	Human poliovirus 2 isolate P2S/Mog65-2 (20077), complete genome
PIPOLS2/ X00595	Poliovirus type 2 genome (strain Sabin 2 (P712, Ch, 2ab).
POL2CG1/	Human poliovirus 2 genomic RNA, complete sequence.
POL2LAN/ M12197	Poliovirus type 2 (Lansing strain), complete genome.
AY184221	Human poliovirus 3 strain Sabin 3, complete genome
HPO293918/ AJ293918	Human poliovirus type 3 complete genome, live-attenuated strain
PIP03XX/ X04468	Poliovirus type 3 strain 23127 complete genome.
PIPO3119/ X01076	Poliovirus type 3 complete sequence (strain P3/119).

PIPOLS3/ X00596	Poliovirus type 3 mRNA (vaccine strain Sabin 3 (Leon 12a1b).
PIPOX3/ X00925/K00043	Poliovirus type 3 leon 12 a1b sequence (P3/Leon 12 a1b).
POL3L37/ K01392	Poliovirus P3/Leon/37 (type 3), complete genome.
AF546702	Human coxsackievirus A21 strain Kuykendall, complete genome.
NC_001428	Coxsackievirus A21, complete genome.
CXA24CG	D90457 Human coxsackievirus A24 DNA, complete genome.
AF499636	Human coxsackievirus A11 strain Belgium-1, complete genome.
AF499637	Human coxsackievirus A13 strain Flores, complete genome.
AF499638	Human coxsackievirus A15 strain G9, complete genome.
AF499639	Human coxsackievirus A17 strain G12, complete genome.
AF499640	Human coxsackievirus A18 strain G13, complete genome.
AF499641	Human coxsackievirus A19 strain 8663, complete genome.
AF499642	Human coxsackievirus A20 strain IH35, complete genome.
AF499643	Human coxsackievirus A22 strain Chulman, complete genome.
AF499635	Human coxsackievirus A1 strain Tompkins, complete genome.

Appendix 2

Species B (n = 68)

Genbank accession number	Comments
AY302560	Enterovirus 69 strain Toluca-1 complete genome
AY302559	Human echovirus 7 strain Wallace complete genome.
AY302558	Human echovirus 6 strain D'Amori complete genome.
AY302557	Human echovirus 4 strain Pesacek complete genome.
AY302556	Human echovirus 33 strain Toluca-3 complete genome.
AY302555	Human echovirus 32 strain PR-10 complete genome.
AY302553	Human echovirus 31 strain Caldwell complete genome.
AY302552	Human echovirus 3 strain Morrisey complete genome.
AY302551	Human echovirus 29 strain JV-10 complete genome.
AY302550	Human echovirus 27 strain Bacon complete genome.
AY302549	Human echovirus 26 strain Coronel complete genome.
AY302548	Human echovirus 25 strain JV-4 complete genome.
AY302547	Human echovirus 24 strain DeCamp complete genome.

AY302546	Human echovirus 21 strain Farina complete genome.
AY302545	Human echovirus 20 strain JV-1 complete genome.
AY302544	Human echovirus 2 strain Cornelis complete genome.
AY302542	Human echovirus 19 strain Burke complete genome.
AY302541	Human echovirus 16 strain Harrington complete genome.
AY302540	Human echovirus 15 strain CH 96-51 complete genome.
AY302539	Human echovirus 14 strain Tow complete genome.
AY186746	Human echovirus 13 strain Del Carmen complete genome.
AY186747	Human coxsackievirus B1
AY186748	Human coxsackievirus B1
NC_001472	Human coxsackievirus B1
AF085363	Coxsackievirus B1, complete genome.
NC_000881	Coxsackievirus B2 strain Ohio-1 polyprotein gene, complete cds.
AF231763	Coxsackievirus B2, complete genome.
AF231764	Coxsackievirus B3 strain 31-1-93, complete genome.
AF231765	Coxsackievirus B3 strain P, complete genome.
CXA3CG	Coxsackievirus B3 strain PD, complete genome.

M33854	Coxsackievirus B3 (CVB3) complete genome
CXA3G M16572.	Coxsackievirus B3, complete genome
CXAB3CG M88483	Coxsackievirus B3 mRNA, complete genome.
CXU57056 U57056	Coxsackievirus B3 Woodruff variant, complete genome.
AF241359 AF241360	Enterovirus CA55-1988, complete genome.
AF311939	Human coxsackievirus B4 strain E2 variant, complete genome.
NC_001360	Coxsackievirus B4, complete genome.
PICOXB4 X05690 D00149	Coxsackievirus B4 complete genome
S76772	polyprotein [Coxsackie B4 virus CB4, host=mice, E2, originally derived from Edwards CB4 h
AF317694	Human echovirus 18 strain Metcalf polyprotein gene, complete cds
AY167107	Human echovirus 19 strain K/542/81 polyprotein gene,
AF311938	Human echovirus 30 strain Bastianni, complete genome.
HEC295172 AJ295172	Human echovirus 30 genomic RNA for partial polyprotein gene, isolate ITA97-002.
NC_000873	Echovirus 30, complete genome.
AF504533	Human enterovirus 73 isolate 2776-82 polyprotein (pol) gene, partial cds
NC_002347	Coxsackievirus A9, complete genome

AF114383	Coxsackievirus B5 strain Faulkner, complete genome.
NC_001342	Coxsackievirus B5, complete genome.
AF105342	Coxsackievirus B6 strain Schmitt, complete genome.
AF114384	Coxsackievirus B6 strain Schmitt, complete genome.
NC_002003	Coxsackievirus B6, complete genome.
NC_003986	Human echovirus 1, complete genome
NC_002601	Echovirus 5, complete genome.
NC_001657	Echovirus 6, complete genome.
AY036578	Human echovirus 7 strain UMMC polyprotein gene, complete cds.
AY036579	Human echovirus 7 strain Wallace polyprotein gene, complete cds.
AF524866	Human echovirus 9 strain Barty, complete genome.
AF524867	Human echovirus 9 isolate DM, complete genome
ECHOV9XX X92886	Echovirus 9 (strain Barty), complete genome.
NC_001656	Echovirus 9 complete genome.
AY167103 AF446118 AF447459 AF447474	Human echovirus 11 strain Hun/90
AY167104 AF446123 AF447462 AF447473	Human echovirus 11 strain Kar/87
AY167105 AF446122 AF447460 AF447477	Human echovirus 11 strain Kust/86
AY167106 AF446113 AF447467 AF447488	Human echovirus 11

E11276224 AJ276224	Human echovirus 11 genomic RNA for polyprotein.
EV11VPCD X80059	Echo Virus 11 genomic DNA.
EC12TCG X77708	Echovirus type 12, prototype Travis complete RNA genome.
EC12TCGWT X79047	Echovirus type 12, prototype Travis wild type genome.

Appendix 3

G A A A A C C T A C A A G G G C A T A G A T T T G G A C C A	PV3wt
G A A A A C C T A C A A G G G C A T A G A T T T G G A C C A	3DNheI
G A A A A C C T A C A A G G G C A T A G A T T T G G A C C A	3DNdeI
G A A A A C T A C A A G G G A T A G A T T T G G A C C A	3DBglII
G A A A A C C T A C A A G G G C A T A G A T T T G G A C C A	3DApoI
G A A A A C C T A C A A G G G C A T A G A T T T G G A C C A	3DNsiI
C T T A A A A A T G A T T G C C T A T G G T G A C G A T G T	PV3wt
C T T A A A A A T G A T T G C C T A T G G G A C G A T G T	3DNheI
C T T A A A A A T G A T T G C T A T G G G A G A T G T	3DNdeI
C T T A A A A A T G A T T G C T A T G G T G A C G A T G T	3DBglII
C T T A A A A A T G A T T G C C T A T G G T G A C G A T G T	3DApoI
C T T A A A A A T G A T T G C C T A T G G T G A C G A T G T	3DNsiI
A A T A G C T T C C T A T C C C C A T G A G G T T G A C G C	PV3wt
T A T G C T C T A T C C C C A T G A G G T T G A C G C	3DNheI
T A T A G C T T C C T A T C C C C A T G A G G T T G A C G C	3DNdeI
A A T A G C T T C C T A T C C C C A T G A G G T T G A C G C	3DBglII
A A T A G C T T C C T A T C C C C A T G A G G T T G A C G C	3DApoI
A A T A G C T T C C T A T C C C C A G A G G T G A G C	3DNsiI
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	PV3wt
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	3DNheI
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	3DNdeI
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	3DBglII
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	3DApoI
T A G T C T C C T A G C C C A A T C A G G A A A A G A C T A	3DNsiI
T G G A C T A A C C A T G A C T C C G G C A G A T A A A T C	PV3wt
T G G A C T A A C C A T G A C T C C G G C A G A T A A A T C	3DNheI
T G G A C T A A C C A T G A C T C C G G C A G A T A A A T C	3DNdeI
T G G A C T A A C C A T G A C T C C G G C A G A T A A A T C	3DBglII
T G G A C T A A C C T G A C T T G G C A A T A A A T T	3DApoI
T G G A C T A A C C A T G A C T C C G G C A G A T A A A T C	3DNsiI
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	PV3wt
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	3DNheI
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	3DNdeI
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	3DBglII
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	3DApoI
T G C C A C T T T T G A G A C A G T C A C A T G G G A G A A	3DNsiI
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	PV3wt
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	3DNheI
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	3DNdeI
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	3DBglII
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	3DApoI
T G T A A C T T T C T T G A A A A G A T T C T T C A G A G C	3DNsiI
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	PV3wt
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	3DNheI
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	3DNdeI
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	3DBglII
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	3DApoI
A G A T G A G A A A T A C C C C T T C C T C A T A C A T C C	3DNsiI
A G	PV3wt
A G	3DNheI
A G	3DNdeI
A G	3DBglII
A G	3DApoI
A G	3DNsiI

The sequence shown encompasses nucleotides 6907 to 7148 of the full-length PV3/Leon/37 genome (PV3wt). The mutations that were introduced in order to destabilise structure I (3DNhe I, 3DNde I and 3DBgl II) and structure II (3DApo I and 3DNsi I) are highlighted in blue.