

# HIV-1 RNA Dimerisation.

Submitted for the degree of PhD.

June 1996.

Malcolm Haddrick.

Department of Microbiology and Immunology  
University of Leicester.

UMI Number: U085658

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.



UMI U085658

Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



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



X753573279

*"Experience is a grindstone and  
it is lucky for us if we get  
brightened by it - not ground."*

♪  
Josh Billings.

UNIVERSITY OF LEICESTER

STATEMENT OF ORIGINALITY

The accompanying thesis submitted for the degree of PhD entitled

**HIV-1 RNA Dimerisation**

is based on work conducted by the author in the Department of Microbiology and Immunology at the University of Leicester mainly during the period between October 1992 and September 1995.

All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed.......... Date.....28/6/96.....

## **Abstract.**

Genomic RNA isolated from retroviral particles is a dimer composed of two identical strands. A region called the dimer linkage site close to the 5' end of the RNA may be involved in forming the dimer. Several models for the formation of the HIV-1 genomic RNA dimer have been proposed. In the "kissing loop" model, dimerisation results from base pairing between homologous sequences in an RNA stem loop. In the guanine tetrad model interstrand guanine contacts form the dimer. Mutations have been made preventing the dimerisation of subgenomic RNAs *in vitro* by these mechanisms. To prevent the "kissing loop" dimer forming the complementary loop sequence 711GCGCGC716 was changed to 711AAACGC716. To prevent the guanine tetrad dimer forming residue G819 was changed to U. These mutations were introduced into a clone of HIV-1<sub>NL4-3</sub> separately and collectively. All three clones produced infectious virions. Dimeric RNA with similar thermal stabilities was isolated from viruses containing either the single or the double mutations. The results suggest that sequences involved in forming a guanine tetrad are not important for HIV-1 RNA dimerisation. In contrast sequences involved in forming a kissing loop complex are not absolutely required, but are important in forming a stable HIV-1 RNA dimer.

## **Publications.**

Haddrick M., Lear A., Cann A. J., and Heaphy S. (1996). Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J. Mol. Biol.* 259: 58-68.

Haddrick M., Liu Z-H., Lau A., Heaphy S., and Cann A. J. (1996) Morphogenesis and RNA packaging in recombinant HIV VLPs. *Meth. Virol.*, *in press*.

## **Acknowledgements.**

Acknowledgements are due to all the staff of the Department of Microbiology and Immunology at the University of Leicester. Specifically I am grateful to Shaun Heaphy and Alan Cann for their assistance throughout the project, with countless valuable contributions in their own inimitable styles. I also thank the members of Lab 136 past and present for their co-operation and help. I am indebted to Andy Lear for his input into this project, particularly the DMS modification analysis. Thanks are also due to Alan Lau for his co-operation in maintaining the CIII containment laboratory. Finally, but most importantly, I am especially grateful to Shital for her support and encouragement throughout this project.

## Abbreviations.

AIDS	Acquired Immune Deficiency Syndrome
AMPS	Ammonium Persulphate
ASV	Avian Sarcoma Virus
BLV	Bovine Leukemia Virus
CMV	Cytomegalovirus
DEPC	Diethyl Pyrocarbonate
DIS	Dimer Initiation Site/Structure
DLS	Dimer Linkage Site/Structure
DMS	Dimethyl Sulphate
DMSO	Dimethyl sulphoxide
DNase	Deoxyribonuclease
GST	Glutathione-S-Transferase
HIV	Human Immunodeficiency Virus
HTLV	Human T-cell Leukemia Virus
IMS	Industrial Methylated Spirit
kb	kilo base
LP-USE	Long PCR- Unique Site Elimination
LTR	Long Terminal Repeat
MLV	Murine Leukemia Virus
MMTV	Mouse Mammary Tumour Virus
MoMLV	Moloney Murine Leukemia Virus
MPMV	Mason-Pfizer Monkey Virus
MW	Molecular Weight
NMR	Nuclear Magnetic Resonance
PBS	Phosphate Buffered Saline
PBS	Primer Binding Site

PCR	Polymerase Chain Reaction
REV	Reticuloendotheliosis Virus
RNAse	Ribonuclease
RSV	Rous Sarcoma Virus
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SD	Splice donor
SDS	Sodium Dodecyl Sulphate
SFV	Simian Foamy Virus
SIV	Simian Immunodeficiency Virus
SNV	Spleen Necrosis Virus
TEMED	N,N,N',N'-tetramethylethylenediamine
VLP	Virus Like Particle

## **List of Figures.**

### **Chapter 1.**

Figure 1.1	Schematic Course of HIV-1 Infection.	6
1.2	Schematic Representation of the HIV-1 virion particle.	8
1.3	The HIV-1 Genome.	10
1.4	Life Cycle of HIV-1.	15
1.5	HIV-1 RNA Splicing.	18
1.6	Features of Retroviral RNA.	20
1.7	Electronmicrograph of RSV Genomic RNA.	23
1.8	Forced Copy Choice Mechanism for Retroviral Recombination.	25
1.9	Guanine Quartet Structures.	38
1.10	Guanine Tetrads in RNA Dimerisation.	40
1.11	Dimerisation Competent HIV-1 RNAs.	43
1.12	Secondary Structure for a 44 Nucleotide RNA.	46
1.13	A-U Base Pairing Mechanism of RNA Dimerisation.	47
1.14	The Kissing Loop Model of RNA Dimerisation.	48
1.15	Kissing Loop Stacking Interaction.	50

### **Chapter 2.**

Figure 2.1	RT-PCR Mutagenesis.	69
2.2	LP-USE Mutagenesis.	71
2.3	Random Primed DNA Probe Profiles.	81
2.4	Detection Limit of Northern Analysis.	82
2.5	The Virus Like Particle Expression Construct pBCCX CSF X4.	90
2.6	The Infectious Molecular Clone of HIV-1 <sub>NL4-3</sub> , pUCΔNL.	91

### **Chapter 3.**

Figure 3.1	Schematic Representation of the HIV-1 RNAs under Study.	94
3.2	RNAs that Dimerised by the Formation of a Guanine Tetrad.	96
3.3	Dimerisation Time Course of RNAs 1, 2, 12 and 13.	97
3.4	DMS Footprinting Analysis of RNA 1 (817GGGG820).	98
3.5	DMS Footprinting Analysis of RNA 1 (732GGG734).	99
3.6	DMS Footprinting Analysis of RNA 12.	101
3.7	Cation Dependent Thermal Stability of RNA 11.	102
3.8	RNA 2 Forms A Two-stranded Dimer.	104
3.9	Cation Dependent Thermal Stability of RNA 2.	106
3.10	DMS Footprinting Analysis of RNA 2.	108
3.11	Possible G-Tetrad Conformations for RNA 2 and DLS 112.	110
3.12	Gag Coding Sequence.	112
3.13	Dimerisation Time Course for RNAs 2 and 3.	113
3.14	RNAs Containing Sequences Upstream of Position 708.	115
3.15	Dimerisation Time Course for RNA 6.	115
3.16	Cation Dependent Thermal Stability of RNA 6.	116
3.17	Dimerisation of RNAs 4, 5, 6 and 7.	118
3.18	Sequence Alignment of RNAs 1, 2, 4, 5, 6 and 7.	119
3.19	StemII and LoopII RNA Sequences.	120
3.20	Tetrad and Kissing Loop RNAs.	128
3.21	Nucleocapsid Peptide ( $\Delta$ NCp7):RNA Interaction.	129
3.22	Effect of $\Delta$ NCp7 on the Dimerisation of RNA 8.	130
3.23	$\Delta$ NCp7 Peptide Assisted Dimerisation of RNA 6.	131
3.24	Specificity of $\Delta$ NCp7:RNA Binding.	133
3.25	Effect of Organic Extraction on RNA 6 Dimerisation.	135
3.26	Organic Extraction Procedures Dimerised RNA 6.	136
3.27	Phenol Induced Heterodimerisation.	139
3.28	Dimerisation of RRE 33 RNA.	140

## **Chapter 4.**

Figure 4.1	The CMV:HIV DNA Junction in the pBCCX CSF X4 Vector.	146
4.2	Restriction Digestion Analysis of pBCCX CSF X4.	147
4.3	Electroporation of pBCCX CSF X4 into COS Cells Produces p24.	148
4.4	Reverse Transcriptase Activity from Electroporated COS Cells.	149
4.5	Electronmicrograph of HIV-1 Virus Like Particles (VLPs).	150
4.6	Sucrose Density Gradient Analysis of VLPs.	152
4.7	HIV-1 gag RNA was Detected with the VLPs.	154
4.8	HIV-1 pol and env RNA was Detected with the VLPs.	155
4.9	VLPs were Disrupted by Treatment with SDS.	157
4.10	Specificity of RNA Packaging into VLPs.	159
4.11	Northern Analysis of VLP RNA.	161
4.12	Northern Analysis of VLP RNA - DNase Digestion.	163
4.13	Northern Analysis of VLP RNA - RNase Digestion.	164
4.14	Restriction Digestion Analysis of pUCΔNL.	168
4.15	Mutagenesis Summary for Viral DNA Clones.	170
4.16	Production of Kissing Loop Mutant DNA Clones.	172
4.17	p24 Production from Cells Electroporated with Viral Clones.	174
4.18	Infection of C8166 Cells with Mutated HIV-1 Viruses.	175
4.19	RT-PCR Analysis of Mutated Viruses.	177
4.20	Schematic Representation of the RT-PCR Product.	178
4.21	BssHIII Restriction Digestion Analysis of the RT-PCR Product.	179
4.22	Sequencing of Recovered Viruses ex Culture.	180
4.23	Northern Analysis of Virion RNA.	183
4.24	Thermal Stability of HIV-1 <sub>NL4-3</sub> Genomic RNA.	186
4.25	Thermal Stability of Mutant HIV-1 Genomic RNA Dimers.	187

## **Chapter 5.**

Figure 5.1	Conservation of Tetrad and Kissing Loop Sequences.	191
5.2	Autocomplementary Kissing Loop Sequences.	192
5.3	Tetrad <sub>817</sub> GGGG <sub>821</sub> RNA Sequences.	193
5.4	Possible Kissing Loop Mutant RNA Dimers.	201

## **Tables.**

Table 1	Classification of Retroviruses.	2
Table 2	Construction of HIV-1 RNAs for Dimerisation analysis.	88

## Contents.

Title page.	i
Statement of Originality.	ii
Abstract and Publications.	iii
Acknowledgements.	iv
Abbreviations.	v
List of Figures.	vii
List of Tables.	x
Contents.	xi
<b><u>Chapter 1 Introduction.</u></b>	<b>1</b>
1.1 Retroviruses.	1
1.2 Human Immunodeficiency Virus Type 1.	3
1.2.1 Discovery of HIV-1.	3
1.2.2 HIV-1 Pathogenesis.	4
Clinical Aspects of HIV-1 Infection.	7
1.3 Molecular Biology of HIV-1.	8
1.3.1 Structure of the Virion.	8
1.3.2 The HIV-1 Genome.	9
1.3.3 Life Cycle of HIV-1.	14
1.4 Retroviral Genomic RNA.	20
1.4.1 Features of Retroviral RNA.	20
1.4.2 Dimerisation of Retroviral RNA.	21
Biochemical Analysis.	21
Electronmicroscopy.	22
1.4.3 Role of RNA Dimerisation in the Viral Life Cycle.	24
Reverse Transcription.	24
Translation.	26
Packaging.	27

1.5 Mechanisms of Retroviral RNA Dimerisation.	30
1.5.1 <i>In vitro</i> RNA Dimerisation.	31
1.5.2 Effect of Nucleocapsid Protein on RNA Dimerisation.	32
1.5.3 Relevance of <i>in vitro</i> RNA Dimerisation Studies.	33
1.5.4 <i>In vivo</i> RNA Dimerisation.	34
1.6 Dimerisation of HIV-1 Genomic RNA.	37
1.6.1 The Guanine Tetrad Model.	38
1.6.2 The A-U Base Pairing Model.	45
1.6.3 The Kissing Loop Model.	48
1.7 Project Outline.	52
<b><u>Chapter 2 Materials and Methods.</u></b>	<b>54</b>
2.1 Oligonucleotides.	54
2.2 <i>E. coli</i> Strains and Cell Lines.	55
2.2.1 Growth Conditions.	56
<i>E. coli</i> culture.	56
Cell line culture.	57
2.2.2 Counting Cells.	57
2.2.3 Storing Cells.	57
2.3 DNA Techniques.	58
2.3.1 Agarose Gel Electrophoresis.	58
2.3.2 Restriction Digestion.	58
2.3.3 Phenol/Phenol Chloroform Extraction.	59
2.3.4 Ethanol Precipitation.	60
2.3.5 Gel Purification of DNA.	60
2.3.6 Ligation.	60
2.3.7 Preparation of Transformation Competent <i>E. coli</i> .	61
2.3.8 Transformation of <i>E. coli</i> .	61

2.3.9	Preparation of Plasmid DNA from <i>E. coli</i> .	61
	Alkaline lysis and PEG precipitation.	62
	Caesium chloride procedure.	63
	Miniprep procedure.	64
2.3.10	Polymerase Chain Reaction.	65
	Estimation of primer concentration.	65
	Polymerase chain reaction.	65
2.3.11	Mutagenesis.	67
	PCR mutagenesis.	67
	RT-PCR mutagenesis.	67
	LP-USE mutagenesis.	70
2.3.12	DNA Sequencing.	70
2.4	RNA Work.	73
2.4.1	Precautions for Working with RNA.	73
2.4.2	<i>In vitro</i> RNA Transcription.	73
2.4.3	<i>In vitro</i> RNA Dimerisation.	74
	Buffer I (without peptide).	74
	Buffer II (with peptide).	74
2.4.4	Thermal Stability of RNA Dimers.	74
2.4.5	Radiolabelling of RNA.	75
2.4.6	Dimethyl Sulphate Modification Protection of RNA.	75
	The DMS reaction.	75
	Aniline cleavage.	76
2.4.7	RT-PCR.	76
2.4.8	RNA Extraction.	77
	RNA Extraction from cells and VLPs with RNazol.	77
	Extraction and Preparation of Virion RNA.	78

2.4.9 Northern Analysis of Viral RNA.	79
Gel electrophoresis, transfer and hybridisation.	79
Random-primed DNA probes.	80
Optimisation of Northern analysis.	81
2.5 Virus/VLP Techniques.	82
2.5.1 Electroporation of Cells.	82
2.5.2 Infection of C8166 Cells with HIV-1 Viruses.	83
2.5.3 Antigen Capture p24 ELISA.	83
2.5.4 Reverse Transcriptase Assay.	85
2.5.5 VLP Concentration with Centriprep Concentrators.	86
2.5.6 Sucrose Gradient Fractionation of VLPs.	87
2.5.7 Nuclease Digestion of VLP Associated Nucleic Acid	87
DNase treatment.	87
RNase treatment.	87
2.6 Vector Construction.	88
2.6.1 Construction of Plasmids to Make RNA.	88
2.6.2 The VLP Construct, pBCCX CSF X4.	89
2.6.3 The HIV-1 <sub>NL4-3</sub> Vector, pUCANL.	91
<b><u>Chapter 3 <i>In vitro</i> RNA Dimerisation.</u></b>	93
3.1 Introduction.	93
3.2 RNA Dimerisation by the Formation of a Guanine Tetrad.	95
3.3 Identification of the Guanines Involved in RNA 1 Dimerisation.	96
3.4 An Alternative Guanine Tetrad Formed in the RNA 12 Dimer.	100
3.5 RNA 11 Dimers Contained a Guanine Tetrad.	102
3.6 RNA 2 Dimerised by Dual Guanine Tetrad Formation.	103
3.6.1 RNA 2 Formed a Two-stranded RNA Dimer.	103
3.6.2 Cation Dependent Thermal Stability of RNA 2 Dimers.	105
3.6.3 Guanine Tracts Responsible for RNA 2 Dimerisation.	106

3.7 RNA Dimers can be Composed of Several Guanine Tetrads.	109
3.8 Mutagenesis of One Guanine Residue Abolished Dimerisation.	111
3.8.1 The Dimerisation of RNA 3.	112
3.9 5' Extended RNAs Failed to Dimerise by a Guanine Tetrad.	114
3.9.1 The Dimerisation of RNA 6.	114
3.9.2 The Dimerisation of RNA 8.	117
3.9.3 The Dimerisation of RNA 4.	118
3.10 Sequences Required for the Alternative Dimerisation Mechanism.	119
3.11 The Kissing Loop Model of HIV-1 RNA Dimer Formation.	121
3.12 Kissing Loop Mutagenesis abolished RNA 4 and 6 Dimerisation.	122
3.13 Summary.	123
3.14 Peptide Assisted Dimerisation of HIV-1 RNAs.	125
3.15 Introduction.	126
3.16 The Peptide $\Delta$ NCp7.	126
3.17 Dimerisation of RNAs 1, 2, 6, 8, 10, 11 and 13 with $\Delta$ NCp7 peptide.	127
3.18 Interaction of the Peptide with Kissing Loop RNAs 6 and 8.	130
3.19 Specificity of the Peptide:RNA interaction.	132
3.20 The Effect of Organic Extraction on RNA Multimerisation.	134
3.21 Phenol Extraction Facilitated Kissing Loop Dimer Formation.	137
3.21.1 Phenol Induced Heterodimerisation of RNAs 4 and 6.	138
3.21.2 Phenol Extraction Failed to Dimerise RRE 33 RNA.	140
3.22 Mechanism of Action of Phenol on RNA Dimerisation.	141
3.23 Summary.	142
<b><u>Chapter 4 <i>In vivo</i> RNA Dimerisation.</u></b>	144
4.1 Introduction.	144
4.2 Virus Like Particles (VLPs).	145
4.3 VLP Production from Electroporated COS Cells.	148

4.4 Physical Analysis of VLPs.	149
4.4.1. Electronmicroscopy.	150
4.4.2. Sucrose Gradient Density Centrifugation.	151
4.5 RNA Content of the VLPs.	153
4.6 RT-PCR Analysis of VLP RNA.	153
4.6.1. VLPs Contain HIV RNA.	153
4.6.2. Detergent Treatment of VLPs Releases HIV RNA.	156
4.6.3. Specificity of HIV RNA Packaging.	158
4.7 Northern Analysis of VLP RNA.	160
4.7.1. Analysis of Denatured VLP Samples.	161
4.7.2. DNase Treatment of VLP RNA.	162
4.7.3. RNase Treatment of VLP RNA.	164
4.8 VLP Summary.	165
4.9 Testing <i>in vitro</i> Models of Dimerisation in HIV NL4-3 Viruses.	167
4.10 Mutagenesis of HIV-1 <sub>NL4-3</sub> .	167
4.10.1 The HIV-1 <sub>NL4-3</sub> Molecular Clone pUCΔNL.	167
4.10.2 Mutagenesis of pUCΔNL.	169
4.11 Electroporation of Mutant DNA Clones.	173
4.12 Infections Using Mutant HIV-1 Viruses.	174
4.13 Confirmation of Viral Sequences.	176
4.13.1 RT-PCR Analysis of Mutant Viruses.	176
4.13.2 Restriction Digestion of RT-PCR Products.	177
4.14.3 Sequencing of Mutant Viruses.	179
4.14 Virion RNA Analysis.	181
4.14.1 Northern Analysis.	182
4.14.2 Effect of Phenol on Virion RNA Dimerisation.	184
4.14.3 Thermal Stability of Viral RNA Dimers.	185
4.15 Mutant Viruses Summary.	188

<b><u>Chapter 5 Discussion.</u></b>	189
5.1 Summary of <i>In vitro</i> RNA Dimerisation.	189
5.2 Conservation of Tetrad and Kissing Loop Sequences.	190
5.3 Specificity of <i>In vitro</i> RNA Dimerisation.	194
5.4 <i>In vivo</i> RNA Dimerisation of Mutant HIV-1 Viruses.	196
5.5 Nature of the 5' Dimer Linkage Site.	204
5.6 Future Experiments.	206
<b>References.</b>	207

## Chapter 1: Introduction.

### 1.1. Retroviruses.

Retroviruses have been classified into three main families, based on pathogenicity (Table 1). The first retrovirus to be discovered was Rous sarcoma virus in 1911, infection by this virus lead to the development of a cancer or sarcoma in chickens (Rous, 1911). Studies on in-bred mice lead to the discovery of murine oncoretroviruses, eg. Murine Leukemia Virus (MLV, Furth et al., 1933). Transformation of cells is not the only consequence of a retroviral infection. Lentiviruses cause long latency disorders eg. encephalopathy and lung disease in sheep infected with visna virus. Immunodeficiency can also result from lentiviral infection. Recently isolated human retroviruses HTLV-1 (Poiesz et al., 1981), and HIV-1 (Barre-Sinoussi et al., 1983; Montangier et al., 1984; Gallo et al., 1984; Popovic et al., 1984), have become the target for extensive research due to the severity of disease they cause and the number of infected individuals. However not all retroviruses are pathogenic, for example Human Foamy Virus (HFV) infection has not been linked with disease. However HFV may actually be SFV acquired from contact with primates in the few individuals infected (Weiss, 1996).

Retroviruses contain two molecules of genomic RNA arranged as a dimer. The RNA dimer structure is unique to retroviruses and consists of two capped, unspliced and polyadenylated molecules associated together by non-covalent interactions. During retroviral replication, information from each RNA strand of the dimer is reverse transcribed into a double stranded DNA copy, or provirus, by the viral reverse transcriptase enzyme. The discovery of this RNA dependent DNA polymerase activity demonstrated that the flow of genetic information can occur from RNA to DNA (Baltimore, 1970; Temin and Mizutani, 1970). The provirus

integrates into the host cell genome where it may remain dormant or become expressed to generate more virus.

Subfamily	Group	Example	Comment
Oncovirinae	Avian leukosis sarcoma	Rous Sarcoma Virus (RSV)	Transforming, contains <i>src</i>
	Mammalian C-type	Moloney murine leukaemia virus (MoMLV)	Causes T-cell lymphoma
	B-type viruses	Mouse mammary tumour virus (MMTV)	Causes mammary carcinoma
	D-type viruses	Mason-Pfizer Monkey Virus (MPMV)	Unknown pathogenicity
	HTLV-BLV group	Human T-cell leukaemia virus (HTLV-1)	Causes T-cell lymphoma
Lentivirinae	Lentiviruses	Visna virus	Causes lung disease in sheep
Spumavirinae	Foamy viruses	Simian Foamy Virus (SFV)	Cytopathic and Syncytium inducing, in vitro

Table 1. Classification of Retroviruses.

Retrovirus particles are around 100nm in diameter. An inner ribonucleoprotein complex contains the genomic RNA dimer, reverse transcriptase, integrase and protease enzymes (described in section 1.3.1.). This nucleocapsid is surrounded by a gag derived capsid protein shell. The matrix protein constitutes the outermost layer of the viral structural proteins, beneath the host cell derived lipid bilayer. The particle envelope contains the viral glycoproteins.

Particle morphogenesis proceeds in two main ways. For the B and D-type viruses eg. MMTV and MPMV respectively, the complete but immature capsid forms intracellularly and acquires its envelope as the particle leaves the cell. For C-type viruses eg. RSV, assembly begins with the appearance of crescent shaped patches at the cell membrane. Incorporation of genomic RNA and recruitment of the viral glycoproteins occurs as the particle buds out from the cell. Release of the immature virion is followed by maturation and completion of morphogenesis (reviewed in Wills and Craven, 1991). Whether a virus assembles in a B/D or C type fashion is dependent on the gag molecule alone, as gag only constructs produce virus like particles of the correct morphology (eg. Smith et al., 1990). A single point mutation near the N terminal of the MPMV gag protein switched assembly from D type to C type morphogenesis (Rhee and Hunter, 1990)

## **1.2. Human Immunodeficiency Virus Type 1 (HIV-1).**

### **1.2.1. Discovery of the Virus.**

HIV-1 is the cause of AIDS. This virus was first identified in 1983 when a reverse transcriptase containing virus was isolated from an individual with lymphadenopathy syndrome (LAS, Barre-Sinoussi et al., 1983). This virus had some similar properties to the previously identified HTLV-1 which had also been isolated from individuals with AIDS (Gallo et al., 1984). However the reduction in CD4 positive T cell count was not

characteristic of an HTLV infection. Montagnier et al. demonstrated that the virus they called LAV (Lymphadenopathy Associated Virus) grew to high titre in tissue culture and destroyed, not immortalised CD4 positive lymphocytes (Montagnier et al., 1984). Similarly a virus distinct from HTLV-1 was detected in immunosuppressed patients and named HTLV-III (Popovic et al., 1984). Also at this time, Levy et al. recovered a virus ARV (AIDS associated Retrovirus) from AIDS patients which cross reacted with the French LAV strain (Levy et al., 1984).

The LAV, HTLV-III and ARV viruses had properties distinct from HTLV, so the virus which leads to AIDS was designated the Human Immunodeficiency Virus (Coffin et al., 1986). In fact HTLV-III and LAV were later shown to be the same virus, culture contamination had presumably occurred with the LAV virus (Chang et al., 1993). HIV isolates were subsequently obtained from many patients with AIDS, AIDS-related complex (ARC) and from asymptomatic individuals. Later a separate virus HIV-2 was identified in Western Africa (Clavel et al., 1986), both virus types lead to AIDS although the pathogenic course of HIV-2 is longer (Whittle et al., 1992).

### **1.2.2. HIV-1 Pathogenesis.**

CD4 positive T helper lymphocytes are the major target for infection by HIV-1 virus, viral replication occurs to the highest titres in these cells (Klatzmann et al., 1984). Macrophages can be infected although only low level viral replication occurs (Gendelman et al., 1989). Along with other haematopoietic cells, HIV-1 also infects neural tissue, eg. resident macrophages, astrocytes and ganglia cells. The presence of HIV-1 has been detected by PCR or *in situ* techniques in a wide variety of other cells (reviewed by Levy, 1994).

The mechanism of immunopathology due to the virus is unknown (Weiss, 1993). Explanations have been proposed based on T cell destruction following HIV-1 infection either as a direct result of virus infection, by the formation of syncytia, or secondary effects like activation of apoptosis (Stanley and Fauci, 1993; Gougeon and Montagnier, 1993). Immune system dysfunction may also occur if dendritic cells are destroyed by HIV (Patterson and Knight, 1987). Alternatively a shift in the T-cell helper Th1/Th2 balance, observed in other chronic diseases, may lead to a depletion in CD4 positive Th1 cells (Cease and Berzofsky, 1994). However it is certain that the escape of the virus from the immune response is important for viral persistence (McNicholl and McDougal, 1994). Shifting antigenic diversity and a high rate of genetic variation is also an important factor in pathogenesis, eg. point mutations in the nef gene product of SIV has a substantial effects on pathogenesis (Kestler et al., 1991).

Recently the discovery that lymphoid organs are major reservoirs for HIV (Pantaleo et al., 1991) and that viral replication occurs preferentially and continuously in these cells (Pantaleo et al., 1993), reinforced the belief that pathogenesis of disease is linked to replication of the virus. Longitudinal studies showed that the viral load both of free virus and that in infected cells increases gradually over months to years, indicating a dynamic viral disease during clinical latency (Wain-Hobson, 1995). The extent of virus replication and clearance by the host has been revealed by Wei et al. and Ho et al. (Wei et al., 1995; Ho et al., 1995). After treatment of asymptomatic patients with potent antiretroviral drugs, the level of free HIV-1 was reduced and a corresponding increase in circulating CD4 positive T cells was observed. It was estimated that around  $10^8$ - $10^9$  virions were being cleared every day, yet the levels of plasma viremia in infected individuals only reduce gradually over many months. This indicated that the production of HIV-1 must be around  $10^8$  to  $10^9$  per day. Eventually the

immune system is probably no longer able to combat the virus at this level, and the gradual decline to AIDS presumably occurs, figure 1.1.

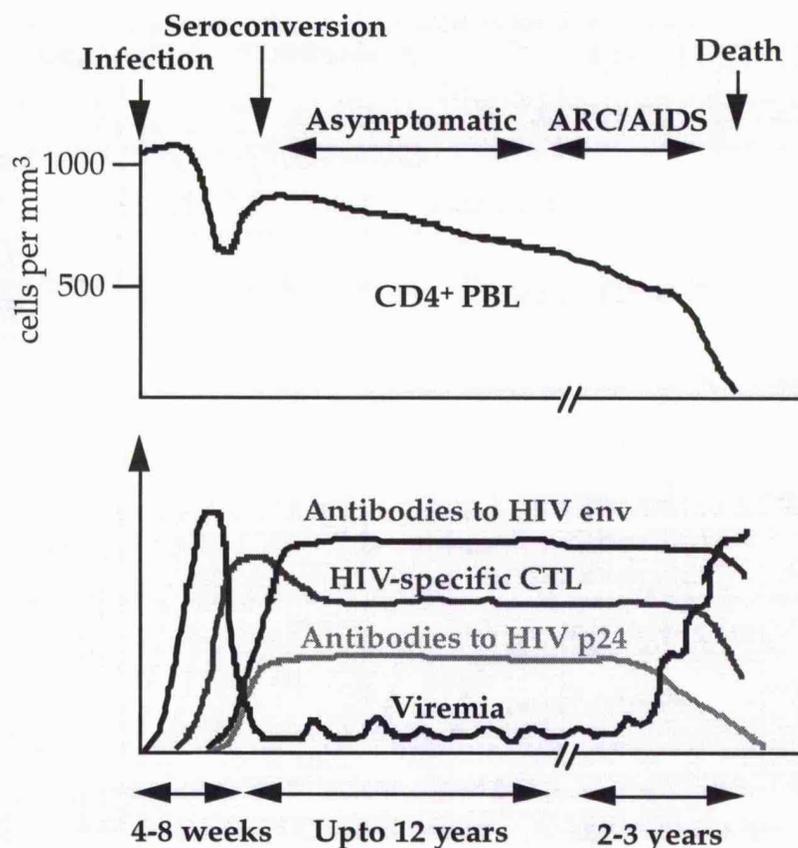


Figure 1.1. Schematic course of HIV-1 infection.

After infection viremia occurs and probably assists dispersion of the virus (Daar et al., 1991; Clark et al., 1991). An immune response begins consisting of HIV-1 specific antibodies (eg, to p24 gag and gp120 env proteins) and HIV-1 specific CD4 and CD8 T cells. This is followed by a clinical latency period, the patient is asymptomatic although the circulating CD4 count falls slowly. However viral replication and viral clearance probably occurs throughout this period (Wei et al., 1995; Ho et al., 1995). As CD4 counts decline clinical complications are observed, along with an increase in viral load (de Wolf et al., 1988). Viral diversity (Hahn et al., 1986) expands

producing complex viral quasispecies (Wain-Hobson, 1992) and altered biological properties of the virus (Groenink et al., 1991) eg. tropism.

#### Clinical Features.

While the molecular mechanisms of HIV pathogenesis may remain unclear, the clinical manifestations of progression to AIDS are well known. Some symptoms are indicated below and are described in detail by (Levy, 1994).

#### Initial Infection.

Following infection by HIV-1 seroconversion usually occurs from 1 to 4 weeks. Typical symptoms include fever, headaches, irritability, depression, nausea, diarrhoea, rashes, vesicular rash and swollen lymph nodes. Clinical recovery usually follows followed by the asymptomatic period which can last from months to years.

#### Symptomatic Infection

A variety of symptoms may present at this stage, eg. chronic fevers, weight loss, oral thrush, and herpes zoster. Once the diagnosis of AIDS has been made survival is often less than one year. Typical symptoms include infection by opportunistic pathogens eg. *Candida albicans*, toxoplasma, and CMV. Kaposi Sarcoma is also observed, and complications can occur in the kidneys (nephropathy), heart and lungs (pneumonia), blood (anemia) and nervous system (encephalitis).

#### Therapy

There is currently no vaccine or cure for AIDS. Drug therapy has had only minor success, the evolution of drug resistant viruses occurs rapidly after administration eg. with AZT (Kellam et al., 1992). Recent drug combination therapies have demonstrated some increase in survival compared to control groups (eg. Delta trial).

## Epidemiology

Current World Health Organisation estimates are a total number of 19.5 million HIV infected individuals, with 2.5 million new infections in 1994. Sub-saharan Africa has been worst hit with about 11 million infected, however the largest increase in infections (1 million in 1993-1995) was in South and South East Asia (Global AIDS News, 1995 No.1).

### 1.3. Molecular Biology of HIV-1.

#### 1.3.1. Structure of the Virion

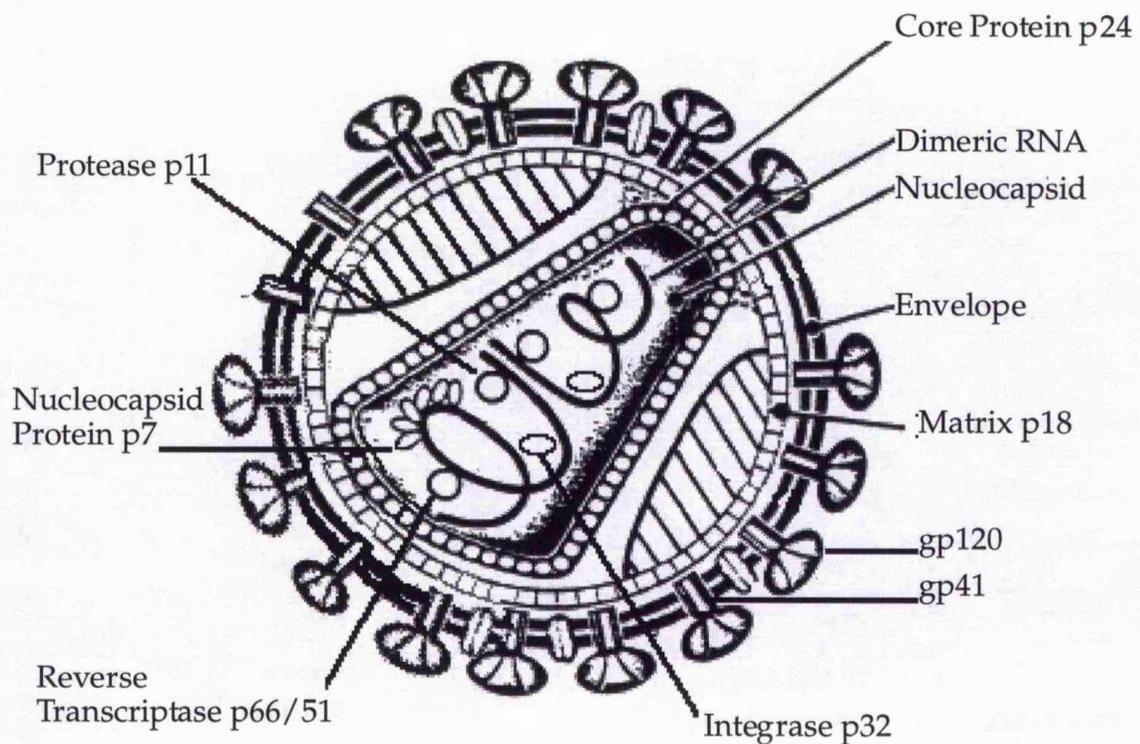


Figure 1.2. Schematic representation of the HIV-1 virion particle (adapted from Gelderblom, 1991).

HIV-1 particles are around 110nm in diameter, they have a condensed conical shaped core consisting of gag derived p24 protein (Gelderblom, 1991). This structure contains the genomic RNA dimer as part of the

nucleocapsid complex. The two 9kB positive sense RNA molecules have an annealed tRNA<sup>Lys,3</sup> replication primer and are associated with nucleocapsid protein (p7), reverse transcriptase (RT p66/51), integrase (IN p32) and protease (p11) enzymes. The capsid is contained within a layer of matrix protein (p18), this is myristoylated and associated with the envelope surrounding the particle. The glycoprotein gp 120 is present on the surface of the virus probably as a trimer or tetramer complex, this is linked to the particle by the gp 41 transmembrane protein. Other molecules have been found within the viroplasm, eg. viral DNA (Lori et al., 1992) and cyclophilin A (Franke et al., 1994; Thali et al., 1994) although their role, if any, is uncertain.

### **1.3.2. The HIV-1 Genome.**

The HIV-1 provirus is flanked by LTR repeat regions and contains the structural proteins gag, pol and env, in an arrangement common to all retroviruses. Additionally HIV-1 codes for the accessory proteins tat, rev, nef, vif, vpr and vpu, figure 1.3.

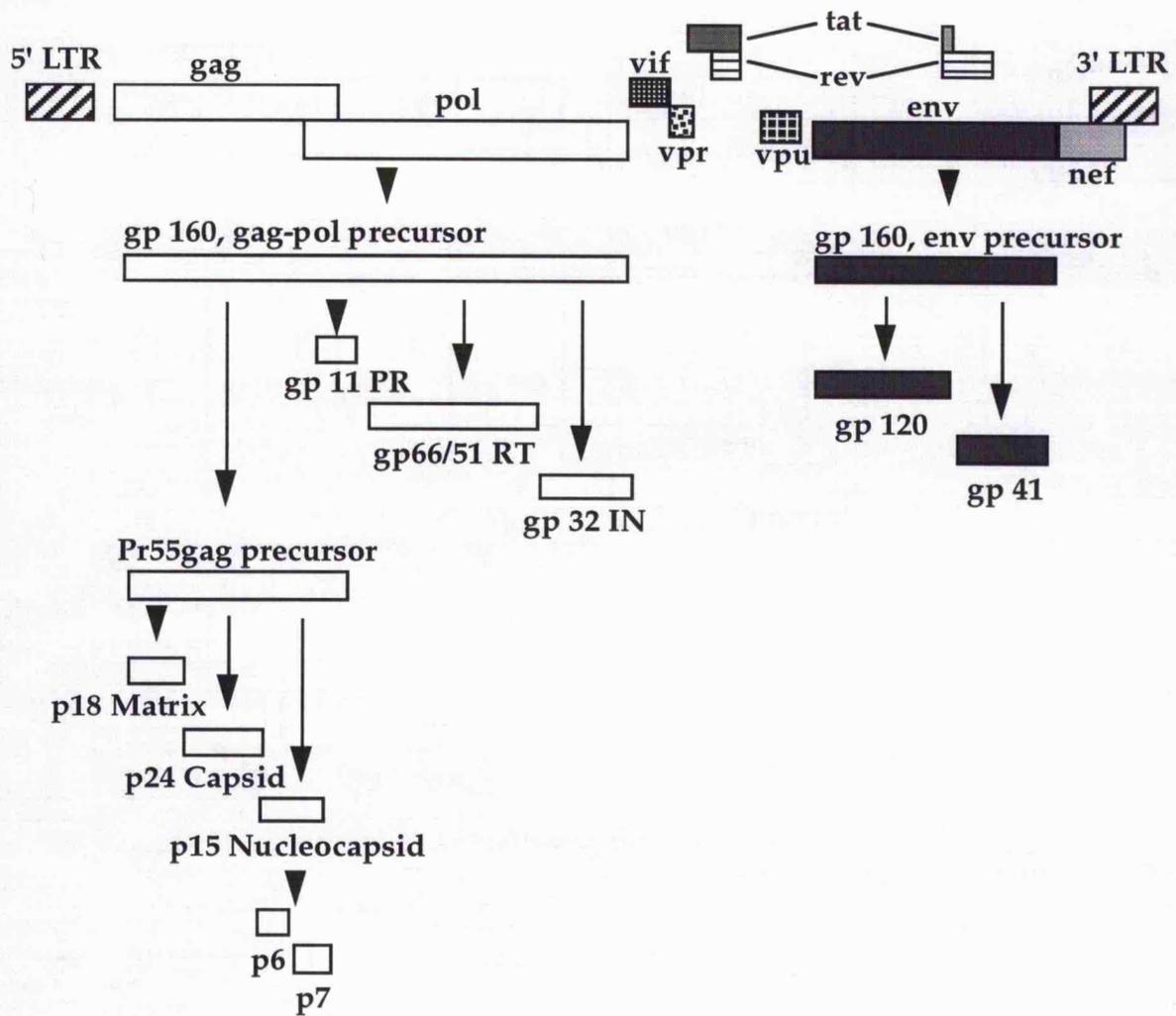


Figure 1.3. The HIV-1 Genome.

### The gag gene.

The gag gene is initially translated into a polyprotein precursor (Pr55<sup>gag</sup>), containing the proteins NH<sub>2</sub>-p18(MA)-p24(CA)-p15(NC)-COOH. Pr55<sup>gag</sup> protein is sufficient to form immature virus like particles when expressed in baculovirus (Gheysen et al., 1989; Overton et al., 1989; Luo et al., 1990), vaccinia (Karacostas et al., 1989; Hu et al., 1990; Shioda and Shibuta 1990), and other expression systems (Mergener et al., 1992). However Pr160<sup>gag-pol</sup> overexpression does not lead to efficient particle production (Park and Morrow, 1991; Karacostas et al., 1993). The individual proteins are released from the polyprotein by cleavage of the precursor with the viral protease

during morphogenesis and assembly (Erickson-Viitanen et al., 1989). The polyprotein gag protein probably has a role in RNA packaging and maturation, demonstrated for RSV genomic RNA (Oertle and Spahr, 1990).

The matrix (MA, p18) protein has several functions. The BLV matrix protein binds to dimeric BLV RNA (Kato et al., 1991). Matrix was essential for the incorporation of HIV-1 env proteins (Yu et al., 1992) and may have a role in the early stages of infection (Yu et al., 1992; Freed et al., 1995). Replacement of the MLV matrix protein by HIV-1 MA allowed efficient infection to occur, although chimeric particle assembly and replication were impaired (Deminie and Emerman, 1994). Matrix is a myristoylated protein, the absence of this signal directed HIV-1 assembly into the endoplasmic reticulum (Facke et al., 1993). However matrix can be deleted and substituted with a myristoylation signal allowing correct morphogenesis at the plasma membrane to occur (Lee and Linial, 1994). Deletions in the matrix domain have also been identified which abrogate particle assembly (Niedrig et al., 1994; Spearmen et al., 1994).

The p24 (CA) protein forms the virion capsid structure. Mutational analysis of this domain identified residues primarily in the C-terminal of the protein that were involved in gag:gag interactions to form the capsid structure (Jowett et al., 1992; Hong and Boulanger 1993; Poblitzki et al., 1993; Chazal et al., 1994; Hockley et al., 1994; Niedrig et al., 1994; Reicin et al., 1995). Gag precursor proteins can form viral capsid structures *in vitro*, although RNA may be required to initiate and organise the structure (Klikova et al., 1995; Campbell et al., 1995). A model based on fullerene like assembly of the gag capsid structure has been proposed (Nermut et al., 1994).

The nucleocapsid (NC) moiety of gag is released from the polyprotein precursor as a p15 protein which is additionally cleaved to produce proteins p7 and p6. This cleavage may require RNA, as demonstrated *in vitro* (Sheng and Erickson-Viitanen, 1994). The nucleocapsid proteins of all retroviruses contain one or two copies of a conserved motif, the Cys-His box (Henderson et al., 1981) reminiscent of a zinc finger like structure (reviewed in Rhodes and Klug, 1993). These proteins contain bound zinc ions in the finger motif and are highly basic. Nucleocapsid protein (p7) binds to and protects the genomic RNA; *in vitro* NCp15 and NCp7 have been reported to facilitate HIV-1 RNA dimerisation (Darlix et al., 1990; Weiss et al., 1992; DeRocquigny et al., 1992, 1993). The nucleocapsid domain is important for the recognition and packaging of genomic RNA into the assembling virion, probably as part of the gag-pol polyprotein (Stewart 1990). The function of the p6 protein from the NC is uncertain, although it may be a tether linking freshly budded particles to the membrane of the infected cell (Gottlinger et al., 1991) and may form a core to envelope link within the virus (Gelderblom, 1991). The p6 protein is also required to incorporate vpr protein into particles (Kondo et al., 1995)

#### The pol gene.

The pol gene is translated from the gag-pol polyprotein precursor (Pr160<sup>gag-pol</sup>). This results from a ribosome frameshift at the gag (NC) to pol (protease) junction into the -1 reading frame. The pol gene contains the viral protease (PR p11), reverse transcriptase (RT p66/51) and integrase (IN p32) enzymes.

The viral protease is an essential enzyme, it is an aspartate protease that acts as a dimer and cleaves the viral polyproteins (Babe et al., 1995). Protease defective viruses have a reduced infectivity and aberrant morphology (Stewart et al., 1990; Kaplan et al., 1993).

The reverse transcriptase enzyme acts as a p66/p51 dimeric complex of reverse transcriptase and RNase H activities (LeGrice and Leitch, 1990). Reverse transcription generates the DNA provirus from the dimeric RNA template, the mechanism is error-prone due to the lack of a 3' to 5' DNA proofreading exonuclease activity by the enzyme (reviewed in Williams and Loeb, 1992).

The final pol gene product is the integrase enzyme (IN p32), which catalyses the non-specific integration of the provirus into the host cell genome. Integrase is present within the virion particle (Gelderblom 1991).

#### The env gene.

The env gene is translated as a gp 160 precursor protein, it is extensively glycosylated and oligomerises in the rough endoplasmic reticulum (Earl et al., 1990). Cleavage by a cellular protease in a lysosomal compartment releases the gp120 and gp41 proteins, which are transported to the cell surface (Wiley et al., 1988; Kozarsky et al., 1989). The gp120 oligomerises probably into trimers or tetramers on the virion particle. A functional gp120 is essential to the infectivity of progeny virus particles.

#### Auxillary gene products.

##### tat, rev and nef.

These are early transcripts following infection. Tat is the transactivator protein, it acts on the viral LTR to increase the initiation of transcription and produce more tat protein (Cullen, 1992). Along with the production of tat, the Rev protein is produced. Rev reduces the level of tat and its own synthesis, then promotes the export of mRNAs from the nucleus (Cullen, 1992). The nef gene product was initially believed to have a negative effect on viral infection, however recently it has been shown to be important for viral pathogenesis (Cullen 1994; Ratner and Niederman, 1995).

#### vif

This protein has a role in increasing viral infectivity and also may have a role in packaging of the core of the virus (Hoglund et al., 1994; Liu et al 1995).

#### vpr

The function of vpr is unclear, it may be a regulatory protein as vpr mutant viruses replicate more slowly (Ogawa et al., 1989).

#### vpu

This protein may have a role in virus release (Klimkait et al., 1990).

The accumulation of mutations in the accessory genes of HIV-1 has been associated with reduced cytopatogenicity of the virus and an increase in viral persistence (Kishi et al., 1995). Two accessory proteins have been shown to be particle associated, vif (Karcezwski and Strebel, 1996) and vpr (Kewalramani et al., 1996).

### **1.3.3. HIV-1 Life Cycle.**

The life cycle of HIV-1 is illustrated in figure 1.4 and is typical of a retrovirus

#### Infection.

Infection of a susceptible cell with HIV-1 begins with the viral gp120 glycoprotein binding to the cellular CD4 molecule, the first retrovirus receptor to be identified (Dalgleish et al., 1984). The CD4 molecule is essential but not sufficient to ensure virus infection, as murine cells expressing CD4 bind HIV-1 but are not infected (Maddon et al., 1986). The entry cofactor "fusin" has recently been identified, expression of this molecule in CD4 positive non-human cell types allowed HIV infection (Feng et al., 1996).

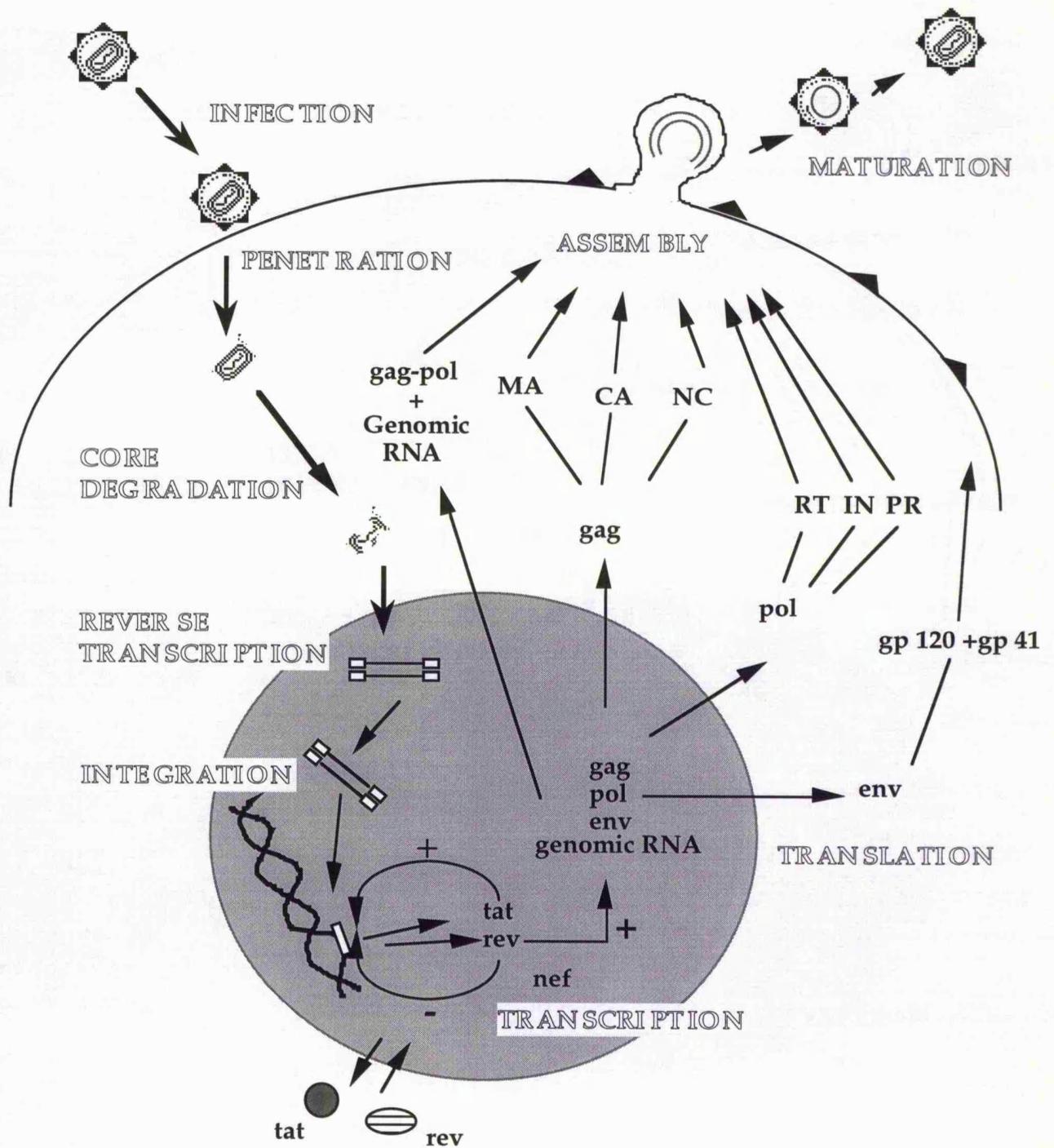


Figure 1.4 . Life cycle for HIV-1.

HIV-1 can enter a cell either by receptor mediated endocytosis using CD4 or by direct membrane fusion (Grew et al., 1990). Receptor binding is

followed by conformational changes of the receptor/viral gp120 complex (Littman, 1995) and penetration of the virion core into the cell.

#### Synthesis of DNA.

Proviral DNA is synthesised from the genomic RNA by reverse transcription. Replication is initiated from the tRNA primer by the viral reverse transcriptase enzyme, some limited DNA synthesis (strong stop cDNA) may occur within the virion before or during attachment of the particle to the cell (Zhang et al., 1995). Proviral DNA synthesis is complete within about 4 to 8 hours after infection, the DNA then associates with the nucleus of the cell. Circular viral DNA forms have been observed, but are thought to be by-products from the reverse transcription process (Brown et al., 1987). Integration of the provirus into the host DNA is catalysed by the viral integrase enzyme. Integration appears to be a random process into the host genome, although some evidence suggests that integration into transcriptionally active, chromatin-free regions of DNA may occur (Rohdewohld et al., 1987). Lentiviruses are unusual in that they can penetrate the nuclear membrane and integrate, they do not require replication of the cell.

#### Transcription.

The integrated provirus is initially transcribed like a cellular gene. The transcription initiation site (+1, position 455 in HIV-1<sub>NL4-3</sub>) is present within the HIV-1 LTR, upstream of this sequence is a proximal (promoter) and distal (enhancer) region. The promoter sequences are typical of those used by RNA polymerase II, eg. the TATA box is between nucleotides -22 to -27. Other transcription factor binding sites are present eg. three SP1 binding sites. In the enhancer region, HIV-1 promoter activity can be stimulated by binding of cellular factors like NFkB in activated T cells

(Nabel and Baltimore, 1987). The interaction of these cellular factors produces only a low, basal level of viral mRNA synthesis. Transcription from the viral LTR is then amplified by the action of the first gene product, the transactivating tat protein.

HIV-1 tat protein binds to the TAR RNA sequence, resulting in an increase in transcription from the viral LTR. Tat is an early gene product, along with rev and nef. Enhanced tat activated transcription leads to the accumulation of a critical level of the rev gene product. Rev acts to reduce transcription of the multiply spliced early transcripts, and stimulates the production of the later singly and unspliced RNAs. Rev facilitates export of these later RNAs from the nucleus into the cytoplasm (Cullen 1992). The role of nef in the early stages of transcription is uncertain, although it forms 80% of the total early transcripts (Cullen 1994).

### Splicing.

Gene expression by HIV-1 is complex and regulated at the level of splicing. There is a major splice donor site (SD) at the 5' end of the genome close to the PBS (GT at position 744-745 HIV-1<sub>NL4-3</sub>), and many other secondary splice donor and acceptor sites exist throughout the genome. The selection of the appropriate splice donor and acceptors is poorly understood, and is probably dictated by sequences within the RNAs. Three classes of transcripts are produced:

1. Unspliced genomic RNA, codes for gag, gag-pol.
2. Singly spliced RNAs eg. env.
3. Multiply spliced RNAs eg. tat, rev and nef.

All retroviral sub-genomic RNAs share common 5' and 3' ends with the genome, figure 1.5.

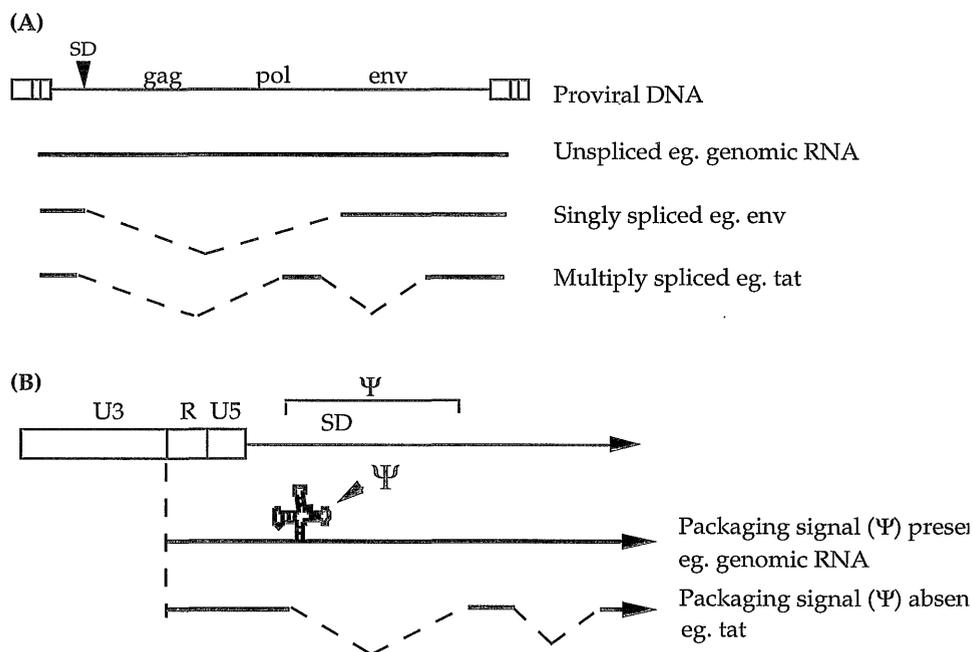


Figure 1.5 HIV-1 RNA Splicing. Panel A) Schematic representation of three HIV-1 mRNAs; unspliced, spliced and multiply spliced. Panel B) Splicing removes sequences between the splice donor (SD) and the start of the gag gene in all sub-genomic RNAs. This removes the packaging signal ( $\Psi$ ), which is only resident on unspliced genomic RNA.

All transcripts contain sequences from the start of transcription (+1 at position 455 in HIV-1<sub>NL4-3</sub>) to the 5' major splice donor site (position 745 in HIV-1<sub>NL4-3</sub>). Only the unspliced genomic RNA contains sequences between the 5' splice donor and the initiation of the gag gene (position 790). The position of the packaging signal ( $\Psi$ ) between splice sites means that it is present only on unspliced genomic RNA which is recognised and incorporated into assembling viruses.

### Translation.

Translation of viral mRNAs produces viral proteins and is also used to regulate gene expression. In common with other viruses, ribosomal frameshifting (Kastelein et al., 1982) occurs in HIV-1 to regulate levels of the gag and gag-pol fusion proteins. Gag is the major capsid antigen and is required in larger amounts than the gag-pol fusion protein which contains the viral enzymes. Ribosomal frameshifting at the 'slippery' U/A rich sequence at the gag:pol junction ensures that the level of gag to gag-pol is about 20 to 1 (Wilson et al., 1988).

Post translational myristoylation of HIV-1 gag and gag-pol proteins ensures targetting of the proteins to the plasma membrane of the cell, although the C-terminal domain of the polyprotein gag may also contain important membrane localisation signals (Platt and Haffar, 1994). The basicity of regions of the gag polyprotein may also be involved with membrane interactions (Ehrlich et al., 1996). The envelope protein is also modified, it is translated on ribosomes associated with the Golgi apparatus where it is extensively glycosylated before export to the cell membrane.

### Virion Assembly.

HIV-1 assembly occurs at the plasma membrane like C-type viruses. The unspliced genomic RNA is selectively recognised and packaged into the virion, probably by the gag-pol polyprotein (Stewart et al., 1990; Richardson and Lever 1996). The assembling particle buds out from the surface of the cell, capturing the env glycoprotein from the host membrane (Hunter, 1994). Following release of the immature virus the viral protease is activated, leading to a rearrangement of the core into a characteristic conical shape and maturation of the genomic RNA dimer (Fu and Rein, 1993; Fu et al., 1994).

## 1.4. Retroviral Genomic RNA.

### 1.4.1. Features of Retroviral Genomic RNA.

Retroviral RNAs range in size from 8 to 10 kb in length, HIV-1 genomes are around 9 to 10 kb eg. 9,709 bp for HIV-1<sub>NL4-3</sub>. The internal regions of the genome contain the structural genes, the cis-acting non-coding RNA sequences (LTRs) are arranged terminally, figure 1.6.

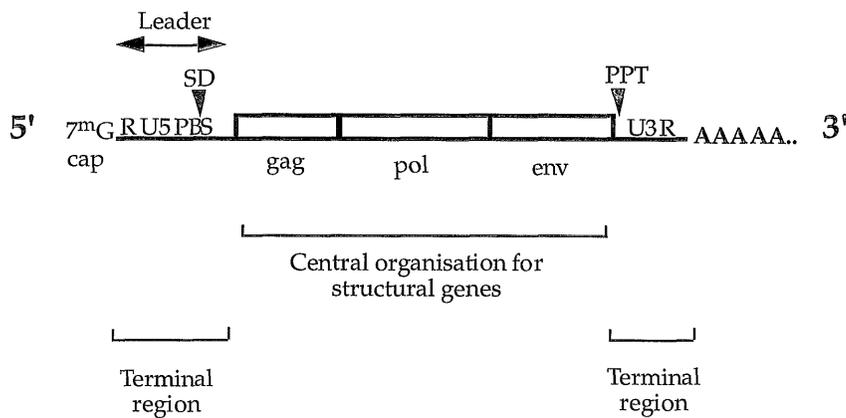


Figure 1.6. Features of retroviral genomic RNA.

#### Sequences at the 5' end

Retroviral genomic RNA molecules are capped at the 5' end by 7-methyl guanosine and contain occasional methylated adenine residues throughout the genome like any cellular mRNA. The genome is terminally redundant with the repeated sequence (R) at each end, this is important for DNA strand transfer during reverse transcription. The U5 sequence is between R and the PBS, this is the first region copied during cDNA synthesis and it is involved in the initiation of reverse transcription. The primer binding site (PBS) is the location of the annealed tRNA replication primer, tRNA<sup>Lys,3</sup> for HIV-1, essential to initiate reverse transcription. The major 5' splice donor site (SD) is present just

downstream from the PBS in HIV-1<sub>NL4-3</sub> (744-745) , although its location varies in other retroviruses. The entire non-coding RNA sequence before the gag gene begins is termed the leader region. This contains the TAR stem loop and sequences important for packaging and dimerisation of the genomic RNA.

#### Sequences at the 3' end.

Non-coding sequences at the 3' end contain the polypurine tract (PPT, position 9059-9073 in HIV-1<sub>NL4-3</sub>) required for the initiation of reverse transcription. This is followed by a U3 region, which ultimately resides at the 5' end of the provirus, and contains important signals for transcription and integration. Polyadenylation occurs onto the 3' end of the retroviral RNA although the significance of this, if any, for the retrovirus is uncertain.

#### **1.4.2. Dimerisation of Retroviral RNA.**

All retroviruses contain a dimeric RNA genome consisting of two unspliced genomic RNA molecules non-covalently linked together. This arrangement is a unique feature to retroviruses and was identified from biochemical and electronmicroscopy analysis of retroviral genomic RNAs.

#### Biochemical analysis.

Examination of RNA tumour viruses showed that their genomes were a 60 to 70 S RNA complex with associated 4S tRNA replication primer molecules (Robinson et al., 1965; Bonar et al., 1967; Bishop et al., 1970 a&b; Erikson and Erikson, 1971). Two configurational variants of oncornavirus genomic RNA were identified, vRNA (native) and vRNA' (denatured). The vRNA complex sedimented more slowly during sedimentation analysis and could be composed of several 35S (vRNA') molecules

(Duesberg, 1968, Bader and Steck 1969, Erickson 1969 and Montagnier et al., 1969). The complex dissociated on heating at temperatures around 60°C or by treatment with reagents that disrupted hydrogen bonds eg. formamide.

In retroviral infected cells the 60-70S RNA complex was primarily cell surface associated with the membrane fraction (Fan and Baltimore, 1973). Treatment of infected cells with actinomycin D subsequently inhibited formation of the complex, showing that RNA was essential (Levin et al., 1974). Estimation of the molecular weight (Luborsky, 1971; Kung et al., 1975) and electron microscopy of vRNA indicated that it consisted of at least two molecules of vRNA', associated together by hydrogen bonds (Bader and Ray, 1976)

#### Electronmicroscopy Studies.

The biochemical analysis of retroviral genomic RNA indicated that it consisted of two or more identical subunits of around 35S, with smaller associated RNAs. Electronmicroscopy studies were performed on a number of retroviral genomic RNAs, eg. RD-114 (Kung et al 1975), MLV (Friend strain, Dube et al., 1976), Moloney sarcoma virus (Maisel et al., 1978), AKR, NZB, REV and wild mouse viruses (Bender et al., 1978), REV (Gonda et al., 1980) and MoMuLV and RSV (Murti et al., 1981), shown in figure 1.7.

Electronmicrographs demonstrated that the RSV genome consisted of two RNAs associated towards their 5' end by a recognisable structure. This region of RNA was termed the dimer linkage site or structure (DLS). Under less stringent conditions the RNAs were observed to be in contact throughout the genome (Mangel et al., 1974) so it appeared that the DLS was probably the most stable point of contact between the two RNAs.

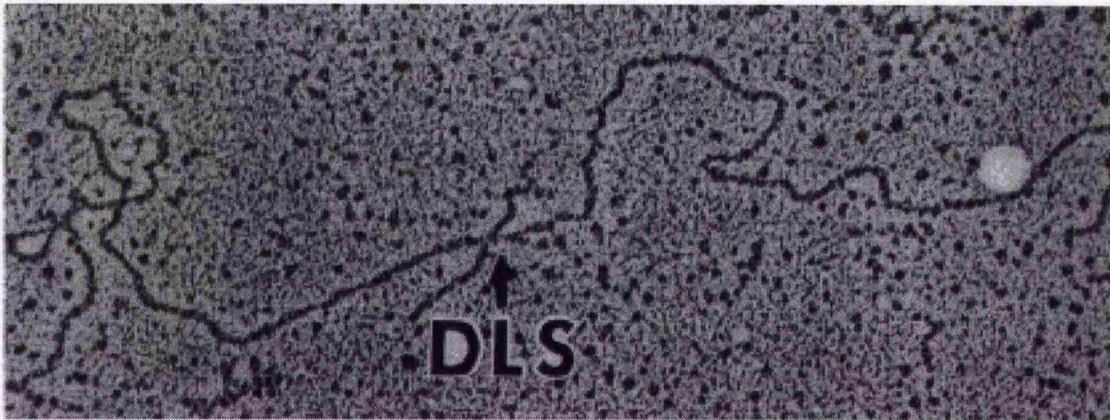


Figure 1.7. RSV genomic RNA analysed by electronmicroscopy (Adapted from Murti et al., 1981). The two RNAs are visible and the dimer linkage site (DLS) is indicated.

In every case where the dimer linkage structure has been accurately mapped by electronmicroscopy, it was located near the 5' end of the genome, downstream (3') of the PBS in the leader RNA region (Murti et al., 1981, Bender et al., 1978; Kung et al., 1976, figure 1.6). For example, the DLS region mapped by electronmicroscopy for RSV was around residues  $511 \pm 22$  nucleotides (Murti et al., 1981). Over 40% of retroviral genomic RNAs examined by electronmicroscopy have a 5' DLS which does not reform after denaturation (Murti et al., 1981).

Further confirmation of the dimeric nature of retroviral RNA genomes has come from electrophoretic (eg. Oertle and Spahr, 1990) and chromatographic analysis (Pager et al., 1993). Genetic studies also showed that frequent recombination occurred within the viral particles and this was correlated with two different genomes within one virion (Weiss et al., 1973). However the best evidence for the existence of the DLS is from electronmicroscopy.

#### **1.4.3. The Role of RNA Dimerisation in the Viral Life Cycle.**

Retroviral RNA dimerisation may precede, occur during or after particle assembly and morphogenesis. Dimerisation may also have a role in reverse transcription, reduction of viral mRNA translation and RNA packaging.

#### **Reverse Transcription.**

During reverse transcription the dimeric RNA genome is copied into a double stranded cDNA provirus, by the viral reverse transcriptase enzyme. Along with replication of the genome, recombination also takes place. Recombination in viruses can occur by reassortment of genome fragments eg. influenza virus. Recombination in retroviruses is different, it occurs between the two strands of the RNA dimer molecule (Panganiban and Fiore, 1988), as no pool of replicative intermediates is formed during reverse transcription. Retroviral recombination is frequent, over 40% of particles have at least one exchange per 10 kb virus genome in each replication cycle (Hu and Temin, 1990).

Retroviral recombination is a valuable mechanism to generate variants in the virus population eg. to assist evasion of the host immune response, and to remove strand breaks which occur throughout the RNA genome (Jones et al., 1990; Coffin 1990; Temin, 1991; Luo and Taylor, 1990; Peliska and Benkovic, 1992). During reverse transcription the nascent DNA strand can transfer to a similar or identical sequence on the other RNA strand and continue primer extension ie. homologous recombination. If the two parental genomes are different, a new recombinant is produced. This can occur if a cell contains two different proviruses. However it has been shown that in the absence of recombination, one viral RNA is sufficient for the synthesis of viral DNA (Jones et al., 1993). Recombination between

the two strands of the retroviral RNA dimer may occur by a forced copy choice mechanism, figure 1.8.

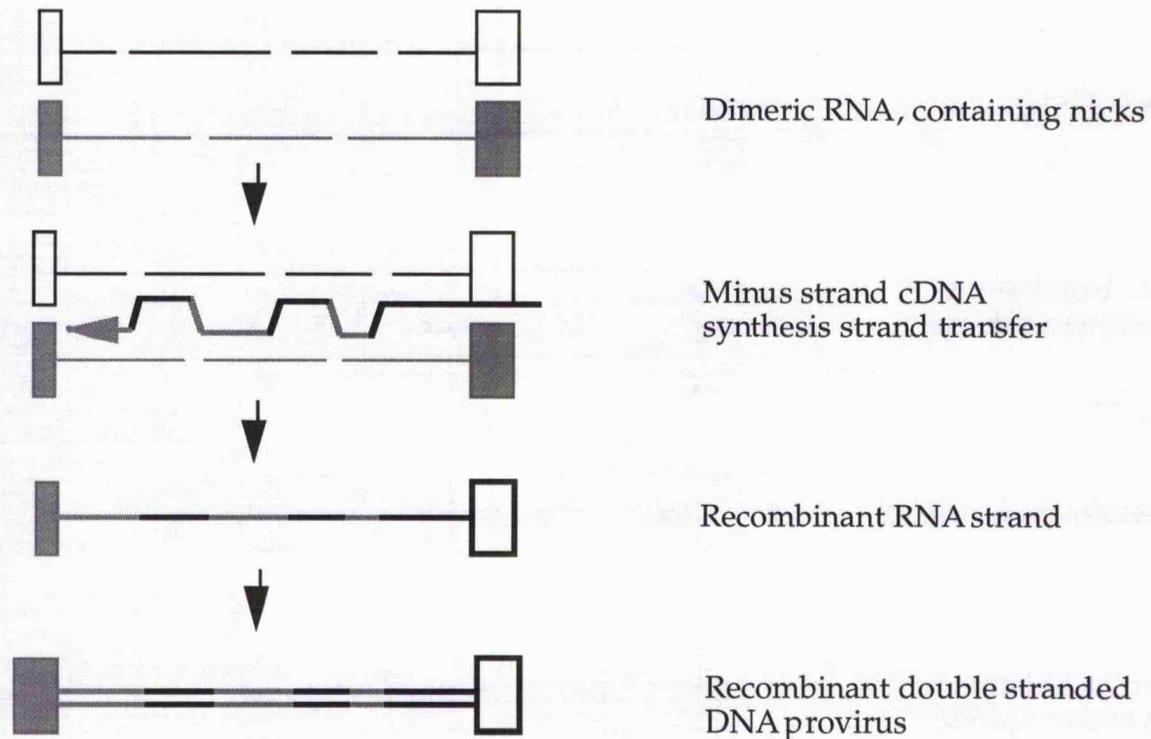


Figure 1.8. Forced copy choice model for retroviral RNA recombination.

In the forced copy choice model (Coffin et al., 1980), strand transfer occurs during minus strand DNA synthesis in response to a strand break in the RNA. Synthesis continues using the information on the alternative RNA strand. Along with strand breaks, stalling of the reverse transcriptase at sequences or structures that make continued synthesis difficult may also force the strand transfer (Wu et al., 1995). Alternatively recombination may occur by the strand aggression model (Skalka et al., 1982). Here synthesis of two complete minus strand DNA molecules is complete, before recombination occurs during plus strand DNA synthesis. For either

model, information on both RNA strands is used to generate the recombinant genome.

Recombination is advantageous for the survival of the virus, and so perhaps this is one of the reasons why retroviruses have evolved to contain two genomic RNAs in their virions (Temin, 1991).

#### Effect of RNA Dimerisation on Viral mRNAs Translation.

According to the ribosome scanning model for the initiation of translation (Kozak, 1989), the ribosome binds to the 5' end of the mRNA and locates the correct AUG initiation codon, unwinding RNA secondary structure as it goes. Extensive secondary structure may be inhibitory to translation, such a conformation may result due to dimerisation of viral RNAs.

Unspliced genomic RNA acts as a template for the translation of gag polyproteins, eg. the 35S RNA of RSV codes for the Pr76<sup>gag</sup> and Pr180<sup>gagpol</sup> proteins. Native RSV RNA was poorly translated, however heat denatured dimer or 35S monomer was efficiently translated (Darlix, 1986). Nucleocapsid protein facilitated dimerisation of the 35S RNA, or sub-genomic RNAs containing the RSV leader, also lead to a reduction in the translation of the RNA (Bieth et al., 1990). Similar observations have been claimed for HIV-1 and MoMuLV RNAs (Darlix et al., 1990). *In vivo* a reduction of translation from the 35S RNA may drive dimer formation and virion assembly by construction of a 'packaging pool' of pre-formed dimers ready to be encapsidated (Bieth et al., 1990).

An alternative explanation for the reduction in translation of unspliced genomic RNA, may be due to the presence of the encapsidation signal on the 35S RNA. For HIV-1, the 5' packaging signal inhibited translation of the gag gene *in vitro* although this was not due to RNA dimerisation (Miele et al., 1996). The 5' packaging signal is a favoured site for interaction

with the viral nucleocapsid protein (Darlix et al., 1990; Sakaguchi et al., 1993), perhaps this binding inhibits translation and ensures the availability of full length message for packaging.

#### Packaging of Retroviral genomic RNA.

The recognition and specific incorporation of genomic RNA into retroviral particles is an essential stage in the retroviral life cycle. Retroviral RNA constitutes about 1% of total cellular RNA, yet it is the predominant RNA incorporated into virions. Selectivity is essential to avoid capture of a cellular mRNA which may become part of the virion genome (Linial and Miller, 1990). Dimerisation of the genomic RNA also occurs, however it is unknown whether this occurs in the cytoplasm and acts as a signal for encapsidation, or in the virion after packaging. There may be a functional relationship between dimerisation and packaging because the sequences thought to be involved in these two processes are co-incident.

A mutant of RSV, SE21Q1b, produced virions that had the full complement of viral proteins (including reverse transcriptase), but were deficient in genomic RNA. Analysis of the provirus showed a small deletion of 180 base pairs at the 5' end of the genome, which was responsible for the lack of RNA incorporation (Linial et al., 1978; Shank and Linial; 1980). The identification of other retroviral RNA packaging signals has been performed using deletion mutagenesis resulting in an RNA encapsidation defect (eg. ASV Katz et al., 1986; HIV-1 Lever et al., 1989; MoMuLV Adam and Miller 1988; VL30 Torrent et al., 1994). Vectors bearing sequences from the candidate packaging regions have also been used to identify the minimum sequences sufficient for encapsidation of an RNA into a particular virus particle (Armentano et al., 1987; Adam and Miller, 1988). The RNA packaging signal was cis-acting, as its addition onto

non-retroviral RNAs allowed their incorporation into particles presumably by the interaction with a trans-acting factor (Rizvi and Panganiban 1993).

#### Cis-acting RNA Packaging Sequences.

RNA packaging signals, or Psi ( $\Psi$ ) sites, have been located within the leader region for several retroviruses, and can be regarded as a 'label' for genomic RNA. For MLV the packaging signal is present downstream from the 5' splice donor, so it is removed by splicing and not found on subgenomic RNAs. The situation is more complex for avian retroviruses. For RSV, the packaging signal is upstream of the splice donor and so is present on spliced mRNAs eg. env mRNA, although this is only rarely incorporated (Schlesinger et al., 1994). Additional packaging signals may be present at the 3' end of the RSV genome, so a complex interaction between the ends may be required for packaging, this only occurs on the full length genome. For ASV a base-paired structure is required for efficient RNA encapsidation (Knight et al., 1994). A discontinuous encapsidation signal has also been proposed for BLV, a region in non-coding RNA between the PBS and gag was the primary signal, with a secondary signal within gag (Mansky et al., 1995). Similarly the packaging signal of MoMLV extends into the gag coding region (Bender et al., 1987). Redundancy in the function of the packaging signal has been observed, HIV-1 viruses were able to recognise and package vectors containing the SIV encapsidation element. The two viral encapsidation sequences were diverse, indicating that similar structures had been recognised (Rizvi and Panganiban, 1993).

The complex nature of the packaging signal has also been demonstrated for HIV-1. Deletion analyses originally located a small region between the 5' splice donor and the gag gene, removal of this region lead to a 98%

reduction in RNA packaging (Lever et al., 1989; Clavel and Orenstein 1990; Aldovini and Young 1990). Structural determinations of the leader region based on chemical probing, enzymatic digestions, phylogeny and computer folding, have produced several structures (Harrison et al 1992; Baudin et al., 1993; Sakaguchi et al., 1992; Luban and Goff 1994; Clever et al., 1994; Harrison et al., 1995; McBride and Panganiban, 1996). Stable stem loops within these structures have been shown to be important for HIV-1 RNA packaging and binding of gag proteins (Hayashi et al., 1992; Luban and Goff 1994; Clever et al., 1994; McBride and Panganiban 1996). However the 5' region may not be sufficient for HIV-1 RNA encapsidation, as wild type HIV-1 failed to package vectors containing RNA sequences from this region (Berkowitz et al., 1995). The encapsidation signal may be multipartite, several hairpin structures which base pair to one another may be required for encapsidation (McBride and Panganiban, 1996). An interaction between two hairpin structures associated with the packaging signal for SNV was important for RNA incorporation (Jones et al., 1993). The integrity of RNA secondary structures appears to be essential for efficient encapsidation (Knight et al., 1994).

#### Trans-acting Proteins Involved in RNA Packaging.

Along with mutations in RNA sequences, retroviruses containing mutations in the nucleocapsid domain of the gag polyprotein are also packaging defective.

The gag polyprotein precursor recognises and binds to genomic RNA (Stewart et al., 1990; Kaye and Lever, 1996). It is sufficient for RNA incorporation as viral particles generated in the absence of pol and env genes encapsidate normal levels of viral genomic RNA (Oertle and Spahr, 1990; Sakalian et al., 1994). Mutations in the conserved Cys-His-Cys-Cys nucleocapsid domain lead to a reduced RNA incorporation for RSV

(Meric and Spahr, 1986), MoMuLV (Gorelick et al., 1988; Meric and Goff, 1989), and HIV-1 (Gorelick et al., 1990). However this interaction may not be sufficient for RNA capture as specific incorporation can occur in RSV if the zinc finger motifs are deleted (Aronoff et al., 1993). Mutations of both zinc finger motifs in the HIV-1 nucleocapsid domain did not eliminate RNA packaging, but decreased specificity (Zhang and Barklis, 1995). Further evidence that RNA recognition is achieved by the nucleocapsid domain of the gag polyprotein comes from the construction of chimeric virus particles. HIV-1 RNA was incorporated into MoMLV particles that contained the HIV-1 nucleocapsid domain replacing the MoMLV sequences (Zhang and Barklis, 1995). Conversely, RNAs containing the MoMLV packaging signal were packaged by an HIV-1 virus containing the MoMuLV nucleocapsid domain. Both the MoMuLV Cys-His boxes and flanking regions were required for the specific recognition of the MoMLV Psi element (Berkowitz et al., 1995).

#### **1.5. Mechanisms of Retroviral RNA dimerisation.**

The physical analysis of retroviral RNAs performed on isolated 20kB RNA molecules by biochemical and electronmicroscopy techniques, has previously been described. These techniques have limitations, as does the practicality of manipulating such a large RNA. Therefore the discovery that small RNAs which corresponded to the DLS and packaging regions dimerise *in vitro*, lead to the analysis of mechanisms for RNA dimerisation.

The two RNA molecules of isolated retroviral genomic RNA remained associated after treatment with denaturants, following extensive proteolysis, and organic solvent extractions designed to destroy RNA-protein interactions (Duesberg, 1968; Mangel et al., 1974; Stoltzfus and

Snyder, 1975; Kung et al., 1976; Bender et al., 1978; Murti et al., 1981; Fu and Rein, 1993). This indicated that a 'protein linker' (Kung et al., 1976) could not be the dimer linkage structure, moreover a direct RNA:RNA interaction was likely. The RNA genomes appeared to be arranged in parallel ie. 5' to 5' and 3' to 3', suggesting that a simple Watson-Crick helical structure did not mediate the strand association (Kung et al., 1976). Further support for a direct RNA:RNA contact has come from observations that RNA sequences corresponding to the DLS region of several retroviruses can dimerise *in vitro* (eg. Darlix et al., 1990). The dimerisation occurred in the absence of any proteins, indicating that dimerisation was an intrinsic property of the RNA.

#### **1.5.1. In vitro RNA Dimerisation.**

*In vitro* transcribed RNAs containing sequences from around the DLS and packaging signal from several retroviruses can dimerise in the absence of any proteins. By deletion mutagenesis of a dimerisation competent 1333 nucleotide RNA encompassing the 5' end of HIV-1<sub>MAL</sub>, a 100 nucleotide RNA was identified that contained a dimerisation domain (Darlix et al., 1990). Similarly an RSV RNA transcript could also form dimers, sequences important were present within the first 600 nucleotides of RSV RNA (Bieth et al., 1990). These results were consistent with the electronmicroscopy localisations of the DLS region. For example electron microscopy of a 1600 nucleotide MoMLV RNA dimer formed *in vitro* had a dimer linkage structure similar to that extracted from virions (Prats et al., 1990). By deletion mutagenesis approaches dimerisation competent RNA transcripts from the DLS of several retroviruses were identified eg. RSV (Bieth et al., 1990; Lear et al., 1995), MoMLV (Roy et al., 1990; Prats et al., 1990; Tounekti et al., 1992; Girard et al., 1995), REV (Darlix et al., 1992), HaSV (Feng et al., 1995), BLV (Kato et al., 1993), HIV-2 (Berkhout et al.,

1993), and HIV-1 (Darlix et al., 1990; Marquet et al., 1991; Awang and Sen 1993; Sakaguchi et al., 1993; Sundquist and Heaphy 1993; Weiss et al., 1993; Laughrea and Jette 1994 & 1996; Marquet et al., 1994; Paillart et al., 1994; Skripkin et al., 1994; Muriaux et al., 1995 & 1996).

RNA dimerisation was a specific process as not all RNA transcripts examined dimerised. The dimerisation of RNAs containing sequences downstream from the major splice donor was dependent on RNA concentration, temperature and on solution monovalent and divalent cations (Marquet et al., 1991). Dimerisation of these RNAs was reported to proceed slowly and require high ionic strength conditions (Marquet et al., 1991; Sundquist and Heaphy 1993; Weiss et al., 1993). The changes in conformation of MoMLV RNAs during dimerisation have been investigated from this region (Tounekti et al., 1992; Mougél et al., 1993). Dimerisation induced specific reactivity changes and rearrangement of the structure around the splice donor and gag initiation region. However RNAs with additional 5' sequences dimerised rapidly in lower ionic strength buffers (Marquet et al., 1994; Skripkin et al., 1994; Paillart et al., 1994). The analysis of longer RNAs has shown that multiple regions of viral RNAs can dimerise eg. HaSV, (Feng et al., 1995), RSV (Lear et al., 1995) and HIV-1 (Paillart et al., 1994). For HIV-1 the originally defined DLS (sequences downstream from the splice donor (Darlix et al., 1990)) has been expanded to include a DIS (dimer initiation site) region (Skripkin et al., 1994). A second DLS like regions has been proposed for HIV-1 5' (upstream) of the DIS site (Laughrea and Jette, 1996).

#### **1.5.2. The effect of nucleocapsid proteins on RNA dimerisation, *in vitro*.**

Barat et al. demonstrated enhanced tRNA annealing to and dimerisation of a 1330 nucleotide RNA from the DLS region of HIV-1. The HIV nucleocapsid protein, as well as that from RSV and MoMLV, promoted

the RNA dimerisation (Barat et al., 1989). RNA dimerisation *in vitro* was also facilitated by NCp10 of MoMLV (Prats et al., 1990) and NCp12 of RSV (Bieth et al., 1990). Similarly HIV-1 NC protein extracted from virions could facilitate dimerisation of a 104 nucleotide HIV-1 RNA (Darlix et al., 1990). The RNA dimer formed by the action of the nucleocapsid protein was indistinguishable in stability and electrophoretic mobility from the RNA dimer formed spontaneously in the absence of protein. Both the NCp15 and NCp7 proteins can facilitate dimerisation of HIV-1 RNA transcripts (DeRocquigny et al., 1991). This alleviated the need for high salt incubation conditions *in vitro*. A recombinant NCp15:GST protein also promoted RNA dimerisation, the effect was due to the NCp15 protein (Weiss et al., 1992). The functional domains of the nucleocapsid protein were shown to be the basic amino acids surrounding the zinc finger motifs of the nucleocapsid protein (DeRocquigny et al., 1992 and 1993). The NC protein positions the tRNA replication primer onto the PBS and stimulates reverse transcription (Prats et al., 1988; Barat et al., 1989; DeRocquigny et al., 1992; Weiss et al., 1992), probably by its helicase activity (Khan and Giedroc, 1991). The way in which nucleocapsid assists dimerisation is uncertain, it was essential for the dimerisation of a 44 nucleotide HIV-1 RNA (Sakaguchi et al., 1993) and may function by binding to the RNA, inducing a conformational change driving the monomer: dimer equilibrium towards dimer formation. It has also been proposed that the NCp7:RNA interaction is followed by NC:NC interactions leading to a condensation of the two viral RNAs into a chromatin like structure perhaps facilitating dimerisation.

### **1.5.3. Relevance of *in vitro* RNA dimerisation analysis.**

It has been assumed that the RNA dimers formed from subgenomic RNAs *in vitro* are accurate models for the dimer linkage structure formed

within virion particles. There is some support for this, RNA dimers formed *in vitro* and *in vivo* are reported to have similar morphologies (Prats et al., 1990), thermal stabilities (MoMLV, Roy 1990; REV, Darlix et al., 1992; HaSV, Feng et al., 1995; HIV-1, Laughrea and Jette, 1994) and chemical modification patterns (Tounekti et al., 1992; Alford et al., 1991). Also an RNA stem-loop proposed as the HIV-1 packaging signal has the same structure in an *in vitro* synthesised RNA transcript (Baudin et al., 1993; Harrison and Lever, 1992) and in the genomic RNA extracted from infected cells (Hayashi and Okamoto, 1993). Furthermore, minimal dimerisation domains mapped by deletion analyses *in vitro* correspond reasonably well to sequences implicated in retroviral RNA packaging and dimerisation *in vivo* (Darlix et al., 1990; Aldovini and Young 1990; Lever et al., 1989; Hayashi et al., 1992; Clavel and Orenstein 1990; Embretson and Temin 1987) although the correspondence is not perfect (Berkhout et al., 1993; Aronoff et al., 1993, Tchenio and Heidmann, 1995).

#### **1.5.4. *In vivo* RNA Dimerisation studies.**

RNA dimerisation may occur during particle morphogenesis or prior to RNA packaging. The analysis of rapid harvest RSV virions showed that the RNA in the particles was monomeric, even after gag polyprotein cleavage was complete (Cheung et al., 1972; Canaani et al., 1973; Lear et al., 1995). Changes in RNA structure have been observed within particles as they age. For RSV, RNA dimerisation was complete within 3 hours of virion formation, for visna virus maturation was complete within one hour (Cheung et al., 1972; Brahic and Vigne, 1975). RNA stabilisation also occurred for B77 sarcoma virus RNA (Stoltzfus and Snyder, 1975), although the RNA may have been packaged as an immature dimer. Protease deficient ALV and RSV viruses packaged RNA which remained monomeric (Stewart et al., 1990; Oertle and Spahr, 1990) suggesting that

dimers were not packaged and could not form without correct morphogenesis of the particle. Similarly nucleocapsid mutant viruses failed to dimerise their packaged RNA (Dupraz et al., 1990; Meric and Spahr, 1986). In fact the release of the nucleocapsid protein from the gag polyprotein precursor was necessary for RNA dimer formation *in vivo* (Oertle and Spahr, 1990).

Alternatively there is some evidence that dimerisation may occur prior to packaging. Actinomycin D treatment of MLV infected cells resulted in severely impaired RNA packaging, however the RNA within the particles was dimeric (Levin et al., 1974; Messer et al., 1981). Similarly, for HIV-1 and MoMLV, protease mutant viruses contained an RNA dimer of immature structure. This dimer had a reduced electrophoretic mobility, sedimentation co-efficient, and thermal stability (Fu and Rein 1993). The dimer then matured into a more compact structure concomitant with particle morphogenesis (Fu et al., 1993; Fu and Rein 1994). Re-interpretation of the earlier rapid harvest studies can be consistent with the immature dimer model, if the failure to detect the packaged "rapid harvest" dimers was due to their instability.

If RNA dimerisation preceeds packaging then RNA dimers should form within the cytoplasm of the infected cell. Conflicting reports exist concerning the detection of cytoplasmic dimeric RNA. An analysis of the RNA metabolism of MLV infected cells showed that the 70S RNA complex was primarily at the surface of the cell, but some was detected within the cytoplasm (Fan and Baltimore, 1973). Dimeric RNA has been isolated from REV infected cells which had been trypsinised to remove adsorbed virions (Darlix et al., 1992). Isolation of cytosolic RSV and MoMLV dimers has also been claimed (Prats et al., 1990). Protease

defective MoMLV and HIV-1 have been shown to contain an immature dimer, which could have formed within the cell cytoplasm. In contrast, Tchenio and Heidmann failed to detect higher ordered MoMLV RNAs within cellular RNA from infected cells, when monomeric RNA was readily detectable (Tchenio and Heidmann, 1995).

It has been proposed that full length genomic RNA is dimerisation competent without any protein involvement (Fu and Rein, 1994) from the analysis of protease defective viruses. This may be the case *in vivo*, however attempts to reconstitute genomic RNA dimers *in vitro* failed probably due to many structural features present along the RNA molecules (Kung et al., 1976; Mangel et al., 1974; Murti et al., 1981).

The functional importance of the 5' DLS region has been questioned. In an SNV vector packaging system, the location of the 5' DLS could be changed. This resulted in a five fold reduction in RNA packaging, however there was no effect on reverse transcription or recombination. This indicated that any alignment or conformation necessary for retroviral recombination or replication was not as a result of the position of the DLS (Jones et al., 1993). Furthermore, in MoMLV removal of the first 6.5kb of the genome (containing the DLS) showed that this region was not absolutely required for packaging, dimer formation or provirus establishment (Tchenio and Heidmann, 1995). Presumably the RNA:RNA interstrand contacts present throughout the genome allowed the dimeric complexes to form, or RNA dimerisation was not necessary for packaging and provirus formation.

### **1.6. Dimerisation of HIV-1 RNA.**

The dimerisation ability of HIV-1 RNA transcripts has been extensively analysed. By deletion mutagenesis of a dimerisation competent 1333 nucleotide RNA encompassing the 5' end of HIV-1<sub>MAL</sub>, a minimal 100 nucleotide dimerisation competent RNA was identified (Darlix et al., 1990). This region was co-existent with those important for HIV-1 RNA packaging (Aldovini and Young, 1990; Lever et al., 1989). Antisense RNA corresponding to this sense RNA failed to dimerise, indicating that a Watson-Crick base-pairing interaction was unlikely to stabilise the dimer. An RNA dimerisation mechanism involving purine tetrad structures was proposed (Marquet et al., 1991; Sundquist and Heaphy, 1993; Awang and Sen 1993). However other RNAs dimerised in the absence of guanine rich sequences ie. HIV-2 (Berkhout et al., 1993) and BLV (Kato et al., 1993) indicating another mechanism of dimerisation.

Two other models for RNA dimerisation by Watson-Crick base pairing have been proposed. An A-U base pairing scheme may be responsible for the dimerisation of a 44 nucleotide HIV-1 RNA for which the nucleocapsid protein was essential (Sakaguchi et al., 1993). This sequence was contained within the previously identified minimal dimerisation domain (Darlix et al., 1990). However sequences upstream of the splice donor were able to dimerise in both HIV-1 (Marquet et al., 1994; Skripkin et al., 1994; Laughrea and Jette 1994 & 1996; Paillart et al., 1994; Muriaux et al., 1995 & 1996), HIV-2 (Berkhout et al., 1993) and MoMLV (Girard, 1995) without any protein cofactors. For HIV-1 the upstream sequences contained an imperfect palindrome, as part of a kissing loop structure. This led to the kissing loop model for HIV-1 RNA dimerisation (Skripkin et al., 1994).

The ability of HIV-1 RNAs to dimerise by the formation of a guanine tetrad, or a kissing loop, was dependent on the RNA sequences examined.

Therefore a bi-partite dimerisation domain where the kissing loop acts as a dimer initiation signal (DIS) which may subsequently be stabilised by the formation of a guanine tetrad in the DLS was proposed (Paillart et al. 1994).

**1.6.1. The Guanine Tetrad Model for HIV-1 RNA Dimerisation.**

Guanine tetrads

A guanine tetrad is a nucleic acid structure. Guanine tetrads, or G-quartets, are composed of planar arrays of four guanines. Each guanine may come from a single folded back nucleic acid strand, but is more usually from two or four strands in a parallel or antiparallel configuration (Sen and Gilbert, 1988; Williamson et al., 1989; Kim et al., 1991; Aboul-ela et al., 1992), figure 1.9.

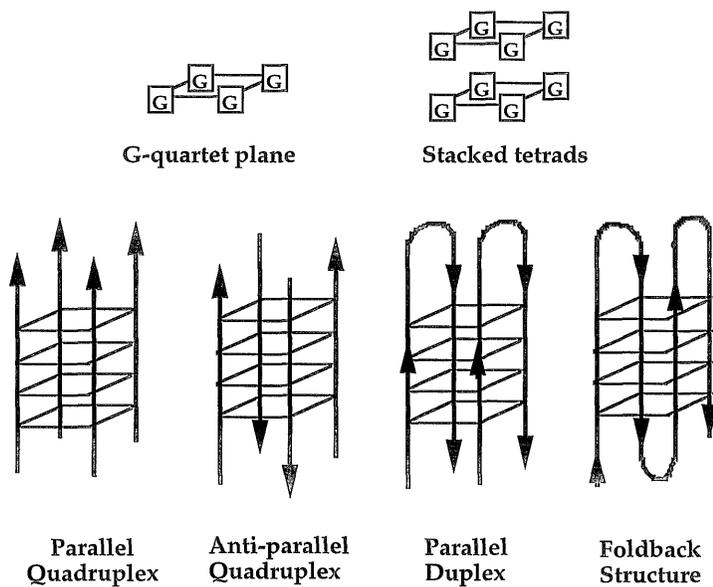


Figure 1.9. G-quartet structures can form with parallel or antiparallel strand arrangements.

The quadruple helical structure is stabilised by Hoogsteen hydrogen bonding between adjacent guanines, figure 1.10 A. This hydrogen bonding scheme involving the N7 to N2 and O6 to N1 positions of each guanine, was first proposed to explain the organisation of 3'-GMP in oriented fibres (Gellert et al., 1962; Lipanov et al., 1990). Several oligonucleotide structures containing multiple G-quartets stacked into short quadruple helices have been characterised (reviewed in Sundquist, 1991; Guschlbauer et al., 1990; Sen and Gilbert 1991). Guanine quadruple helices are further stabilised by metal ion co-ordination within the axial channel of the quadruplex (figure 1.10 B). The eight guanine O6 atoms from two adjacent quartets have unpaired electrons which chelate the metal cation. Potassium ions have an optimal ionic radius for this interaction and therefore increase the thermal stability of G-quartet structures to a greater extent than do other group I cations such that  $K^+ > Na^+ > Cs^+ \approx Li^+$  (Sundquist and Klug, 1989; Raghuraman and Cech, 1990; Pinnavaia et al., 1978; Oka and Thomas, 1987; Kang et al., 1992; Detellier and Laszlo, 1980). Enhanced thermal stability in the presence of potassium is a diagnostic feature of a guanine quartet structure. Stacking and dipole interactions lead to further increases in stability.

The association of two nucleic acid chains (dimerisation) by the tetrad mechanism can occur if each RNA strand folds back on itself and contributes two strands to the quadruplex (figure 1.10 C). Two guanine tracts separated in the linear sequence are required. The RNA adopts a fold back structure and intra-strand hydrogen bonds between the guanines in each tract form. Dimerisation is complete by inter-strand hydrogen bond formation with another RNA molecule in a similar conformation. Therefore a nucleic acid strand must contain at least two separate tracts of guanines in order to dimerise by G-quartet formation. Such quadruple-

helical dimers have been structurally characterised for telomeric DNA sequences (Kang et al., 1992; Smith and Feignon, 1992; Scaria et al., 1992).

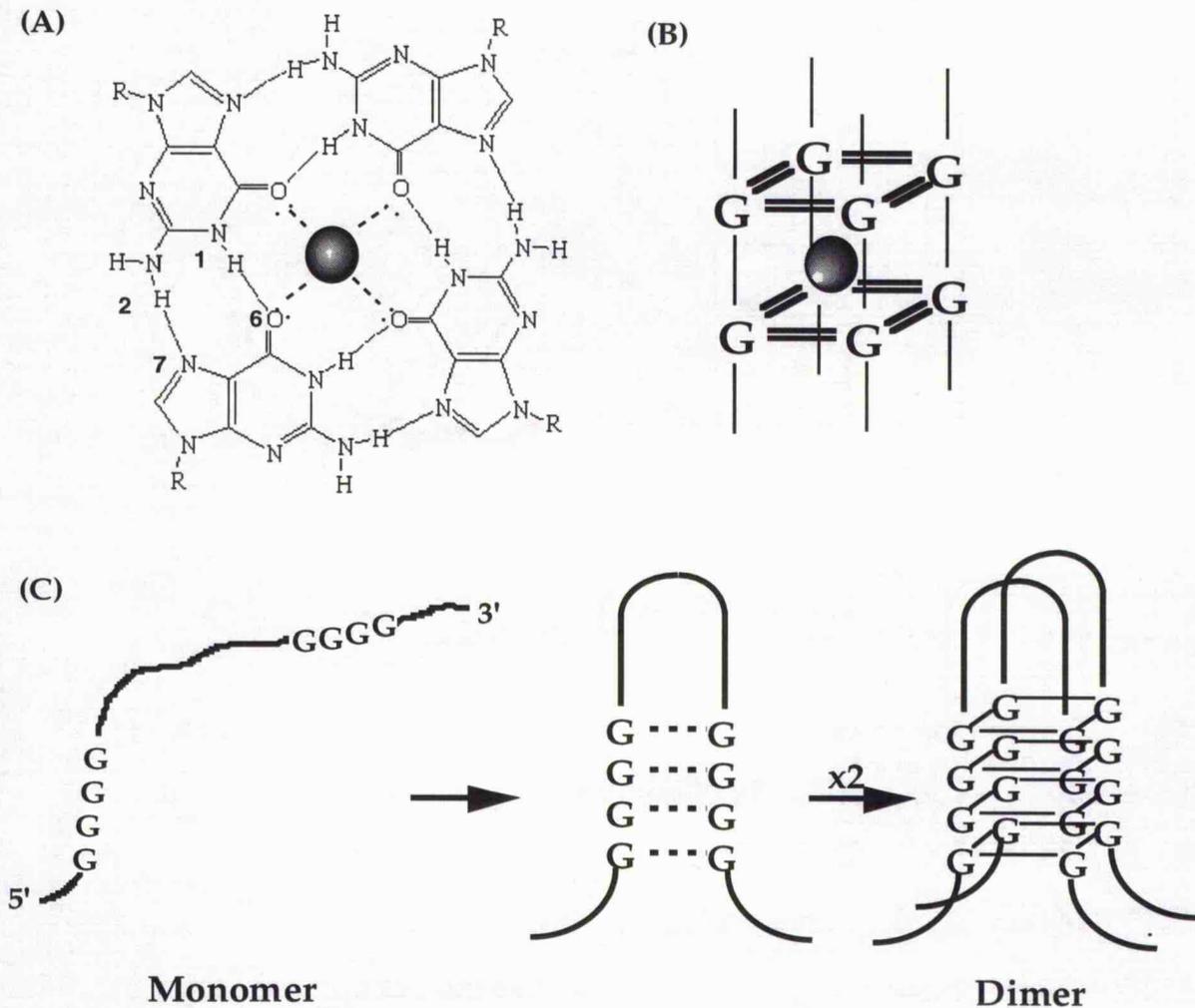


Figure 1.10. A) Hoogsteen hydrogen bonding scheme for one G-quartet stack. The O6 oxygen electrons co-ordinate the metal cation. B) Cations are co-ordinated into the central cavity of stacked tetrads. C) Quadruple helix dimer formation from the association of two hairpin foldback structures.

There is currently no known biological role for a G-quartet structure, however it has been suggested that telomeric DNA sequences may arrange into G-quartets. For the *Oxytricha* telomere sequence  $d(G_4T_4G_4)$ , the NMR

(Smith and Feignon, 1992) and crystal structure (Kang et al., 1992) confirm the tetrad arrangement of the guanines. Similar quartet sequences may be involved in the recombinational events required for immunoglobulin maturation (Sen and Gilbert 1988, 1990) and a potential role in chromosome alignment during meiosis has been proposed (Sen and Gilbert, 1988). The quartet structure also formed in an RNA oligonucleotide derived from *E.coli* 5S RNA sequence (Kim et al., 1991).

The quadruplex may act as a binding site for proteins, eg. MyoD (a transcription factor involved in myogenesis) binds to a G tetrad structure *in vitro* (Walsh and Gualberto, 1992). The  $\beta$  subunit of Oxytricha telomere binding protein greatly accelerated G-quartet formation *in vitro* where the protein acted as a molecular chaperone (Fang and Cech, 1993 a and b). This interaction between a telomeric protein and G-quartet formation is some evidence that the DNA structure may exist *in vivo*. Similarly a nuclease capable of recognising the tetrad structure has been identified in *Saccharomyces cerevisiae* (Liu et al., 1993). However telomere terminal transferase, another telomeric protein, was inhibited by G-quartet structures (Zahler et al., 1991). Similarly Hoogsteen base pairing was dispensible for telomere healing onto synthetic plasmid DNA templates introduced into the yeast *Saccharomyces cerevisiae* (Lustig, 1992). The proposal that a purine quartet structure (Marquet et al., 1991; Sundquist and Heaphy 1993; Awang and Sen, 1993) may mediate retroviral RNA dimerisation would be a function for this structure, *in vivo*.

#### Guanine Tetrads in RNA Dimerisation.

Purine quartets were initially proposed for the mechanism of HIV-1 RNA dimerisation *in vitro* by Marquet et al. (Marquet et al., 1991). The dimerisation of HIV-1 RNAs, containing the previously identified dimerisation domain (Darlix et al., 1990) and extending into the gag gene,

was analysed (figure 1.11, sequence 1). Dimerisation depended on RNA concentration, solution cations, temperature and pH. As the concentration of the RNA increased in the dimerisation reaction, the amount of RNA dimer also increased. The effect of solution cations was complex. Dimerisation was more efficient when magnesium (5mM) was present in the buffer. The percentage of dimer also increased with concentration of potassium ions (upto 1000mM), and spermidine at around cellular concentrations also assisted dimer formation. The optimum temperature for dimerisation was 37°C and at acidic pH (less than 6) dimerisation was inhibited indicating that C·C<sup>+</sup> and C·G<sup>+</sup> pairing was probably not involved (Gehring et al., 1993). Antisense RNAs failed to dimerise, and heterodimers could be formed between HIV-1 RNA and MoMLV or RSV RNA. The heterodimer formation of one HIV-1 RNA and one MoMLV (or RSV) RNA molecule suggested a common mechanism for retroviral RNA dimerisation.

These observations were inconsistent with a Watson-Crick based mechanism for dimerisation, however they were consistent with purine quartets. A conserved PuGGAPuA sequence was identified and proposed to be involved in quartet formation, involving both guanine and adenine residues (Marquet et al., 1991). Facilitated dimerisation with Li<sup>+</sup> as the solution cation indicated that guanine only quartets were not present for the RNA dimerisation investigated by Marquet et al. Higher ordered RNA structures with a molecular weight greater than the dimer were also observed. These trimers or tetramers can be accounted for by quartet formation, eg. a tetramer could be composed of four strands, each one contributing a single purine tract.

Further evidence for tetrad mediated dimerisation was produced from the analysis of HIV-1 RNA (732-858, figure 1.11 sequence 3) by Sundquist and Heaphy (Sundquist and Heaphy, 1993). By deletion mutagenesis the 817GG-

```

UACGCCAA--UUUUUGACUAGCGGAGCCUAGAAAGGAGAGAGAUAGGGUGGCGGAGCGGUCAGUUAUAAAGCGGGGCAAAAUUAGAUGCAUGGGAGAAAAUUCCGUUAAGG 1
GGUGAGUAAGCCAA--UUUUUGACUAGCGGAGGCTUAGAAAGGAGAGAGAUAGGGUGGCGGAGCGGUCAGUUAUAAAGCGGGGCAAAAUUAGAUGCAUGGGAGAAAAUUCCGUUAAGG 2
GGGCGCCGACUGGGUGAGUACCCCAAAAUUUUUGACUAGCGGAGGCTUAGAAAGGAGAGAGAGAUAGGGUGGCGGAGCGGUCAGUUAUAAAGCGGGGCAAAAUUAGAUAUAAAUGGGGAAAAUUCCGUUAAGGCCA 3

```

- Sequence**
1. Darlix et al., 1990. (311-415 HIV-1<sub>MAL</sub>)
  2. Awang and Sen, 1993. (DLS 112 HIV-1<sub>MAL</sub>)
  3. Sundquist and Heaphy, 1993 (732-858 HIV-1<sub>NL4-3</sub>)

Figure 1.11. Sequence alignment of dimerisation competent HIV-1 RNAs from the DLS region. Purine tracts identified as important for the dimerisation of each RNA are underlined.

GGGAGAA<sub>825</sub> sequence was essential for RNA dimer formation. Although sequences downstream from position 825 were not essential, their presence facilitated dimerisation. Furthermore the RNA dimers formed were preferentially stabilised when potassium was the solution cation,  $T_m \approx 80^\circ\text{C}$ . When  $\text{Na}^+$ ,  $\text{Li}^+$  and  $\text{Cs}^+$  were the coordinating ions the thermal stability of the RNA dimers was reduced accordingly to  $55^\circ\text{C}$ ,  $40^\circ\text{C}$  and  $30^\circ\text{C}$ . This cation dependent thermal stability suggested a quartet structure consisting entirely of guanine residues, in contrast to the mixed tetrad proposed by Marquet et al. The observation that ammonium ions stimulated RNA dimerisation was further support for the purine quartet model for HIV-1 RNA dimerisation (Weiss et al., 1993).

Confirmation that a sub-genomic HIV-1 RNA could dimerise by the formation of a quartet structure containing guanine residues was shown by Awang and Sen (Awang and Sen, 1993). An HIV-1 RNA from sequences 743-855 (figure 1.11, sequence 2, DLS 112) was analysed, this was similar to the 104 nucleotide RNA sequence 1 (figure 1.11) analysed by Darlix et al. The DLS 112 RNA dimerised in a high salt buffer and the dimers were also significantly stabilised by potassium ions. By a combination of antisense oligonucleotide annealing and DMS modification protection studies, two guanine tracts involved in the dimerisation of this RNA were identified. Nucleotides <sub>817</sub>GGGGG<sub>821</sub> were protected from DMS alkylation upon dimerisation, in accord with the deletion analysis of Sundquist and Heaphy. The second protected guanine tract required by the tetrad mechanism was found to be <sub>836</sub>GGG<sub>838</sub>, this was in contrast to the results of Sundquist and Heaphy. They showed that sequences downstream from nucleotide 825 were not essential for RNA dimerisation, however their presence assisted dimerisation. This suggested that another guanine tract was involved in the RNA dimerisation observed by Sundquist and Heaphy, not requiring <sub>836</sub>GGG<sub>838</sub>.

To determine if a guanine tetrad was involved in HIV-1 virion RNA dimerisation, Fu and Rein analysed the thermal stability of isolated HIV-1 genomic RNA in solutions containing Li<sup>+</sup>, Cs<sup>+</sup>, Na<sup>+</sup> and K<sup>+</sup> ions (Fu and Rein, 1994). If a tetrad was part of the virion DLS, then an enhanced thermal stability in the presence of potassium may be expected. In fact, no tetrad features were evident in the thermal dissociation of the HIV-1 genomic RNA dimer (Fu and Rein, 1994). This indicated that a tetrad was probably not part of the retroviral DLS. However it is subject to the caveat that a 20kB RNA dimer may not behave as a 300 nucleotide dimer does *in vitro*. For example ion exchange may not occur in the virion dimer, accounting for the lack of cation dependent thermal stability. To determine if a tetrad is essential within virion RNA, mutagenesis of the guanines involved in the context of an infectious clone would be a better approach.

### **1.6.3. The A-U Base Pairing Model for HIV-1 RNA Dimerisation.**

The purine tetrad model could not account for the dimerisation of HIV-2 and BLV RNAs because these dimerised in the absence of any PuGGAPuA or similar consensus (Berkhout et al., 1993; Katoh et al., 1993). An alternative base pairing mechanism has been proposed for HIV-1 RNA, requiring the nucleocapsid protein and an RNA stem loop structure (Sakaguchi et al., 1993). A 44 nucleotide RNA (corresponding to 736-779 of HIV-1<sub>NL4-3</sub>) was identified which could adopt a two stem loop structure, based on computer predictions and RNase probing analysis, figure 1.12.

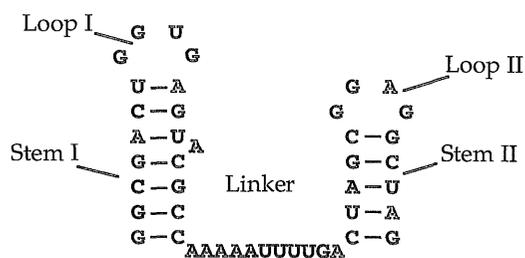


Figure 1.12. Predicted secondary structure for the 44 nucleotide RNA (Sakaguchi et al., 1993).

The nucleocapsid protein p7 of HIV-1 was shown to bind specifically to this RNA structure (Sakaguchi et al., 1993), the binding site had previously been identified as CCAAUUUUUG in HIV-1<sub>MAL</sub> (Darlix et al., 1990). The model proposed was that the NC binds to the stem loop 1 structure of one RNA, and acts as a protein bridge when it binds to either stem loop 1 or 2 of the second RNA, figure 1.13. The dimer linkage in this sequence was hydrogen bonding between residues UUUUAAAA induced by the nucleocapsid protein.

However mutagenesis of the linker sequence failed to prevent dimerisation of the RNA (Sakaguchi et al., 1993). For this reason, and the lack of sequence conservation between viruses, this mechanism was not examined in detail in this study, although the 44 nucleotide sequence is present in all of the HIV-1 RNAs analysed. Furthermore, longer RNAs dimerised in the presence or absence of the the A-U base pairing and tetrad sequence tracts, demonstrating that these sequences were not essential (Clever et al., 1995). This suggests another mechanism of dimerisation.

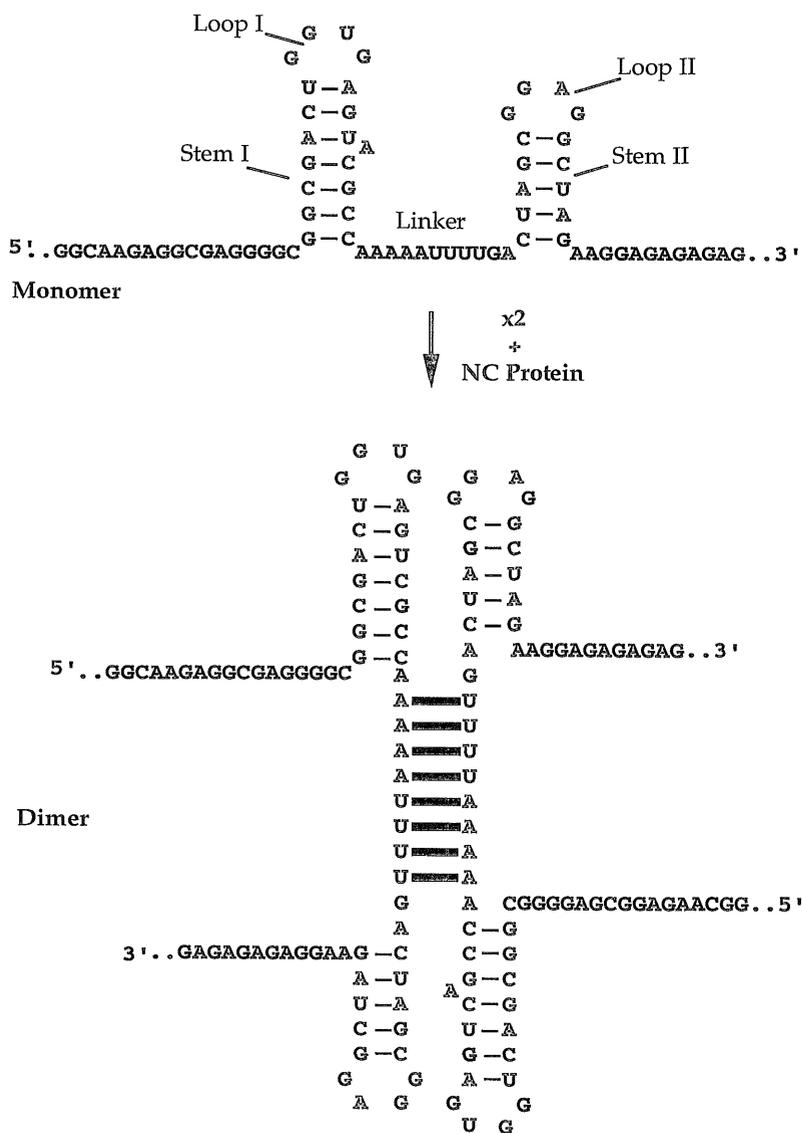


Figure 1.13. A-U base pairing mechanism of nucleocapsid induced RNA dimerisation.

#### 1.6.4. The Kissing Loop Model for HIV-1 RNA Dimerisation.

Sequences upstream from the originally defined DLS were shown to be important for dimerisation of HIV-1 RNAs *in vitro* (Marquet et al., 1994). RNAs extending from position 1-707, 1-615 and 1-311 of HIV-1<sub>MAL</sub> dimerised ten times faster than RNAs 311-612 and 311-415. However the thermal stability of the longer RNAs was reduced. Therefore transcripts lacking the sequences upstream from the splice donor (position 305 in HIV-1<sub>MAL</sub>) may not be an accurate model for RNA dimerisation (Marquet et al., 1994).

An HIV-1<sub>MAL</sub> RNA from 1-707 was analysed by chemical interference analysis which showed that the nucleotides 274GUGCAC<sub>279</sub> were required in an unmodified form to allow RNA dimerisation (Skripkin et al., 1994). This sequence was palindromic, and disruption of the self complementarity by mutagenesis eg. to 274AAACAC<sub>279</sub> prevented RNA dimerisation. The kissing loop model for HIV-1 RNA dimerisation was proposed. Dimerisation was initiated by Watson-Crick hydrogen bonding between two complementary sequences exposed at the tip of a stem loop on the two RNA monomers, figure 1.14.

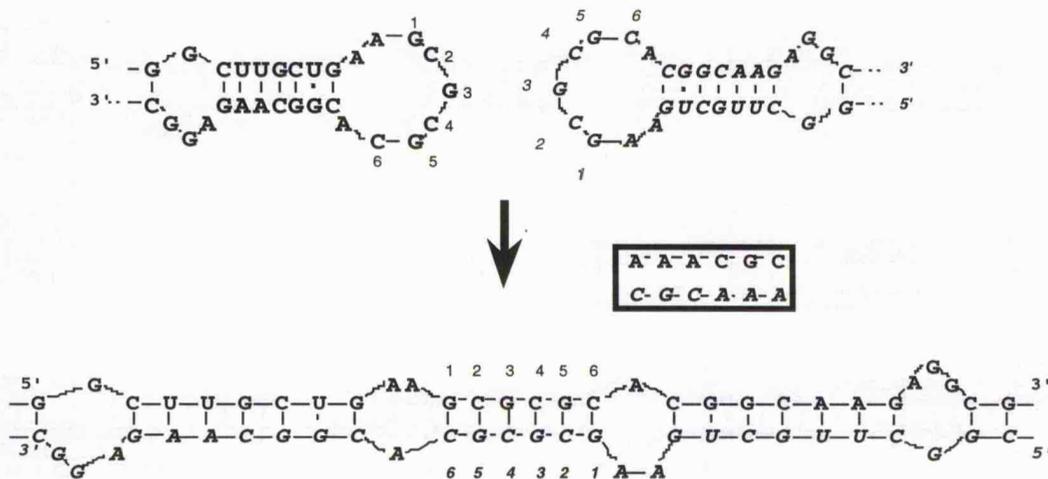


Figure 1.14. The Kissing Loop Model for RNA Dimerisation.

The initial inter-strand hydrogen bonding interaction (residues 1 to 6) resulted in intra-strand stem base pairs dissociating and reforming in the RNA dimer as inter-strand base pairs. Mutations (boxed in figure 1.14) which abolished the complementarity of residues 1 to 6 prevented dimerisation. This base pairing interaction was essential for the initiation of dimerisation (Skripkin et al., 1994), which did not require any protein cofactors. The sequences corresponding to the kissing loop were designated the DIS or dimer initiation site. This model for HIV-1 RNA dimerisation has been confirmed *in vitro* (Laughrea and Jette, 1994 and 1996; Paillart et al., 1994; Muriaux et al., 1995 & 1996). Mutagenesis of the RNA sequences in the loop or those that affected the stem structure, abolished dimerisation. The compensatory changes restored the ability to dimerise the RNA (Paillart et al., 1994). Additionally heterodimer formation between RNAs of different lengths occurred using the kissing loop structures on each RNA (Muriaux et al., 1995). A short autocomplementary sequence in the leader of the MoMLV has been identified and was important for the dimerisation of *in vitro* MoMLV transcripts (Girard et al., 1995).

To discriminate between the contribution from the purine quartets in the DLS and the kissing loop in the DIS, mutagenesis was performed in these regions (Paillart et al., 1994). RNAs containing the DIS region could dimerise, as could those restricted to the DLS. However when both the DIS and DLS were present on the same RNA, those dimers were more stable implicating two purine tracts in the DLS around the start of the gag gene which may be responsible for dimer stabilisation. Deletion of the purine rich sequences failed to prevent RNA dimerisation (Laughrea and Jette, 1994; Muriaux et al., 1995), however some reduction in thermostability

was observed for kissing loop RNAs containing tetrad mutations (Pailliant et al., 1994).

Two possible configurations of the kissing loop structure have been proposed. As described above the initial loop:loop base pairing was followed by extensive inter-strand hydrogen bond formation (Skripkin et al., 1994). Additionally the loop:loop interaction only, or loop stacking, may occur in which the only hydrogen bonding occurs between the palindromic sequences exposed in the loop (Chang and Tinoco 1994; Laughrea and Jette 1996; Muriaux et al., 1996), figure 1.15.

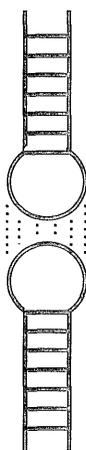


Figure 1.15. The kissing loop interaction may occur by loop stacking and hydrogen bond formation between exposed residues only.

In support of this a low stability and high stability HIV-1 RNA dimer have been characterised *in vitro* (Laughrea and Jette, 1996; Muriaux et al., 1996). The more stable conformation resisted semi-denaturing conditions which the low stability dimer could not. At physiological temperatures the low stability dimer was usually formed, as if dimerisation without nucleocapsid protein corresponded to the loop:loop interaction only. The

high stability dimer may be formed by the action of the nucleocapsid protein, eg. during dimer maturation. This may correlate to the observation of immature HIV-1 genomic RNA dimers *in vivo* which matured concomitant with particle morphogenesis (Fu et al., 1994). Another DLS 'like' domain upstream of the kissing loop has also been proposed (Laughrea and Jette, 1996).

Unlike the guanine tetrad structure, kissing loop RNA interactions have been identified which are biologically significant. A kissing loop interaction may occur between HIV-1 TAR and TAR\* RNA sequences (Chang and Tinoco, 1994). The TAR element is a hairpin structure with a six nucleotide exposed sequence at the tip of the loop, CUGGGA. Within gag coding a TAR\* structure may form, the sequence is highly conserved and has a similar loop with the UCCCAG sequence exposed. These sequences are complementary. A 'stacked' kissing loop complex may form based on evidence from gel electrophoresis, absorbance melting curves, enzymatic probing and imino proton NMR spectra. The biological role for this interaction in the retroviral life cycle is uncertain, although it could be the recognition site for a protein, similar to the replication of the ColEI plasmid.

Replication of the ColEI plasmid is regulated by an interaction between RNAs I and II. The two RNAs adopt a stem loop structure, and hybridise with each other to prevent plasmid replication by the formation of a kissing loop intermediate (Simon and Kleckner, 1988; Marino et al., 1995). The duplex formation occurs through a mechanism of gradual intra-strand breaking and inter-strand formation of hydrogen bonds (Persson et al., 1990). This structure, and not the individual loops, is specifically recognised by the protein dimer ROM (RNA-one modulator), preventing initiation of plasmid replication (Banner et al., 1987).

Another RNA:RNA interaction involving a kissing loop was identified in the recombination of Turnip Crinkle Virus RNAs (Cascone et al., 1993). Single base mutations in the stem or loop of the structure prevented recombinant formation. By phylogenetic analysis, a long range intramolecular *E.coli* 23S ribosomal RNA contact involves a kissing loop structure (Leffers et al., 1986).

### **1.7. Project Outline.**

The aim of this project was to identify nucleotides involved in the dimerisation of HIV-1 genomic RNA. Based solely on *in vitro* RNA dimerisation analysis, three mechanisms of RNA dimerisation have been proposed. The guanine tetrad model had just been proposed for retroviral RNA dimerisation at the start of this project (Sundquist and Heaphy, 1993). This was followed by the A-U base pairing (Sakaguchi et al., 1993) and later the kissing loop model (Skripkin et al., 1994). By introducing specific mutations into these RNAs, dimerisation was prevented probably by disrupting the guanine tetrad or kissing loop structures.

To determine whether nucleotides involved in the guanine tetrad or kissing loop structures *in vitro* were also involved in the dimerisation of genomic HIV-1 RNA *in vivo*, mutations were introduced into an infectious molecular clone of HIV-1<sub>NL4-3</sub>. Prior to this study, we were unaware of any point mutations which affected RNA dimerisation *in vivo*. Structural analyses of the leader region of HIV-1 either do not support (Baudin et al., 1993) or are consistent with the kissing loop structure (Harrison and Lever., 1992). Similarly isolated HIV-1 RNA did not display cation dependent thermal stability suggesting that a guanine tetrad was not part of virion RNA, although this result was not conclusive (Fu et al., 1994). Therefore to determine if these mechanisms were involved in RNA dimerisation *in vivo*, mutagenesis and RNA

dimerisation analysis was performed on infectious HIV-1<sub>NL4-3</sub> viruses (Haddrick et al., 1996a). In anticipation that mutations which abolished RNA dimerisation may prevent viral replication *in vivo*, a virus like particle system was also developed where replication was decoupled from particle assembly (Haddrick et al., 1996b).

## Chapter 2: Materials and Methods.

### 2.1 Oligonucleotides.

The oligonucleotides used are shown below. They were made in the Protein and Nucleic Acid Chemistry Laboratory, University of Leicester, using an Applied Biosystems Synthesizer model 394-08. Oligonucleotides were precipitated with ethanol and resuspended in TE. Mutagenic oligonucleotides were gel purified on denaturing polyacrylamide gels prior to use.

#### Oligo Sequence 5' to 3'.

A3 AATCCTCTAGAATTAGTCAGTGCTGGAA (4210-4227 CSF)  
A4 TCGAGGATCCTGCTTCTATATATCC (4227-4210 CSF)  
Act 1 CGGGACCTGACTGACTACC (588-606 actin)  
Act 2 CTGTGTGGACTTGGGAGAG (1522-1504 actin)  
AL2 TCAGAATTCATTTAGGTGACTACTATAGAACGGCGACTGGTGAGTACGCC (732-754)  
MH6 TCAGAATTCAATACGACTCACTATAGGGACTTGAAAGCGAAAAG (649-666)  
MH7 TCAGAATTCATTTAGGTGACTACTATAGAAGCGCGCACGGCAAGA (708-725)  
MH8 TCAGAATTCAATACGACTCACTATAGGCAAGAGGCGAGGGGCG (719-736)  
MH10 TCGAAGCTTGGCCTTAACCGAATTTTTCCATTATCTAATTCTC-  
CACCGCTTAATACC (858-807)  
MH14\* CAGTATTAAGCGGTGGGAGAATTGGATAGG (806-834 CSF)  
MH17 CGCCTCTGCCGTGCGTTTTTCAGCAAGCCGAGTCCTG (729-692)  
MH18 GCGAGGGGCGGCGACTGG (727-744)  
MH22 GGTATTAAGCGGTGGGAGAATTAGATAAAATG (807-836)  
MH23 CTGAAGGGTACTGGTAGTTCCTGCTATG (1521-1494)  
MH24 TCGGCTTGCTGAAAAACGCACGGCAAGAGGCG (698-729)  
PCR5 CCAACCCACAAGAAGTAGTATTGG (6468-6491 CSF)  
PCR7 AGTAGCTGAGGGGACAGATAG (8796-8816 pBCCX CSF X4)  
PCR20 GAATCGCAAAAACAGCCGGGGCAC (6905-6882 CSF)  
PCR31 GCAGAGCTCGTTTGTGAACC (656-676 pBCCX CSF X4)  
SH29 TCAGAGCTCGGTCTCTCTGGTTAGACC (454-472)  
SH175 TCGAAGCTTGGCCTTAACCGAATTTTTTC (858-838)  
SH241 TCGAAGCTTAATACCGACGCTCTCG (813-797)  
WS109 TCGGAATTCAATACGACTCACTATAGGGCGGCGACTGGTGAGTAC (732-751)

Oligonucleotide sequences are shown along with their annealing positions to HIV-1<sub>NL4-3</sub> DNA (unless otherwise stated). Forward primers are indicated by increasing numerical annealing positions eg. SH29 (454-472). For backward primers the order is reversed eg. SH175 (858-838). Sequences between the underlined residues annealed to the target. Other sequences correspond to restriction sites or RNA polymerase promoters, eg. MH6

▼ Eco RI ▼ T7 RNA polymerase promoter ▼ Target sequence hybridisation ▼  
 5'-TCA GAATTC TAATACGACTCACTATA GGGACTTGAAAGCGAAAG -3'

## 2.2 Escherichia coli Strains and Cell Lines.

*E. coli* strains used are shown below:

<u><i>E.Coli</i></u>	<u>Source</u>	<u>Genotype</u>
DH5 $\alpha$	Gibco BRL	<i>supE44 DlacU169(<math>\phi</math>80 lac ZDM15) hsdR17 recA1 end A1 gyrA96 thi-1 relA1</i>
JM109	Promega	<i>recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi <math>\Delta</math> (lac-pro AB)</i>
XL1 mutS	Stratagene	<i><math>\Delta</math>(mcrA) 183 <math>\Delta</math>(mcrCB-hsdSMR-mrr) 173 end A1 supE44 thi-1 gyrA96 relA1 lac mutS::Tn10 (Tet<sup>r</sup>) [F' proAB lacI<sup>q</sup>Z<math>\Delta</math>M15 Tn5 (Kan<sup>r</sup>) ]</i>
XL1 Blue	Stratagene	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI<sup>q</sup> Z<math>\Delta</math>M15 Tn10 (Tet<sup>r</sup>) ]</i>

### Cell Lines.

The mammalian cell lines used are described below. Cell lines were obtained from the MRC AIDS Directed Programme reagent repository, NIBSC, Potters Bar.

<u>Cell Line</u>	<u>Source</u>	<u>Characteristics</u>
COS	MRC	African Green Monkey Kidney cell line, which has SV40 large T antigen. (Gluzman, 1981)
C8166	MRC	Human T-lymphoblastoid cell line. Expresses tat from HIV-1. (Salahuddin et al., 1983 and Lee et al., 1984)

### **2.2.1. Growth Conditions.**

#### *E.coli* culture details.

Media used to grow *E.coli* was prepared using distilled water and autoclaved at 121°C for 20 min.

#### 2TY Broth (1 litre)

Yeast extract (Oxoid)	10 g
Tryptone (Oxoid)	16 g
NaCl (BDH)	5 g

Ampicillin was added to 100µg/ml final concentration before use. For 2TY agar, 15g of bacteriological agar (Oxoid) was included per litre before sterilisation. *E.coli* cultures were initiated by the addition of single colonies or an appropriate dilution of an overnight culture. Flasks were incubated at 37°C and agitated at 200 rpm. The growth of the culture was followed by measuring the change in optical density at 595nm.

### Establishment of Cell Cultures.

All cell lines were initiated from frozen cell stocks stored under liquid nitrogen. The cells were thawed to room temperature and were added dropwise into fresh medium. They were then resuspended three times in an excess of fresh medium to remove any traces of dimethylsulphoxide.

### Maintenance of Cell lines.

Cell lines were grown in Nunculon tissue culture flasks at 37°C, 5% CO<sub>2</sub> in a humid incubator under containment level III conditions. African green monkey COS cells (an adherant cell line) were grown in Dulbecco's Modified DMEM medium, supplemented with 10% (v/v) heat-inactivated newborn calf serum (NCS), penicillin (100U/ml), streptomycin (100mg/ml) and 2mM glutamine (Gibco BRL). At 80% confluency cells were trypsinised, counted, and resuspended in fresh medium. Human T-cell lines were maintained in RPMI 1640 medium, with 10% (v/v) heat inactivated foetal calf serum (FCS), penicillin (100U/ml), streptomycin (100mg/ml), and 2mM glutamine.

### **2.2.2. Counting cells.**

Cells were counted and their viability determined by trypan blue stain exclusion. 0.2ml of the cell suspension cells was mixed with an equal volume of trypan blue stain (Gibco BRL) and placed into the haemocytometer. The cell count/ml = no. of cells counted in quadrants x 2 x 2500. Cells were maintained at 0.5 to 1 x 10<sup>6</sup> cells per ml.

### **2.2.3. Storing cells.**

Cells were pelleted by a 5 min spin at 1,250g. They were washed once in PBS then pelleted again. The cell pellet was resuspended in 1 ml of freezing solution (0.9 ml fresh medium, 0.1 ml of DMSO) and placed into a

bio-freeze vial (Costar). This was incubated overnight in dry ice and was then transferred to liquid nitrogen.

### **2.3. DNA Techniques**

#### **2.3.1. Agarose Gel Electrophoresis of DNA.**

<u>1X TBE.</u>	<u>Loading Buffer</u>
90 mM Tris pH 8.0	0.25% Bromophenol Blue
90 mM Boric Acid	0.25% Xylene Cyanol
2 mM EDTA	15% Ficoll (MW. 400,000 Pharmacia)

#### 1X TAE

40 mM Tris-Acetate pH 8.3  
1 mM EDTA

Gels were made using Pharmacia agarose containing 0.5 ug/ml ethidium bromide and run in 1X TBE buffer at a constant voltage of 40-80 V. DNA was mixed with loading dye prior to electrophoresis and run alongside molecular weight markers. Markers used were HindIII cut phage  $\lambda$  DNA, or HaeIII digested  $\phi$ X174 DNA, obtained from GIBCO BRL.

#### $\lambda$ Hind III markers.

23,130; 9,416; 6,557; 4,361; 2,322; 2,027; 564 and 125 bp.

#### $\phi$ X174 HaeIII markers.

1,353; 1,078; 872; 603; 310; 281; 271; 234; 194; 118 and 72 bp.

#### **2.3.2. Restriction Digestion.**

Restriction digestion of DNA was performed according to the enzyme manufacturer's instructions (GIBCO BRL). Typically this was the addition of a 5 fold excess of enzyme to the DNA solution followed by digestion at

37°C for at least 2 hours in a 1X buffer, diluted from the 10X buffer supplied. The glycerol concentration was kept to less than 5% of final volume. Restriction enzymes used are listed below.

<u>Enzyme</u>	<u>REact Buffer</u>
BamHI	3
BssHII	2 (incubate at 55°C)
EcoRI	3
HindIII	2
SacI	2
SpeI	4
SphI	6
XhoI	2

REact Buffers at 1X concentration.

REact 2	50mM Tris pH 8, 10mM MgCl <sub>2</sub> , 50mM NaCl
REact 3	50mM Tris pH 8, 10mM MgCl <sub>2</sub> , 100mM NaCl
REact 4	20mM Tris pH7.4, 5mM MgCl <sub>2</sub> , 50 mM KCl
REact 6	50mM Tris pH7.4, 6mM MgCl <sub>2</sub> , 50mM KCl, 50mM NaCl

After digestion the restriction enzymes were inactivated either by heating at 65°C for 10 min, or by phenol/chloroform extraction. DNA was recovered by ethanol precipitation.

**2.3.3. Phenol/Phenol Chloroform Extraction.**

To denature proteins associated with DNA an equal volume of phenol (equilibrated with 0.1M Tris pH 8.0), or phenol:chloroform:isoamylalcohol (25:24:1) was added to the sample. The mixture was vortex mixed for approx 30sec then spun in a microfuge at 13,000g for 1 min. The

upper aqueous layer containing the DNA was carefully removed, and the procedure repeated as required.

#### **2.3.4. Ethanol Precipitation.**

DNA or RNA was precipitated out of solution by the addition of 0.1 times the sample vol of 3M sodium acetate (pH 5.2), followed by 2.5 times sample volume of ethanol. After vortex mixing the tube was chilled on dry ice for 15min, then spun at 13,000g for 10min in a microcentrifuge (microfuge). The supernatant was removed and the pellet washed with a 70% ethanol solution. After an additional 5 min spin all ethanol was removed, and the tube inverted to dry the pellet on the bench for 15-20 min. The sample was resuspended in TE or nanopure water.

#### **2.3.5. Gel purification of DNA from Agarose Gels.**

DNA was recovered from TAE agarose gels using GeneClean, according to the manufacturers instructions. Briefly the DNA band was excised from the gel, and excess agarose removed. Three volumes of 3M sodium iodide was added and the agarose dissolved by incubation at 60°C for 5 min. The resin slurry used to bind DNA was added, and left at room temperature for 5 min. The DNA and resin were pelleted by brief centrifugation and the supernatant removed. The pellet was resuspended and wash buffer applied. This was repeated an additional two times. Finally TE was added and the DNA eluted from the resin by a 5 min incubation at 65°C. The DNA was then ethanol precipitated and recovered.

#### **2.3.6. Ligations**

Ligations were set up in 20µl of 50mM Tris pH 7.6, 5 mM MgCl<sub>2</sub>, 1 mM ATP, 1mM DTT, and 25% (wt/vol) PEG-8000 with 50ng of vector DNA. Usually a 3:1 molar excess of insert:vector was included. One unit of T4

DNA Ligase was added (Gibco BRL) and the reaction incubated overnight at 4°C. 10µl of the ligation mix was usually transformed into 150µl of transformation competent *E.coli*.

#### **2.3.7. Preparation of Transformation Competent *E.Coli*.**

To prepare transformation competent *E.coli* 250ml of 2TY was inoculated with a single colony or a 1:50 dilution of an overnight culture of *E.coli* DH5α. The cells were grown at 37°C to an A<sub>595nm</sub> of 0.6. All of the following stages were done on ice. The cells were spun down at 5,000 rpm for 5min in a Sorvall GS-3 rotor. The pellet was resuspended in 125ml of sterile cold 10mM NaCl. Another spin at 5,000 rpm for 5min was followed by resuspension in 62.5ml of cold 30mM CaCl<sub>2</sub> and incubation for 20min on ice. Finally the cells were pelleted and resuspended in 10ml of cold 30mM CaCl<sub>2</sub> containing 15% glycerol. The transformation competent *E.coli* were aliquotted into sterile eppendorfs and snap frozen in a dry ice/ethanol bath and stored at -70°C.

#### **2.3.8. Transformation of *E.coli*.**

150µl of frozen competent *E.coli* were thawed on ice. The DNA in a volume of less than 20µl, was added to the cells then left on ice for 10min. The tubes were heat shocked for exactly 2min at 42°C then returned to ice for an additional 10min. 200µl of 2TY medium (without antibiotic) was added and the transformants were incubated at 37°C for 30min. After another 10min on ice, 150µl of the transformants were plated out onto antibiotic containing medium.

#### **2.3.9. Preparation of plasmid DNA from *E.coli*.**

Plasmid DNA was prepared from *E.coli* by the procedures described below.

Alkaline Lysis and PEG Precipitation.

<u>GTE</u>	<u>NaOH/SDS</u>	<u>Acetate Solution.</u>
50mM Glucose	0.2 M NaOH	3M Potassium Acetate
25mM Tris pH 7.4	1.0% SDS	2M Acetic Acid
10mM EDTA		

A 500ml overnight culture of *E.coli* carrying a plasmid was pelleted by spinning at 6,000rpm for 10 min in a Sorvall GS-3 rotor. The pellet was resuspended in 7.5ml of GTE solution and incubated on ice for 10min. To lyse the cells 15ml of the NaOH/SDS solution was added followed by a further 10min incubation on ice. For neutralisation 11ml of the Acetate solution was added and incubated as above. The mixture was then spun at 15,000 rpm for 30min at 4°C in a Sorvall SS34 rotor. The cleared supernatant was decanted off and 17ml of propan-2-ol was added. Total nucleic acids were allowed to precipitate for 15min at room temperature then spun at 10,000rpm for 20min. The supernatant was discarded and the pellet resuspended in 3ml of TE buffer. Large RNAs were precipitated out of solution by the addition of 3ml of 5M LiCl followed by a 10,000 rpm spin for 10min at 4°C. The remaining nucleic acid was again precipitated by the addition of an equal volume of propan-2-ol and centrifugation at 10,000 rpm for 10min at room temperature. The supernatant was removed and the pellet washed with 70% ethanol, then air dried. It was then resuspended in 0.5ml of TE (10mM Tris pH 7.5, 1mM EDTA), containing RNaseA (BDH) at 20µg/ml, and incubated for 15min at 37°C. An equal volume of 1.6M NaCl/13% PEG 8,000 was added and the tube contents were thoroughly mixed. After spinning at 13,000g for 5min in a microfuge the pellet was resuspended in 400µl of TE. Following one phenol, phenol/chloroform then chloroform extraction, plasmid DNA was

precipitated by the addition of 0.1ml of 10M Ammonium Acetate and 1ml of cold absolute ethanol. After a 13,000g spin for 10min (microfuge) and a 70% ethanol wash, the DNA was dried and resuspended in TE. DNA concentration was determined by the absorbance of the solution at 260nm and it's purity assessed from the  $A_{260}/A_{280}$ nm ratio.

#### Caesium Chloride Procedure.

<u>ST</u>	<u>Triton lysis Mix</u>
25%(w/v) Sucrose	1.5%(v/v) Triton X 100
50mM Tris pH 8.0	50mM Tris pH 8.0
	50mM EDTA pH 8.0

From 500ml of an overnight culture the cells were harvested by spinning at 5,000 rpm for 10min at 4°C in a Sorvall GS-3 rotor. The pellet was resuspended in a final volume of 2.5ml of ST buffer. 1ml of a 20mg/ml lysozyme (BDH) in ST solution was added and incubated on ice for 10min. Next 1.5ml of 0.5M EDTA was added and incubated as above. The cells were lysed by the addition of 2.5ml of Triton Lysis mix with gentle mixing. After a spin at 20,000 rpm for 45 min at 4°C in a Sorvall SS34 rotor, the cleared supernatant was carefully removed. The approximate volume of the supernatant was determined and CsCl was added at 1g per ml. After a 5min incubation at 37°C the tubes were spun at 2,500 rpm for 5min at 20°C and the supernatant recovered. This was added to 3.9ml Beckman heat seal centrifuge tubes along with 25µl of a 10mg/ml ethidium bromide solution. The tubes were spun in a Beckman Ti75 rotor either overnight at 80,000, or at 100,000 rpm for 4-5 hours at 20°C. The plasmid DNA band was extracted using 20 gauge needles by side puncture using a 2ml syringe. The ethidium bromide was removed by extraction with an equal volume of water/CsCl saturated propan-2-ol until this solution was colourless. The

sample was diluted with 3 volumes of TE and 2 volumes of ethanol. Plasmid was recovered by centrifugation at 2,500 rpm for 30min at 4°C in an MSE Mistral bench top centrifuge. The pellet was washed with 70% ethanol, then dried and resuspended in TE. The resulting DNA was quantitated by measuring the  $A_{260\text{nm}}$  and its purity assessed by measuring the  $A_{260}/A_{280\text{nm}}$  ratio.

#### Miniprep Procedure.

For small scale preparations of plasmid DNA, minipreps were performed using the triton lysis method.

#### STET

8%(w/v) Sucrose

50 mM EDTA pH 8.0

50 mM Tris pH 8.0

0.5% Triton X 100

1mg/ml Lysozyme (BDH)

1.5ml of an overnight culture was taken and the cells pelleted by a 1 min spin at 13,000g in a microfuge. The supernatant was completely removed and the cells resuspended on ice in 200µl of STET buffer. The cells were then placed into a boiling water bath for 2min, then rapidly plunged into ice and left for 10min. After centrifugation at 13,000g in a microfuge for 10min, the resulting cell debris pellet was removed using a toothpick. The remaining supernatant was adjusted to 100µl with TE. Total nucleic acids were precipitated by adding 10µl of 10M Ammonium acetate and 200µl of ethanol, incubating on ice for 10min then spinning at 13,000g for 10min.

The pellets were washed with 70% ethanol then air dried and redissolved in 20  $\mu$ l of TE containing 50ug/ml RNaseA.

### 2.3.10. Polymerase Chain Reaction (PCR).

#### Estimation of Primer Concentration.

Primer concentrations were determined by measuring the  $A_{260nm}$  of primer dilutions to within the 0.1 to 1.0 range. The number of A, C, G and T residues were scored and the concentration in pmol/ $\mu$ l was determined calculated from the absorbance of a 1,000 pmol/ $\mu$ l solution of each base.

eg. MH14\*

Sequence 5'- CAGTATTAAGCGGTGGAGAATTGGATAGG -3'

$A_{260nm}$  (100x dilution) = 0.425

9 A @ 15.4 = 105.3

11 G @ 11.7 = 80.3

2 C @ 7.3 = 17.6

7 T @ 8.8 = 107.8

$\Sigma$  311.0

Therefore [MH14\*] =  $\frac{42.5}{311.0} \times 1000 = 136.6$  pmol/ $\mu$ l (136.6  $\mu$ M)

#### Polymerase Chain Reaction (PCR)

The polymerase chain reaction (Saiki et al., 1986) was performed to amplify target DNA between a 5' and 3' primer. PCR reactions used either Taq DNA polymerase (Promega) or Vent DNA polymerase (NEB). For the detection of RNA within the VLPs Taq DNA polymerase was used. For other uses where sequence fidelity was important, eg. the production of DNA for cloning, Vent DNA Polymerase was used due to its lower

misincorporation rate. Optimisation of Mg<sup>2+</sup> concentration and annealing temperatures were performed for each primer pair.

	<u>Taq</u>		<u>Vent</u>
10X Buffer	5.0µl		5.0ul
MgCl <sub>2</sub>	as reqd.	MgSO <sub>4</sub>	as reqd.
DNA Template	1-20ng		1-20ng
5' Primer	25µM		25µM
3' Primer	25µM		25µM
4 dNTPs (2.5mM)	5.0µl		5.0µl
Water	to 50µl		to 50µl
DNA polymerase	0.5µl		0.5µl

Taq 1 X buffer: 50 mM KCl, 10mM Tris pH (9.0), 1.0 % Triton X 100.

Vent 1X buffer: 10 mM KCl, 20mM Tris pH (8.0), 0.1 % Triton X 100,  
2 mM MgSO<sub>4</sub>, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>

The reactions were set up on ice then overlaid with an equal volume of mineral oil (Sigma). For a hot start PCR the enzyme was added after the tubes had been incubated at 94°C for 1 min. PCR cycles were then performed, typically 94°C for 1 min denaturing, 45-60°C for 1 min annealing, followed by 72°C for 1 min per kilobase amplified. After the desired number of cycles the reaction was held at 55°C for 5 min. 5µl of the resulting mixture was electrophoresed on agarose gels. The PCR product was recovered by organic extraction and ethanol precipitation.

### 2.3.11. Mutagenesis.

Three methods were used to introduce site directed mutations:

1. PCR using mismatch containing primers.
2. RT-PCR using a mutagenic primer for cDNA synthesis followed by PCR.
3. LP-USE using mutagenic primers to generate a 'long' mutagenic primer PCR product containing the mutation and restriction site elimination changes.

#### 1. PCR Mutagenesis using mismatch containing primers.

RNA 12 was made using an oligonucleotide primer containing the mutation (AL2) along with the reverse SH175 primer and HIV-1<sub>NL4-3</sub> plasmid DNA template. The forward primer (AL2) contained an SP6 promoter 5' of the HIV-1 sequences, along with an EcoRI restriction site. SH175 was the reverse primer and contained a HindIII site. The amplified PCR product was cloned into Eco-HindIII cut pUC119 (Messing, 1983). Cloned sequences were verified by DNA sequencing using Sequenase.

#### 2. RT-PCR mutagenesis

Plasmids encoding mutant RNAs 3, 5 and 7 were made by an RT-PCR approach. RNA 3 was made using oligonucleotide MH10 and RNA 2 as a template. RNA 5 was made using oligonucleotide MH17 and RNA 4 as a template, and RNA 7 was made using MH17 and RNA 6 template. The method is illustrated for RNA 7 in figure 2.1.

Oligonucleotide MH17 containing the desired mutation was annealed to *in vitro* transcribed RNA 6 and used as a primer for cDNA synthesis by MLV reverse transcriptase (1). 0.5 ug of RNA 6 was dissolved in 50ul of 50mM KCL, 10mM Tris pH 8.3, 1.7 mM MgCl<sub>2</sub>, 200 uM dNTPs and 1.25 uM

MH17. This was heated to 90°C for 2 min and cooled on ice. 200 units of MLV RT (Gibco BRL @ 200 U/ul), and 10 units of RNasin were then added and the reaction incubated at 37°C for 60 min. Following heat denaturation of MLV RT at 90°C for 5 min (2), 0.5ul of the cDNA containing the desired mutation was used as a template in an asymmetric PCR reaction along with oligonucleotide MH6. After 20 cycles an extended forward primer containing the mutation was generated (3). This was then used in a standard PCR reaction using HIV-1<sub>NL43</sub> DNA as a template and a reverse oligonucleotide eg. SH241 for RNA 5 and SH175 for RNAs 3 and 7. The PCR products were cloned into pUC119 and the DNA sequences verified eg. producing pMH95.003 for RNA 7.

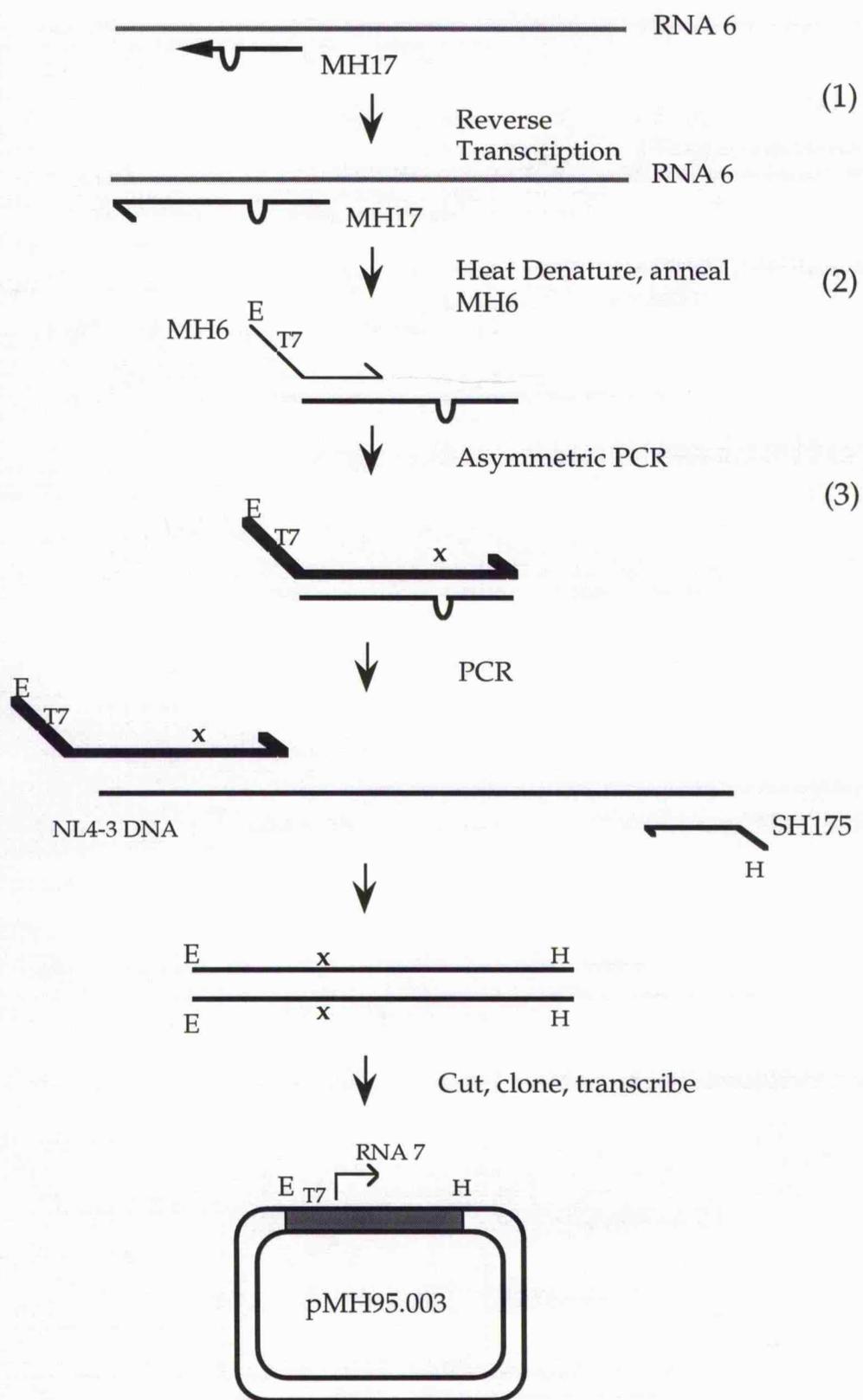


Figure 2.1. RT-PCR mutagenesis.

### 3. LP-USE Mutagenesis.

Mutations were introduced into pUCΔNL by LP-USE (Long PCR - Unique Site Elimination) mutagenesis (Ray and Nickloff, 1992), figure 2.2.

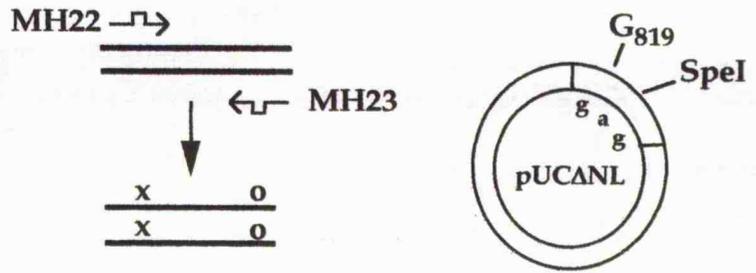
PCR was performed (1) using primers which introduced the desired mutation (G<sub>819</sub> to T) and eliminated a unique restriction site eg. SpeI at 1507-1512 in pUCΔNL (T<sub>1509</sub> to C). 0.5 pmol of this primer (2) was mixed with 0.025 pmol of pUCΔNL DNA in 10mM Tris pH 7.5, 10mM MgCl<sub>2</sub> and 50mM KCl. The mixture was held at 100°C for 3 min, then immediately cooled on ice. The extension, ligation and selection of plasmids carrying the desired mutations was performed according to the manufacturer's instructions (Stratagene Chameleon Mutagenesis).

Following DNA synthesis and ligation (3) from the primer containing the mutations, mutated sequences were selected for by digesting the reaction with SpeI (4). This was performed prior to the first transformation into *E.coli* XLmutS (5) and removes any wild type plasmid. Following transformation into XLmutS *E.coli* and after growth at 37°C overnight, miniprep DNA (6) was recovered from the transformant pool. This was again restricted with SpeI to remove wild type sequences (7). Finally the reaction mixture was transformed (8) into *E.coli* XL-1 Blue competent cells, and miniprep DNA prepared. Mutated viral sequences and the absence of any other changes in the amplified region, were identified by restriction digestion and confirmed by sequencing of the resulting plasmid DNA.

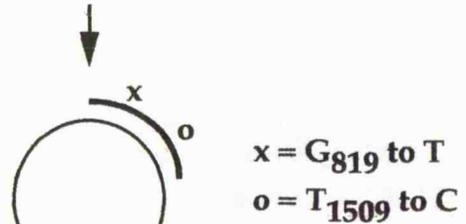
### 2.3.12. DNA Sequencing.

Sequencing of DNA was performed using Sequenase ver 2.0 (Amersham) according to kit instructions, based on the chain termination dideoxy method (Sanger et al., 1977). Double stranded DNA was made single stranded by alkali denaturation. 10µg of plasmid DNA was incubated in 0.4M NaOH and 0.25 mM EDTA for 30 min at 37°C. The denatured DNA

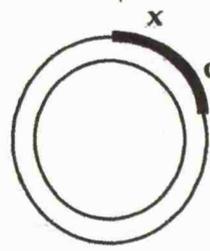
**1.** PCR using the mutagenic oligos MH22 (G819 to T) and MH23 which removes the SpeI site (T1509 to C).



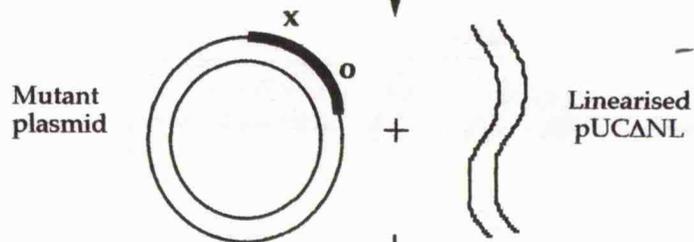
**2.** Denature and anneal the long mutagenic primer.



**3.** Incubate with T7 DNA polymerase and T4 DNA ligase to extend and ligate.

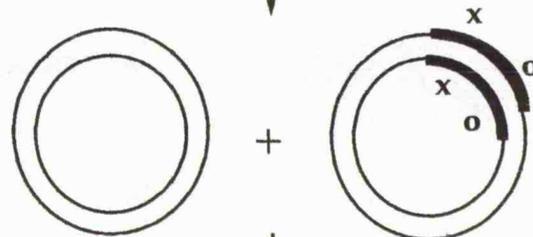


**4.** Perform a SpeI restriction digestion to linearise the parental plasmid.

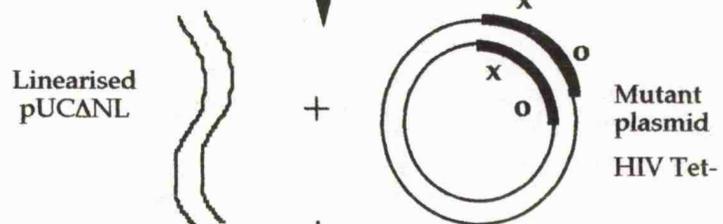


**5.** Transform into competent XLmutS *E. coli* and grow in liquid culture.

**6.** Recover DNA from the pool of transformants.



**7.** Perform a second SpeI digestion.



**8.** Transform into XL1-Blue *E. coli* and analyse the colonies by restriction digestion/sequencing of miniprep DNA.

Figure 2.2. LP-USE Mutagenesis.

was precipitated by the addition of ammonium acetate to 0.9M and the addition of 2.5x volume of ethanol. The tube was placed in dry ice for 15 min, spun at 13,000rpm in a microfuge for 10 min then washed with 70% ethanol. The pellet was dried and finally resuspended in TE so that the DNA concentration was approx 0.4 µg/µl. 2µl of this DNA was used in the sequencing reactions, according to the kit instructions.

The reactions were analysed on a 6% denaturing polyacrylamide gel with a TBE buffer gradient.

<u>40 % Acrylamide Stock (1l)</u>	<u>6% Top (1l)</u>	<u>6% Bottom (500ml)</u>
380 g Acrylamide	150 ml 40% stock	75 ml 40% stock
20 g Bis-acrylamide	50 ml 10X TBE	125 ml 10X TBE
	460g Urea	230g Urea

Clean glass plates (20 x 39cm) were separated by 0.3mm spacers and taped together. 32ml of top solution and 14ml of bottom solution (coloured with a little Bromophenol Blue) was measured out. 64µl and 28µl of 25% AMPS, and 64µl and 28µl of TEMED were added to each respectively. 25ml of the top solution was drawn up into a 25ml glass pipette, then 7ml of bottom solution was carefully drawn in. Several bubbles were allowed into the pipette to form the TBE gradient. The gel was cast, the comb inserted and the gel left to polymerise. Before loading the comb was removed and the wells were thoroughly washed.

Gels were electrophoresed at a constant 37W. The plates were then separated and the gel fixed by immersion in a 10% IMS/10% Acetic acid solution for 10min. The gel was transferred onto Whatman 3MM paper and dried under vacuum at 80°C for 1 hour. The bands were visualised by autoradiography.

## **2.4. RNA WORK.**

### **2.4.1. Precautions for working with RNA.**

To minimise RNA degradation by RNAses care was taken during RNA manipulations. All tubes and tips were autoclaved then baked overnight at 100°C, gloves were worn and regularly changed during all procedures. Dedicated gel tanks etc. were cleaned with a 3% H<sub>2</sub>O<sub>2</sub> solution and rinsed well before use. RNA was stored as an ethanol precipitate at -70°C until required and then resuspended into DEPC treated water or other solutions made with DEPC treated water. (DEPC was added to 0.1% (v/v) in nanopure water, vigorously mixed, then left for several hours at room temperature. To remove unreacted DEPC the solutions were autoclaved). RNAsin (Promega) a protein which binds and inactivates RNAses, was included in some experiments.

### **2.4.2. *In vitro* RNA Transcription.**

Plasmid templates were linearised with HindIII giving all the RNAs the 3' sequence AGCU, unless otherwise indicated. Cut plasmids were then transcribed with T7 or SP6 RNA polymerase using Ribomax Kits (Promega) according to the manufacturer's instructions. RNAs were recovered by extraction with phenol (RNA grade, pH 4.3 Sigma) and ethanol precipitated. Some RNAs were additionally purified on denaturing polyacrylamide gels. The bands were visualised by UV shadowing, excised and eluted by soaking in 0.5M sodium acetate pH 5.2, 1mM EDTA, 1% (wt/vol) SDS. RNA was finally recovered by ethanol precipitation and the concentration determined from the A<sub>260nm</sub> reading.

#### **2.4.3. In vitro RNA Dimerisation.**

The dimerisation ability of RNAs from the DIS/DLS of HIV-1 was analysed by incubating the RNA recovered after transcription under two sets of conditions:

##### **Buffer I. (Without Peptide).**

RNAs were incubated (at 1-5 $\mu$ M) in standard dimerisation buffer I (10mM Tris pH7.9, 50mM MgCl<sub>2</sub>, 250mM NaCl and 0.1% SDS) for time, t (Sundquist and Heaphy, 1993). Loading buffer was added, and the samples electrophoresed.

##### **Buffer II. (With peptide).**

RNAs were incubated (at 1-5 $\mu$ M) in low salt buffer II (25mM Tris pH 7.5, 60 mM NaCl, 0.2mM MgCl<sub>2</sub>) for time, t (DeRocquigny et al. 1992). Loading buffer was added (1%SDS, 10mM EDTA, 20mM Tris pH 7.5, 20% glycerol, 0.01% Bromophenol Blue), phenol extraction was performed if required, then the samples were electrophoresed.

Both native and denatured RNAs were analysed by electrophoresis on non-denaturing agarose or polyacrylamide gels run in TBE buffer. Denatured RNA samples were prepared by adding an equal volume of formamide loading dye (Sambrook et al.) and heating at 100<sup>o</sup>C for 2 min. RNAs were visualised under UV light after staining gels in 0.5 $\mu$ g/ml ethidium bromide in water, or after fixing and drying polyacrylamide gels by autoradiography.

#### **2.4.4. Thermal Stability of RNA Dimers.**

RNA dimers formed *in vitro* or viral RNA was resuspended in 10mM Tris pH 7.5, 100mM MCl, and 5mM MgCl<sub>2</sub> where M = K<sup>+</sup>, Na<sup>+</sup>, Cs<sup>+</sup> or Li<sup>+</sup> ions. Aliquots were taken and incubated at temperatures from 4 to 75<sup>o</sup>C for

5 min, ice cooled and prepared for electrophoresis as described (Lear et al., 1995).

#### **2.4.5. Radiolabelling of RNA.**

5' <sup>32</sup>P end labelled RNAs were prepared by incubating 200 pmol of *in vitro* transcribed RNA in 1X buffer (1mM ZnCl<sub>2</sub>, 1mM MgCl<sub>2</sub>, 10mM Tris pH 8.3, 40U RNasin (Promega)) with 4 U calf intestinal alkaline phosphatase (Boehringer) for 30 min at 50°C. Following phenol extraction and ethanol precipitation, the dephosphorylated RNA was kinased using 20 U of Polynucleotide Kinase (Boehringer) in 1X buffer (50mM Tris pH 8.2, 10mM MgCl<sub>2</sub>, 0.1mM EDTA, 5mM DTT and 0.1mM spermidine) with 5ul of γ<sup>32</sup>P dATP (370 MBq/ml; 111 TBq/mmol, Amersham) for 30 min at 37°C. Unincorporated counts were removed by double ethanol precipitation. The RNA was resuspended in TE.

#### **2.4.6. Dimethylsulphate (DMS) Modification Protection of RNA.**

##### **Dimethyl Sulphate Reaction.**

5' <sup>32</sup>P-labeled RNAs were incubated under dimerisation conditions for 90 min. Monomers and dimers were separated by electrophoresis on 4% or 5% nondenaturing polyacrylamide gels. Monomer was eluted into 80mM HEPES (pH 7.5), 1mM EDTA, 100U/ml RNasin, ethanol-precipitated, and resuspended in the same buffer. Dimer was eluted into 80mM HEPES (pH 7.5), 5mM MgCl<sub>2</sub>, 25mM KCl, 100U/ml RNasin. Monomers and dimers were both modified under native and denaturing conditions. Modifications were carried out essentially as described (Peattie and Gilbert 1980). 5×10<sup>5</sup> cpm of RNA (approximately 1 pmol) was mixed with 10μg yeast tRNA in 30μl of monomer or dimer buffer. 0.3μl of 1M MgCl<sub>2</sub> was added to monomer samples being modified under native conditions. 0.6μl of

0.5M EDTA was added to dimer samples being modified under denaturing conditions, which were heated at 90°C for 5 min and rapidly cooled. DMS was diluted 1:20 in ice-cold ethanol and 2µl added to the reaction mixes at 0°C. Reactions under native conditions were carried out at 37°C for 6 to 18 min. Reactions under denaturing conditions were carried out at 90°C for 15 to 60 sec. Reactions were terminated by placing at 0°C and adding 30µl of solution containing 2M β-mercaptoethanol, 0.6M sodium acetate (pH 5.2), 2 mM EDTA. Reactions were subsequently ethanol precipitated.

#### Aniline Cleavage.

Dried pellets were resuspended in 20µl freshly made ice-cold 200mM NaBH<sub>4</sub> and incubated in the dark at 0°C for 30 min. The RNA was precipitated by adding 200µl of 600mM sodium acetate/600mM acetic acid and 600µl of ethanol on ice. The dried pellet was resuspended in 20µl of 1M aniline (prepared by mixing 558µl H<sub>2</sub>O, 60 µl aniline, and 36µl glacial acetic acid) and incubated in the dark at 60°C for 20 min. The reaction was terminated by snap freezing at -70°C. Aniline was removed by repeated lyophilization. Pellets were resuspended in 10µl of deionized formamide and prepared for electrophoresis on 7M urea polyacrylamide gels. Gels were exposed to film.

#### 2.4.7. RT-PCR

This is a single tube method, in which all of the cDNA reaction is substrate for the PCR amplification. The method described here was used for detection of RNA within the VLPs and wild type virus particles. The reverse transcriptase stage was identical to that used for RT-PCR mutagenesis. The mixes are shown for 1 sample containing RNA and were scaled accordingly.

<u>2X PCR Buffer.</u>	<u>3' Primer Supermix.</u>	<u>5' Primer/Taq DNA pol mix.</u>
100 mM KCl	25.0 µl 2X PCR	10.0 µl 2X PCR
20 mM Tris-Cl	2.5 µl 3' primer (25 µM)	2.5 µl 5' primer (25 µM)
3.4 mM MgCl <sub>2</sub>	17.5 µl Water	0.5 µl Taq pol (Promega 10U/ul)
400 µM dNTPs		7.0 µl Water

MLV RT Supermix.

<u>-RT</u>	<u>+RT</u>
5.0	5.0 µl 2X PCR
	1.0 µl AMV RT (Gibco 200 U/µl)
0.25	0.25 µl RNAsin (Promega 40 U/µl)
4.75	3.75 µl Water

The RNA was dissolved in 5µl water and kept on ice. 45µl of the 3' mix was added and the 3' primer was annealed by an incubation at 90°C for 2 min and rapidly placed into ice. cDNA synthesis was initiated by adding 10µl of the MLV RT mix which was then held at 37°C for 60 min. The RT was denatured by a 90°C for 5 min step. To this 20µl of the 5' Primer/Taq mix was added, the reaction overlaid with 70µl of mineral oil (Sigma), and PCR was performed according to optimised conditions.

**2.4.8. RNA Extraction.**

RNA Extraction from Cells and VLPs using RNAzol.

Cellular RNA was extracted following removal of the culture supernatant and washing the cells once with 1XPBS buffer. Cells were then resuspended, or lysed if adherent, by the addition of 0.2ml of RNAzol (Biogenesis) per 10<sup>6</sup> cells. For RNA extraction from VLPs, serum free culture supernatant was removed from the cells and concentrated using a Centriprep C-100 concentrator (Amicon). A minimum of three times the

volume of RNazol was added to the concentrated VLPs, and RNA extracted.

Cellular and VLP RNA samples were shaken vigorously after the addition RNazol and 0.1x the volume of chloroform. After a 5 min incubation on ice the samples were spun at 13,000 rpm, 4°C, for 15 min in a microfuge. The upper RNA containing layer was carefully removed avoiding the proteins and DNA at the interphase. RNA was precipitated by the addition of an equal volume of propan-2-ol followed by incubation on ice for 15 min. This was then centrifuged at 13,000g, 4°C for an additional 15 min. The resulting pellet was washed with 70% ethanol, dried and resuspended in DEPC treated water. The RNA was additionally purified by a second precipitation step. Sodium chloride was added to 0.2M then twice the volume of ethanol was added, and the tubes incubated at -20°C for 60 min and spun and washed as above.

#### Extraction and Preparation of Virion RNA.

<u>20% Sucrose in TNE</u>	<u>Lysis Mix</u>
20%(w/v) Sucrose (BDH)	1% SDS
10 mM Tris pH 8.0	50 mM Tris pH 7.5
100 mM NaCl	100 mM NaCl
1 mM EDTA	

To harvest virion RNA supernatant was removed from infected cells typically after 16 hours. Cell debris was removed by spinning at 10,000 rpm for 10min in a Sorvall SS34 rotor, then filtered through a 0.45µm Acrodisk. 3ml of a 20% Sucrose in TNE solution was carefully layered underneath the culture medium. The tubes were then spun in a Sorvall AH629 rotor at 26,000 rpm for 3 hours at 4°C, to pellet the viral particles through the sucrose. Medium was removed by pipette and the sucrose

cushion was carefully poured off to waste. Viral pellets were resuspended and lysed in 300µl of Lysis Mix solution then transferred into fresh eppendorfs. ProteinaseK (Boehringer) was added to 200ng/µl along with Yeast tRNA to 12µg/µl. After incubation at 37°C for 30min, viral RNA was recovered following one phenol and three phenol/chloroform extractions (RNA grade phenol, pH 4.8, Sigma). The RNA was stored as an ethanol precipitate at -70°C.

#### **2.4.9. Northern Analysis of viral RNA.**

<u>Transfer Buffer</u>	<u>20X SSC</u>
25 mM Na <sub>2</sub> HPO <sub>4</sub> pH 6.5	3 M NaCl
25 mM NaH <sub>2</sub> PO <sub>4</sub>	0.3 M sodium citrate
<u>Prehybridisation Buffer</u>	<u>TP5</u>
15ml Formamide (BDH)	1.0% Poly Vinyl Pyrrolidone (MW. 30,000)
6ml 50% Dextran sulphate	1.0% Bovine Serum Albumin (BDH)
6ml TP5	1.0% Ficoll (Mol wt. 400,000)
3ml Water	0.5% Sodium pyrophosphate
1.74g NaCl.	5.0% SDS
	250mM Tris pH 7.5
	Filter through 0.45µm Acrodisk.

Virion RNA was resuspended from ethanol in TE buffer pH 7.5, and was electrophoresed through 0.8% agarose (Pharmacia) 0.5X TBE gels in 0.5X TBE running buffer at 4°C for 5-6 hours at 5-6W. Electrotransfer of the RNA onto Hybond N (Amersham) was achieved using a Hoefer electroblotter running at 200mA overnight at 4°C, through 25mM sodium phosphate buffer (pH 6.5). The membrane was recovered, and RNA was crosslinked to it using the Stratagene UV crosslinker autocrosslink

function. The membrane was washed for 3 x 15min in 25mM sodium phosphate buffer (pH 6.5), and then for 10 min in 1% SDS. Prehybridisation was performed using 30 ml of prehybridisation buffer for at least 4 hours at 42°C. A freshly prepared random primed DNA probe was denatured at 100°C for 3 min and then added to the membrane in a volume of prehybridisation buffer to give 1x10<sup>6</sup> counts per ml. Hybridisation was performed by incubating at 42°C overnight in a rotating hybridiser (Hybaid). Unbound probe was washed away using a 2XSSC/1% SDS wash for 5 min, then 30 min at 42°C; 30 min at 65°C, and finally at higher stringency, 0.2XSSC/1% SDS at 65°C for 15 min. The filter was allowed to air dry, covered in Saran wrap, and the RNA visualised by autoradiography.

#### Random Primed DNA Probe.

Random primed probes were prepared using the "Rediprime" system, from Amersham according to the kit instructions. To remove unincorporated counts and purify the probe, the reaction mixture was fractionated down a Sephadex G100 column.

A pasteur pipette was drawn out to form a capillary end, and was filled with the Sephadex G100/TE slurry. The resin was washed several times with TE. The crude radiolabelled mixture was applied to the column, along with 150µl of TE buffer. Fractions were collected into numbered eppendorf tubes. Additional 150µl TE aliquots were added, and the activity of the eluted fractions was monitored using a hand-held Geiger counter. The first high activity fractions were combined to form the probe. These consist of just the longer random primed fragments, unable to penetrate into the pores of the G100 resin, which hybridise better to the target RNA. Unincorporated label was not present in the probes.

The probe was mixed with formamide and TP5 solution (to a final concentration of 50 and 20 % respectively), then denatured for 3 min at 100°C and rapidly placed on ice.

#### Optimisation of Northern Analysis.

The detection limit of the Northern procedure was assessed and optimised using an *in vitro* transcribed RNA 3kB HIV-1<sub>NL4-3</sub> RNA. This was produced from Xho cut pSH91.136, transcribed by T7 RNA polymerase. The RNA corresponded to residues 5942 to 8887 of the HIV-1<sub>NL4-3</sub> env gene with an additional 55 nucleotides from pBLUESCRIPT KS+. The RNA concentration was determined by UV absorbance at 260 nm and adjusted by dilution to the amounts shown in figure 2.4.

After electrophoresis and electrotransfer onto a nylon Hybond N membrane (Amersham), the blot was probed overnight using a random primed DNA probe purified down a Sephadex G100 column. An example of two probe profiles is shown in figure 2.3. Fractions from Sephadex G100 columns were counted in a scintillation counter.

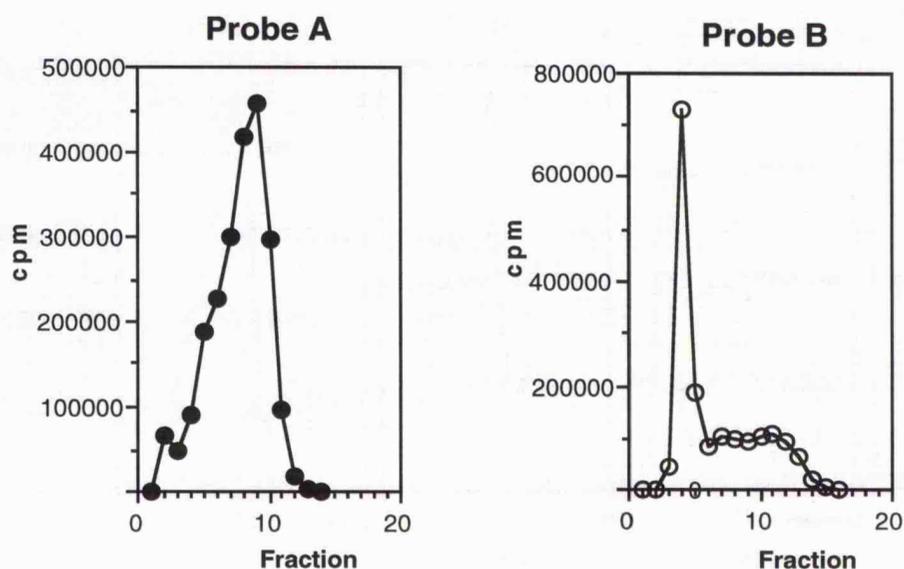


Figure 2.3. Random primed DNA probe profiles following Sephadex G-100 size exclusion chromatography.

Probe A consists of a broad size range of fragments, with the majority of the label in higher fractions corresponding to lower molecular weight probes and unincorporated label. Probe B was ideal, there was a narrow peak, fractions 4, 5 and 6, which corresponds to longer random primed products (average length 400 nt) of high activity. Probe B was a DNA fragment from HIV-1<sub>NL43</sub> from sequences 1443 to 8887 (SphI to XhoI), and was used for all hybridisations to detect HIV-1 RNA. To detect RSV RNA the plasmid pLGPE (containing a complete RSV provirus) was used.

After washing away unbound probe and autoradiography, figure 2.4 shows that the detection limit for an *in vitro* transcribed RNA was around 10 to 25 pg.

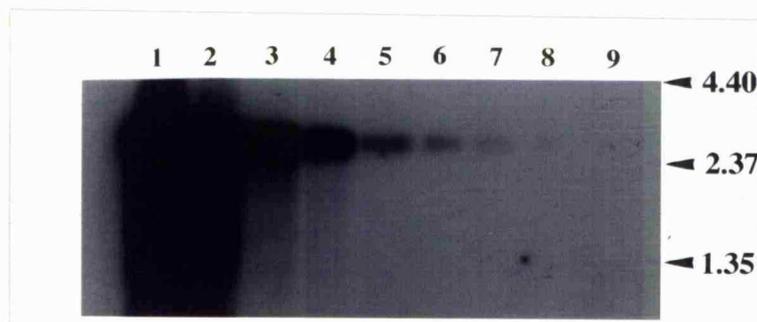


Figure 2.4. Northern Analysis of *in vitro* transcribed RNA. Lanes 1 (100ng), 2 (20 ng), 3 (2 ng), 4 (500 pg), 5 (100 pg), 6 (50 pg), 7 (25 pg), 8 (10 pg) and 9 (1 pg) RNA.

## **2.5 VLP/Virus Techniques.**

### **2.5.1. Electroporation of Cells.**

Circularly permuted HIV-1 plasmids were introduced into cells by electroporation as described by Cann et al. (Cann et al., 1988). Twenty four hours prior to electroporation fresh medium containing serum was added to the cells, and their density adjusted if required to  $0.5 \times 10^6$  per ml. Immediately before electroporation the culture supernatant was removed,

and the cells washed with 1XPBS. Cells were released from the tissue culture flask by treatment with trypsin if adherent, resuspended in fresh medium and counted. The cells were adjusted so that 0.25ml contained  $2 \times 10^6$  cells. The 0.25ml of cells were placed into an electroporation cuvette, along with 25 $\mu$ g of plasmid DNA ensuring no trapped air bubbles. The cuvette was left on ice for 10min and the cells electroporated using a BIO-RAD Gene pulser apparatus, set at 960 $\mu$ F/150mV. Time constants were recorded. After an additional 10 min on ice the cells were removed from the cuvette and placed into the tissue culture flask, along with fresh medium.

#### **2.5.2. Infection of C8166 cells with HIV-1 Viruses.**

Titred virus stocks were recovered from storage in liquid nitrogen, and allowed to thaw slowly to room temperature. Infections were initiated with 5.0 ng of p24 added to 10ml of C8166 cells at  $5 \times 10^5$  per ml. Virus was allowed to infect by incubation at 37°C overnight. The following day the cells were placed into fresh medium to  $5 \times 10^5$  cells per ml. Cultures were inspected daily and samples of culture supernatant were removed for p24 assay, and viral stocks. Infections were maintained by the addition of uninfected cells and adding or removing cells and medium to keep the cell density between 0.5- $1 \times 10^6$  cells per ml.

#### **2.5.3. Antigen capture p24 ELISA.**

This method is a double antibody "sandwich" ELISA for the quantitation of viral capsid antigen p24, against added external standards (Conley et al., 1993).

<u>TBS</u>	<u>Diluent</u>	<u>TMT/SS</u>
144 mM NaCl	TBS	2.0% Skimmed Milk Powder
25 mM Tris pH 7.5	0.1% Empigen	20% Sheep Serum
	10% Sheep Serum	0.5% Tween 20
		in TBS

#### Plate Preparation.

The capture antibody (D7320) was diluted from the stock (in water at 1mg/ml) to a final concentration of 5-10 $\mu$ g/ml in 100 mM NaHCO<sub>3</sub>, pH 8.5. 100 $\mu$ l aliquots of this solution was used to coat the surface of the wells in the microtitre plate, and incubated overnight at room temperature. The plate was washed three times with 200 $\mu$ l per well of TBS. The wells were then blocked by the addition of 200 $\mu$ l of "Blotto" which was 2% (w/v) Skimmed milk powder (Marvel) in TBS, and incubated for a further 30 min. The wells were then washed three times with 200 $\mu$ l of TBS.

#### Sample Preparation.

Samples containing HIV-1 virus or VLPs were inactivated by the addition of Empigen (Calbiochem) detergent to a final concentration of 1%, and incubation at 56°C for 30 min. These samples were then diluted accordingly from the expected p24 titre, using the TBS/Empigen/sheep serum diluent. p24 antigen standards (MRC ADP Reagent Repository) were prepared using the stock solution at 10 ng/ml. 100 $\mu$ l of diluent was dispensed into 5 eppendorf tubes, labelled S1 to S5. Next 46.25 $\mu$ l of p24 stock was added into tube S1, mixed to give a final p24 concentration of 3.1 ng/ml. This same volume was removed from S1 and added to S2, producing a 1 ng/ml p24 solution. The dilution procedure was repeated to give p24 standards S3, S4 and S5, each with a p24 concentration of 0.31, 0.1 and 0.03 ng/ml of p24 protein.

Duplicate 100 $\mu$ l standards were placed into the wells of the microtitre plate. Two diluent samples were prepared to act as negative (blank)

controls. 100µl of the diluted samples containing HIV-1 virus or VLPs were also added into the wells of the microtitre plate. Antigen capture was then allowed to occur for 3 hrs at 30°C.

Detection.

The microtitre wells were washed three times with 200µl of TBS. The detection antibody (EH12E1-AP) was used at about 0.5µg/ml in TMT/SS, and 100µl of this solution was added into each well. This was followed by a 1 hour incubation at 30°C. The plate was then washed five times with 200µl of 1X AMPAK (Novo Biolabs, Cambridge) wash buffer (from 20x concentrate supplied). 50µl of AMPAK substrate solution (reconstituted according to the manufacturers instructions and used at room temperature) was added per well, and incubation allowed for 1 hour at 30°C. Next 50µl of AMPAK amplifier solution (at room temperature) was added, and colour allowed to develop for 10 to 15 min. The reaction was stopped by the addition of 25µl of AMPAK stop solution, and the absorbance of each well was measured at A<sub>495nm</sub> within 1 hour using a Dynatech Microplate Reader.

p24 titres were determined by plotting the absorbance of the known standards S1 to S5 against their p24 concentration (ng/ml), and reading off the p24 titre for the virus/VLP samples from their absorbances.

**2.5.4. HIV-1 Reverse Transcriptase Assay.**

<u>PEG Solution</u>	<u>Solubilising Solution</u>	<u>Supermix/per reaction</u>	
30% (w/v) PEG 6,000	50mM Tris-Cl pH8.2	1M Tris-Cl pH 8.2	5µl
0.4M NaCl.	10mM PMSF	0.2M MgCl <sub>2</sub>	5µl
	0.5% Triton X-100	0.01M DTT	20µl
	5% Glycerol	3H-TTP*	10µl
	80mM NaCl	polyrA + poly dT**	10µl
	2mM DTT	Water	50µl

\*<sup>3</sup>H-TTP (dTTP), 1 mCi/ml, ≈40-70 Ci/mmol (Amersham)

\*\* Template-primer (poly rA + poly dT; 5 ODU/ml; Pharmacia)

Culture supernatant was removed from the cells and passed through a 0.45µm Acrodisk filter. 0.5ml of PEG solution was added to each sample of 1 ml of supernatant, mixed gently, then incubated overnight at 4°C. A positive (infected culture supernatant) and negative control (uninfected cell supernatant) were also included. After centrifugation at 2,250g, 4°C for 30 min, all of the supernatant was carefully removed. 25µl of solubilising reagent was added per tube and well mixed. Tubes were decontaminated using 70% ethanol and removed from the CIII containment facility. 100µl of the reaction supermix was added per tube, mixed, and incubated at 37°C for 60 min. The reactions were then placed on ice, and 5µl of 0.5M EDTA was added. Numbered DEAE (DE81) filters were prepared and the samples were carefully spotted onto the filters and allowed to air dry. The filters were washed for approx 5 min in a 5% Na<sub>2</sub>HPO<sub>4</sub> solution (w/v) with gentle agitation. The wash was repeated six times. The filters were dried under a heated lamp, transferred to scintillation vials, covered with 5 ml of scintillation fluid and tritium counts were measured using a scintillation counter.

#### **2.5.5. VLP Concentration using Centriprep concentrators.**

Centriprep concentrators (Centricon, C-100) were used to concentrate culture supernatant containing VLPs after filtration through a 0.45µm Acrodisk. The unit was filled with approx 15ml of supernatant, and spun at 2,500g until liquid levels had equilibrated across the one way (100,000 MW cut off) selective membrane. The filtrate was discarded and the centrifugation repeated until the residual volume was 15-20 times less than the starting volume.

#### **2.5.6. Sucrose Gradient Fractionation of VLPs.**

Sucrose gradients were prepared by carefully placing one volume of 60% sucrose in PBS into the centrifuge tube. Additional volumes of 50, 40, 30 and 20% sucrose solutions were then added as carefully as possible. The concentrated VLP fraction (in PBS) was layered onto the gradient, and in parallel the same volume of PBS was layered onto a second gradient. After centrifugation at 166,000g for 3hrs at 4°C, the gradients were recovered and fractionated into 150ul aliquots by hand using a P200 Gilson. The fractions from the VLP gradient were diluted and quantitated for p24 antigen. Exactly 100ul volumes were removed using a P200 Gilson from the empty gradient and weighed to calculate the sucrose density.

#### **2.5.7. Nuclease Digestion of VLP associated Nucleic Acid.**

To identify the nature of the nucleic acid associated with the VLPs, DNase and RNase digestions were performed as described below.

DNase: 1 unit of RQ1 RNase free DNase (Promega at 1U/ul) was added to the sample in a 40mM Tris-Cl pH 7.9, 6mM MgCl<sub>2</sub>, 10 mM NaCl, and 10mM CaCl<sub>2</sub> buffer. The reaction was incubated at 37°C for 15 min, then phenol chloroform extracted.

RNase: 0.5 units of DNase free RNase ONE (Promega at 10U/ul) was added to the sample in a 10mM Tris-Cl pH 7.5, 5mM EDTA, 200 mM sodium acetate buffer. The reaction was incubated at 37°C for 15 min, then phenol chloroform extracted.

For RNase and DNase digested samples both enzymes were added into the DNase buffer, this was found to successfully degrade RNA and DNA.

## **2.6. Vector Construction**

### **2.6.1. Construction of Plasmids for use in RNA Transcription.**

RNAs were made from plasmids constructed by PCR or RT-PCR as shown in the table.

<u>RNA</u>	<u>Oligonucleotides</u>	<u>Method</u>	<u>Plasmid</u>
RNA 1 (732-858)	WS 109, SH175	PCR	pSH 91.200
RNA 2 (708-858)	MH6, SH175	PCR	pMH 93.017
RNA 3 (708-858)	MH10, MH6, SH175	RT-PCR	pMH 94.005
RNA 4 (649-813)	MH6, SH241	PCR	pMH 94.004
RNA 5 (649-813)	MH17, MH6, SH241	RT-PCR	pMH 95.001
RNA 6 (649-858)	MH6, SH175	PCR	pMH 93.009
RNA 7 (649-858)	MH17, MH6, SH175	RT-PCR	pMH 95.003
RNA 8 (649-825)	MH6, SH241	PCR	pMH 93.006
RNA 10 (732-825)	WS109, SH241	PCR	pSH 91.219
RNA 11 (719-858)	MH8, SH175	PCR	pMH 93.015
RNA 12 (732-858)	AL2, SH175	PCR	pAL 93.007
RNA 13 (708-825)	MH7, SH241	PCR	pMH 93.011

Table 2. Construction of HIV-1 RNAs used for *in vitro* dimerisation analysis.

Sequences from HIV-1<sub>NL4-3</sub> (Adachi et al., 1986) were amplified by PCR with primers that contained restriction sites to allow cloning, and RNA polymerase promoters 5' of the HIV-1 sequences. Unless otherwise indicated all 5' primers contained EcoRI sites, all 3' primers contained a HindIII site. The amplified products were cloned into EcoRI-HindIII cut pUC119. Cloned sequences were verified by DNA sequencing using Sequenase. Transcription start sites always utilised authentic HIV-1 sequences.

#### **2.6.2. The VLP Particles Construct, pBCCX CSF X4.**

pBC12/CMV/IL-2 (Cullen, 1986) was digested with HindIII and SmaI, to remove the IL-2 gene. The 3327 bp vector fragment was recovered by gel purification on a 1% agarose TAE gel using GeneClean. A HindIII-ClaI-XhoI-SmaI linker (5' AGCTTATCGATTTTCTCGAGCCC 3') was ligated on to form the expression construct pBCCX. The vector was linearised by digestion with HindIII, the single strand overhang was removed by treatment with Klenow DNA polymerase generating blunt ends. Digestion with XhoI followed by dephosphorylation with CIAPase, completed the preparation of the vector, able to receive the HIV-1JR-CSF insert.

HIV-1JRCSF DNA (Koyanagi et al., 1987) was cut with SacI at position 488, and the overhang was removed by digestion with mung bean nuclease. This fragment was restricted with XhoI to create the HIV-1 insert containing sequences from 488 to 8914. This was ligated into pBCCX, and transformed into frozen competent DH5 $\alpha$  *E.coli*.

The final construct pBCCX CSF X4 (figure 2.5), was identified by restriction digestion of miniprep DNAs, and checked by sequencing CsCl prepared plasmid DNA across the Xho and (HindIII/SacI) blunt ends used for cloning, with primers PCR 7 and PCR 31 respectively. Additional restriction digestion confirmed the construct was as expected. (pBCCX CSF was constructed by J.Perry, and S. Chavda. It was completed by M.H. producing pBCCX CSF X4).

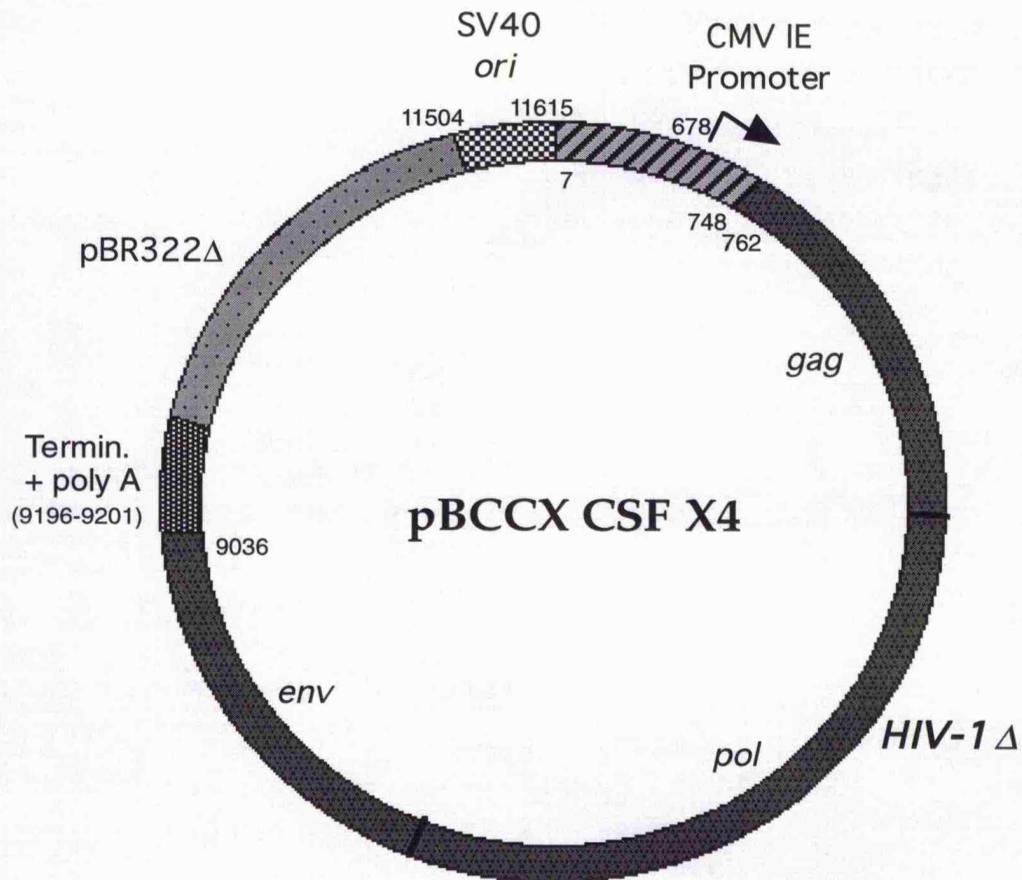


Figure 2.5. The VLP expression construct, pBCCX CSF X4.

The HIV-1<sub>JRCSF</sub> genome (638 to 8914) is expressed from the very strong CMV immediate early promoter (7-748). The construct produces RNA where +1 is at position 678, so that there is a vector encoded 84 nucleotide leader before the HIV-1 sequence begins at position 762 (in the PBS). The vector has an SV40 origin of replication (11 504 to 11 615) which allows for replication up to high copy number in cells transformed with the SV40 large T antigen (eg. COS cells). The polyadenylation signal of SV40 at position 9196AATAAA9201, termination signals, and the ampicillin resistance gene and bacterial origin of replication are present from pBR 322. This vector and can be manipulated and amplified in *E. coli*.

The HIV-1 fragment from 762 to 9036 includes all of the virus genes except the 3' terminal portion of the non-essential *nef* gene. It does not

contain any sequences from the viral LTRs. The PBS is incomplete, sequences 637UGG639 are missing from the 5' end. The construct pBCCX CSF X4 can therefore only produce non-infectious particles.

### **2.6.3. The HIV-1NL4-3 Vector, pUCΔNL.**

pUCΔNL is a circularly permuted clone (Cann, 1990) of HIV-1<sub>NL4-3</sub> interrupted by pUC18, figure 2.6.

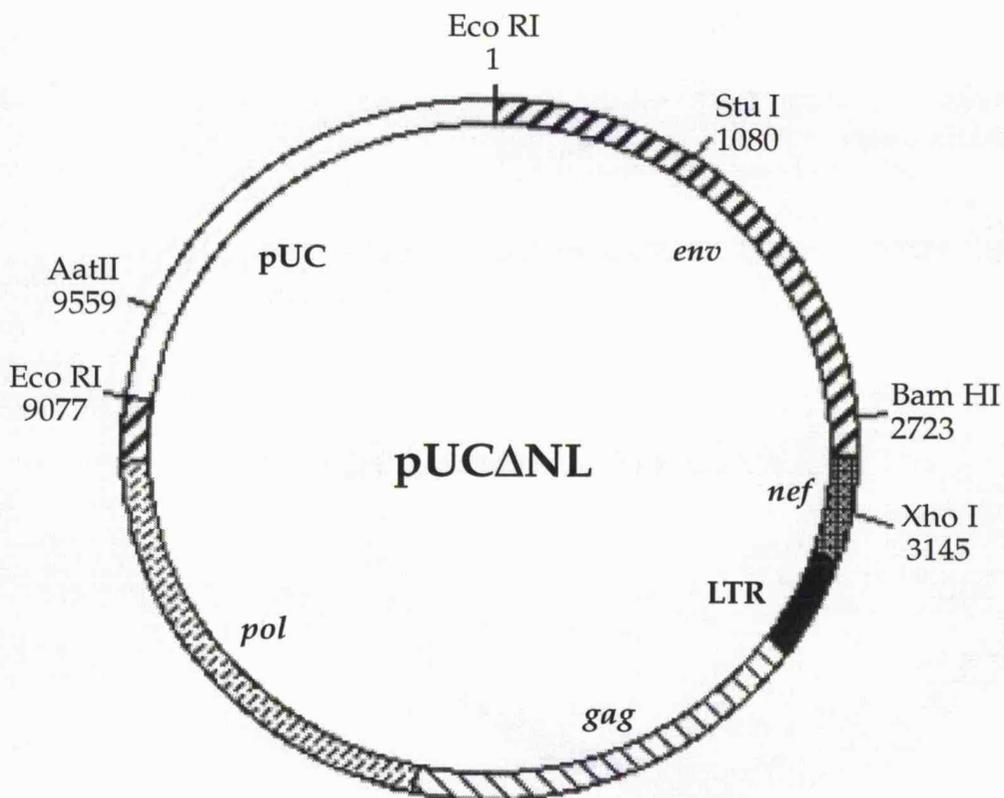


Figure 2.6. pUCΔNL plasmid. The HIV-1 genome is circularised and interrupted by pUC 18 DNA at the EcoRI sites.

The plasmid was constructed by a *SacI* partial digest of HIV-1<sub>NL4-3</sub> DNA cutting the sites within the LTR regions. This was then ligated to form a circle and digested with *EcoRI*. *EcoRI* cut pUC18 was prepared and the HIV-1<sub>NL4-3</sub> insert was cloned into the pUC18 vector to generate pUCΔNL

(M. Johnson, unpublished results). Restriction digestion with EcoRI followed by ligation at high DNA concentration generates concatemers of proviral DNA some of which are able to integrate and establish a productive infection, when introduced into cells by electroporation (Cann, 1990). pUC $\Delta$ NL plasmid DNA was used as the wild type vector for mutagenesis as it was previously demonstrated to produce infectious, HIV-1<sub>NL4-3</sub> viruses following electroporation into T cells (M. Johnson, A. J. Cann, unpublished results).

## Chapter 3: Mechanisms of Dimerisation for HIV-1 RNAs *in vitro*.

The dimerisation ability of RNAs corresponding to the DIS and DLS regions of HIV-1<sub>NL4-3</sub> was analysed *in vitro* with and without a nucleocapsid derived peptide,  $\Delta$ NCp7.

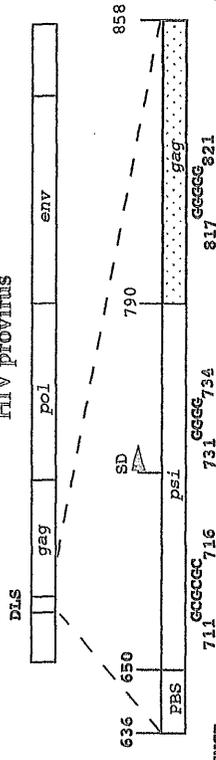
### 3.1. Introduction.

Based on the *in vitro* studies described in the introduction, three mechanisms for HIV-1 RNA dimerisation have been proposed; the Guanine Tetrad, the A-U Base Pairing, and the Kissing Loop models. The *in vitro* dimerisation of HIV-1 RNAs according to the Guanine Tetrad and Kissing Loop mechanisms were investigated here. Initially RNA dimerisation in the absence of any viral proteins was analysed, this was followed by dimerisations including the  $\Delta$ NCp7 peptide.

The RNAs under study are shown schematically in figure 3.1. The sequences (of the proviral DNA) they correspond to are shown, along with residues important for the mechanisms of dimerisation. The first nucleotide of the RNA sequence (+1) for HIV-1<sub>NL4-3</sub> is at position 455 ie. the first residue of the TAR structure. The ability of the RNAs to dimerise is indicated by a + (dimerisation competent) or a - (failed to dimerise).

The nucleotides essential for the dimerisation of the tetrad RNAs were investigated first. This analysis led to the identification of other sequences important for kissing loop mediated RNA dimerisation.

HIV provirus



RNA SEQUENCE	711	716	731	734	817	821	DIMER
1 732-858	GGG	GGGG	GGGG	GGGG	GGGG	GGGG	+
2 708-858	GGGG	GGGG	GGGG	GGGG	GGGG	GGGG	+
3 708-858	GGGG	GGGG	GGGG	GGGG	GGGG	GGGG	-
10 732-825	GGG	GGGG	GGGG	GGGG	GGGG	GGGG	+
11 719-858	GGGG	GGGG	GGGG	GGGG	GGGG	GGGG	+
12 732-858	GAA	GGGG	GGGG	GGGG	GGGG	GGGG	+
13 708-825	GGGG	GGGG	GGGG	GGGG	GGGG	GGGG	+
4 649-813	GGCGC	GGGG	GGGG	GGGG	GGGG	GGGG	+
5 649-813	AAACG	GGGG	GGGG	GGGG	GGGG	GGGG	-
6 649-858	GGCGC	GGGG	GGGG	GGGG	GGGG	GGGG	+
7 649-858	AAACG	GGGG	GGGG	GGGG	GGGG	GGGG	-
8 649-825	GGCGC	GGGG	GGGG	GGGG	GGGG	GGGG	+

Figure 3.1. Representation of the RNAs under study. The RNA number, sequence and ability to dimerise is indicated. Two classes of RNAs were identified, those which dimerised by the formation of a guanine tetrad, and those which dimerised by the kissing loop structure. Nucleotides involved in RNA dimerisation are shown.

### **3.2. RNA Dimerisation by the Formation of a Guanine Tetrad Structure.**

Several lines of evidence supporting the involvement of guanine quartets in the dimerisation of HIV-1 RNA *in vitro* have been obtained. The stability of several HIV-1 RNA dimers of different length varies dramatically with the ionic radius of solution cations (Marquet et al 1991; Sundquist and Heaphy 1993; Weiss et al 1993). A conserved PuGGAPuA sequence was noticed and proposed as a candidate purine tract involved in tetrad formation (Marquet et al. 1991). The rate of formation of G-quartet structures is also cation dependent (Sundquist, 1991; Guschlbauer et al., 1990; Sen and Gilbert, 1991). Once formed RNA dimers of the HIV-1 sequences 732-858 (RNA 1) and 743-855 (DLS 112), exhibit potassium-dependent thermal stability, providing further evidence for intermolecular G-quartet formation (Sundquist and Heaphy, 1993; Awang and Sen, 1993). Moreover, the guanine-rich sequence  $817\text{GGGGGAGAA}_{825}$  was shown by deletion analysis to be required for dimerisation of HIV-1 RNA fragments that began at position 732 (Sundquist and Heaphy, 1993). Awang and Sen performed chemical probing experiments on the dimer formed by the sequence DLS 112 743-855. This demonstrated that nucleotides  $817\text{GGGGG}_{821}$  and  $836\text{GGG}_{838}$  were protected from DMS modification upon dimerisation, implicating these nucleotides in G-quartet formation (✓ in figure 3.2, and Awang and Sen, 1993). However, HIV-1 RNA fragments longer at the 5' end but terminating before the  $836\text{GGG}_{838}$  sequence (eg. RNA 10 732-825, figure 3.2) also formed dimers (Sundquist and Heaphy, 1993), suggesting the possibility that more than one dimeric structure may be formed *in vitro*.

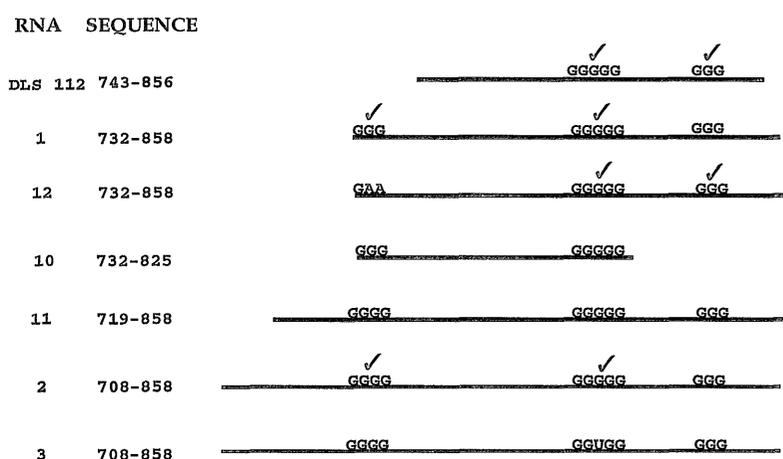


Figure 3.2. RNAs that dimerise by guanine tetrad formation, the guanine tracts used are indicated by a (✓).

### **3.3. Identification of the Guanine Tracts Involved in RNA 1 Dimerisation.**

The guanine tetrad model for HIV-1 RNA dimerisation was pursued here, following on from the work of Sundquist and Heaphy (Sundquist and Heaphy, 1993). RNA 10 (not shown) and RNA 1 (figure 3.2 and figure 3.3, lane 1), dimerised for which a  $_{817}\text{GGGGGAGAA}_{826}$  sequence was essential.

To determine if the  $_{817}\text{GGGGGAGAA}_{826}$  sequence was involved in a guanine tetrad structure for RNA 1, DMS modification protection was performed (Dr. A. Lear). DMS alkylates RNA at the N7 position of guanines, which then permits strand cleavage by subsequent treatment with aniline. The N7 position of guanines in a tetrad are Hoogsteen hydrogen bonded and are not available for modification. Therefore comparison of the DMS protection profile of native and denatured RNA monomers and dimers identifies those guanine residues involved in the tetrad.

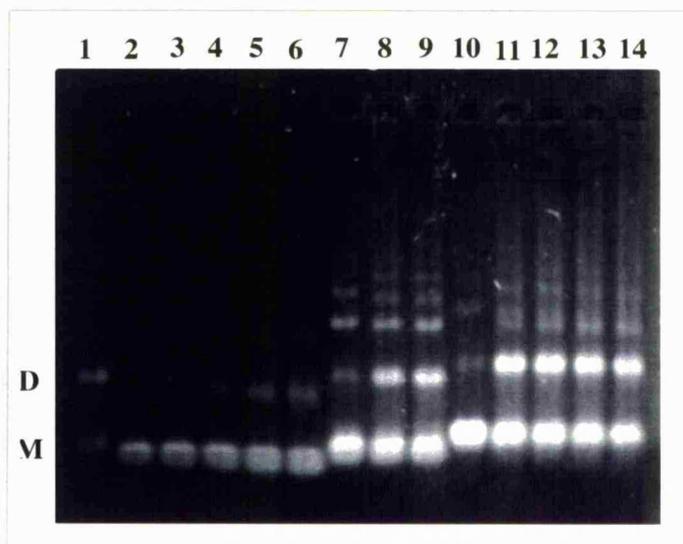


Figure 3.3. Dimerisation time course of RNAs 1, 2, 12 and 13. RNAs were incubated at  $3\mu\text{M}$  in dimerisation buffer I. Samples were removed and analysed by electrophoresis. Lane 1, RNA 1 at  $t=60$  min. Lanes 2 to 6, RNA 13 at times  $t=0, 15, 30, 60$  and  $90$  min. Lanes 7 to 9 was RNA 12 dimer formation at times  $15, 30$  and  $60$  min. Lanes 10 to 14 was RNA 2 analysed at times  $t=0, 15, 30, 60$  and  $90$  min. Samples were run on a 3% GTG agarose gel in TBE buffer.

DMS modification of RNA 1, figure 3.4, showed that guanine residues 817-819 were strongly protected, and G<sub>820</sub> was partially protected against alkylation in the native dimer (lanes 11 to 14). These guanine residues were reactive on denaturation (lanes 8 to 10), and in the RNA monomer (lanes 1 to 7). This suggested that the guanine residues 817-820 could be involved in guanine tetrad dimer formation. This was consistent with the requirement of these sequences by deletion analysis (Sundquist and Heaphy, 1993), and their protection in the DLS 112 RNA dimer (Awang and Sen, 1993). Residue G<sub>821</sub> was not protected, showing that it was not involved in the tetrad. Also the <sup>836</sup>GGG<sub>838</sub> sequence modified in the DLS 112 RNA dimer (Awang and Sen, 1993), was not protected in the RNA 1 dimer.

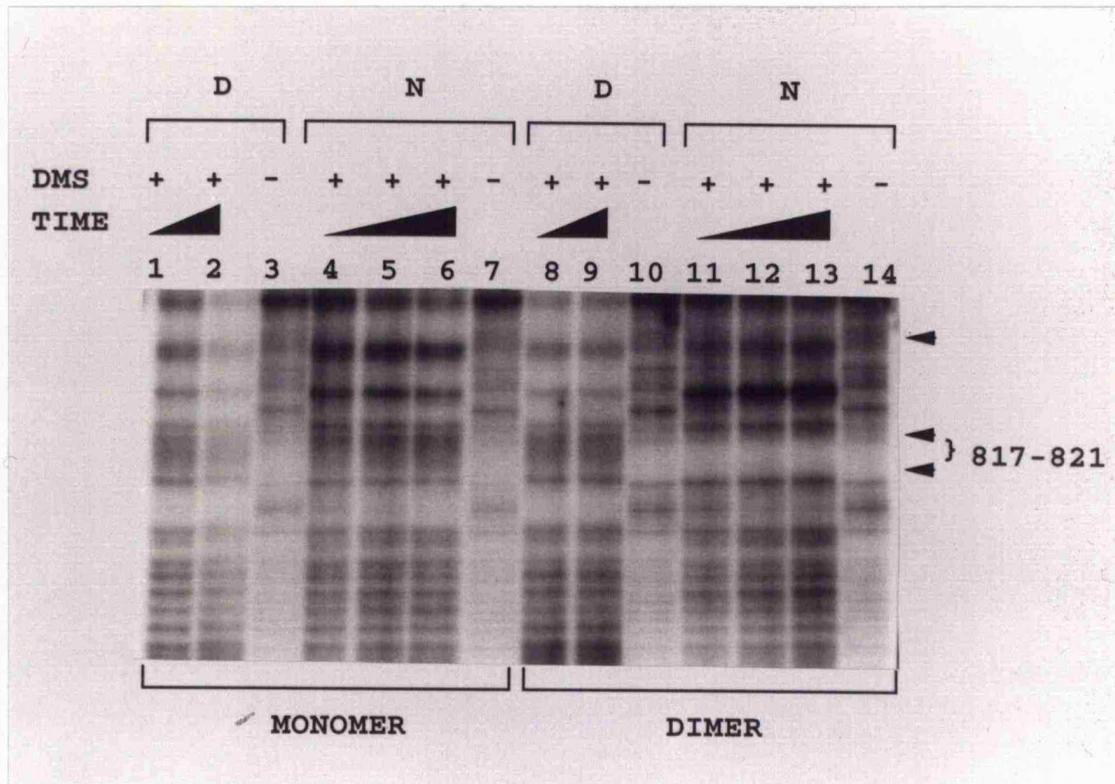


Figure 3.4. DMS footprinting of RNA 1. Monomer or dimer RNA was reacted with DMS under native or denaturing conditions as described in materials and methods. Following cleavage with aniline, samples were electrophoresed on 8% denaturing polyacrylamide gels. Lanes 1 to 7, monomer, lanes 8 to 14 dimer; lanes 1 to 3 and 8 to 10 were incubated under denaturing conditions, lanes 4 to 7 and 11 to 14 were incubated under native conditions. Lanes 3, 7, 10 and 14 were not treated with DMS. The positions of 817GGGGG821 and 836GGG838 are indicated.

The tetrad model requires two separate guanine rich regions. To locate the second guanine tract, inspection of the sequence of RNA 1 suggested the candidate regions 732GGG734, and 792GGG794. The 836GGG838 sequence identified by Awang and Sen was unlikely to be involved, because it was not modified in RNA 1, and RNA 10 (732-825) dimerised which does not contain the 836GGG838 sequence (Sundquist and Heaphy, 1993).

DMS analysis of RNA 1 (figure 3.5) revealed the guanine tract at position 732GGG734 was protected against DMS modification and could be the second guanine sequence involved in tetrad formation. Residues 732 to 734 were modified by DMS in the denatured and native monomer (lanes 1 to 7). They were protected from DMS alkylation only in the native dimer (lanes 11 to 13).

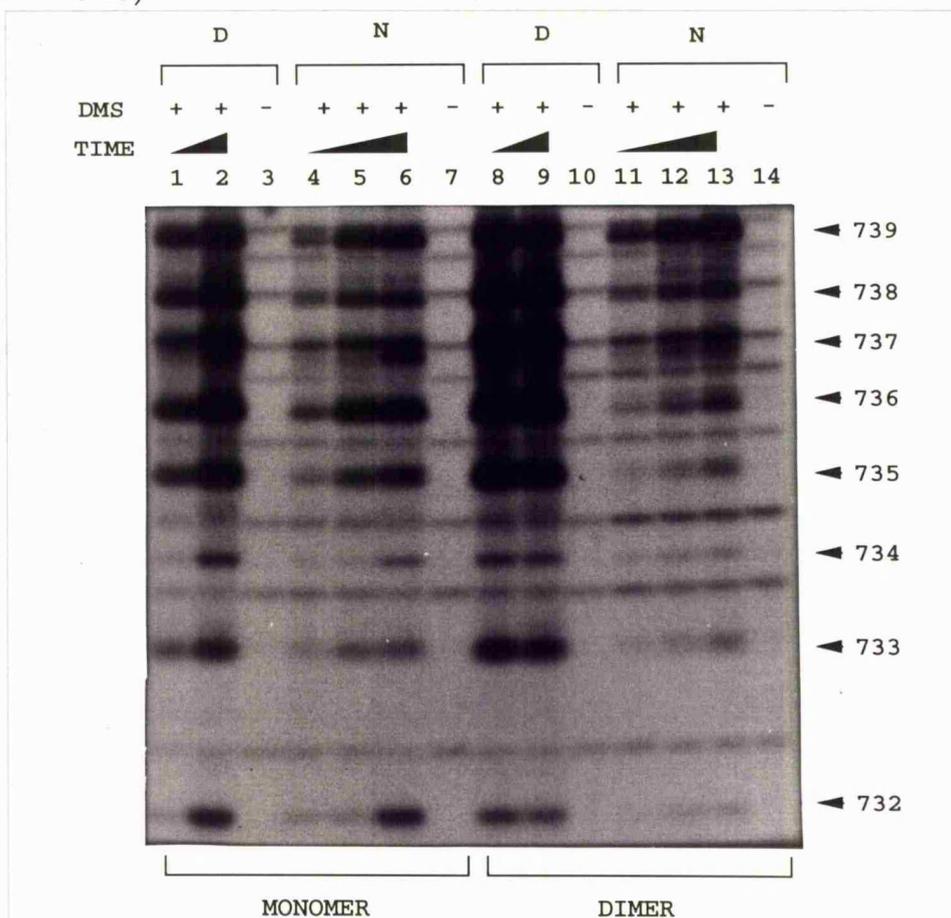


Figure 3.5. DMS footprinting of RNA 1. Monomer or dimer RNA 1 was reacted with DMS under native and denaturing conditions. Lanes 1 to 7 monomer, 8 to 14 dimer. Lanes 1 to 3 and 8 to 10 were analysed under denaturing conditions, lanes 4 to 7 and 11 to 14 native. Lanes 3, 7, 10 and 14 were not DMS treated. The position of residues 732 to 734 is indicated.

This DMS analysis was not conclusive proof for the identification of the guanine tetrad sequences responsible for the dimerisation of RNA 1. For the 732GGG734 sequence at the 5' end of the RNA, the DMS reaction profile

could be due to stacking interactions of these guanines. To more directly test the involvement of the  $732\text{GGG}_{734}$  sequence, it was mutated producing RNA 12. Also the role of this tract in the dimerisation of longer RNAs 11 and 2, where the  $732\text{GGG}_{734}$  sequence was internalised, was determined.

#### **3.4. An Alternative Guanine Tetrad Formed in the RNA 12 Dimer.**

To determine if the  $732\text{GGG}_{734}$  sequence was involved in the dimerisation of RNA 1, it was mutated producing RNA 12. The  $732\text{GGG}_{734}$  sequence was changed to  $732\text{GAA}_{734}$  (figure 3.1 and 3.2). This mutation should prevent the formation of a tetrad requiring  $732\text{GGG}_{734}$  by the introduction of the adenines. However RNA 12 formed a dimer on incubation in buffer I (figure 3.3, lanes 7 to 9). This suggested that  $733\text{GG}_{734}$  was not involved in tetrad formation, therefore another guanine rich sequence may be responsible for dimerisation of RNA 12. This was investigated by DMS modification.

DMS analysis of RNA 12 (figure 3.6) produced an alternative reaction profile compared to RNA 1. The  $836\text{GGG}_{838}$  tract was reactive in the presence of DMS for the monomer RNA 12 (lanes 1 to 3), and was unreactive in the dimer (lanes 5 to 7). The protection of  $836\text{GGG}_{838}$  indicated that a tetrad had formed in RNA 12 between sequences  $836\text{GGG}_{838}$  and  $817\text{GGGG}_{820}$ , figure 3.6.

This was a similar result to that obtained by Awang and Sen, although RNA 12 does not correspond exactly to the DLS 112 RNA, figure 3.2. For RNAs 1 and 12, the  $817\text{GGGG}_{820}$  sequence was involved in tetrad formation. However the second guanine tract was variable.  $732\text{GGG}_{734}$  was used in RNA 1, but  $836\text{GGG}_{838}$  could substitute when the  $732\text{GGG}_{734}$  sequence was mutated in RNA 12. These results indicated some flexibility for tetrad formation within a fairly short RNA.

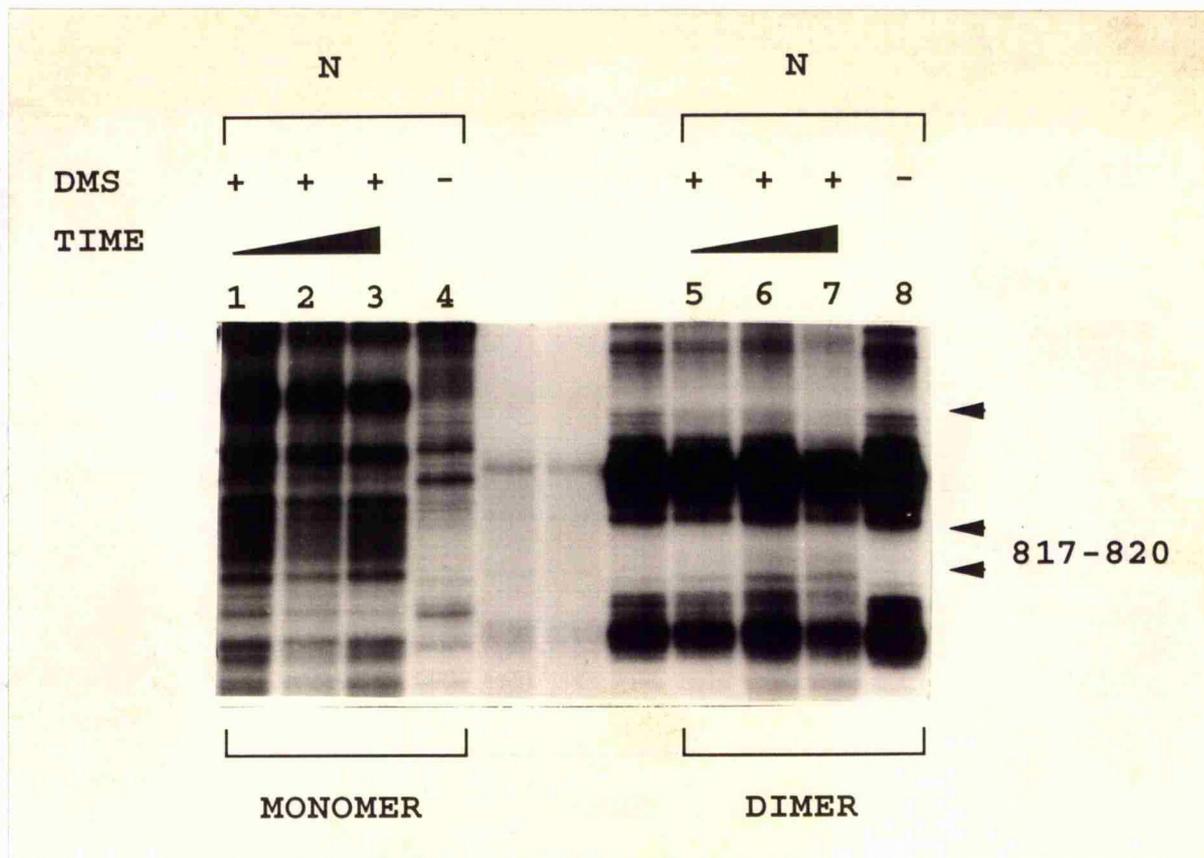


Figure 3.6. DMS footprinting of RNA 12. Monomer and Dimer RNA 12 was reacted with DMS under native conditions, as described in materials and methods. Lanes 1 to 4 monomer, lanes 5 to 8 dimer; lanes 1 to 3 and 5 to 7 were reacted with DMS, lanes 4 and 8 were not DMS treated. The positions of 817GGGG<sub>820</sub> and 836GGG<sub>838</sub> are indicated.

The involvement of the 732GGG<sub>734</sub> sequence in tetrad formation was questioned due to its location at the 5' end of the RNA. At this position it may be structurally less constrained than in an internal location. Therefore the two or three guanines at the start of a T7 RNA polymerase transcribed RNA eg. RNA 1, may be able to fold back and form a tetrad with 817GGGG<sub>820</sub>. To re-address the role of 732GGG<sub>734</sub>, it was internalised by the construction of RNAs with additional 5' sequences. This was performed to determine if the 732GGG<sub>734</sub> tract, or 836GGG<sub>838</sub> was used in tetrad dimer formation for longer RNAs.

### 3.5. RNA 11 Dimers Contained a Guanine Tetrad.

To determine whether longer RNAs would dimerise by a guanine tetrad, RNAs were constructed with additional 5' sequences compared to RNA 1. RNA 11 (719-858) began with a GGC transcription start sequence, figure 3.2, and contained an additional 5' 13 nucleotides compared to RNA 1. RNA 11 formed dimers that showed maximum resistance to heat denaturation in the presence of potassium, consistent with a tetrad in this RNA, figure 3.7.

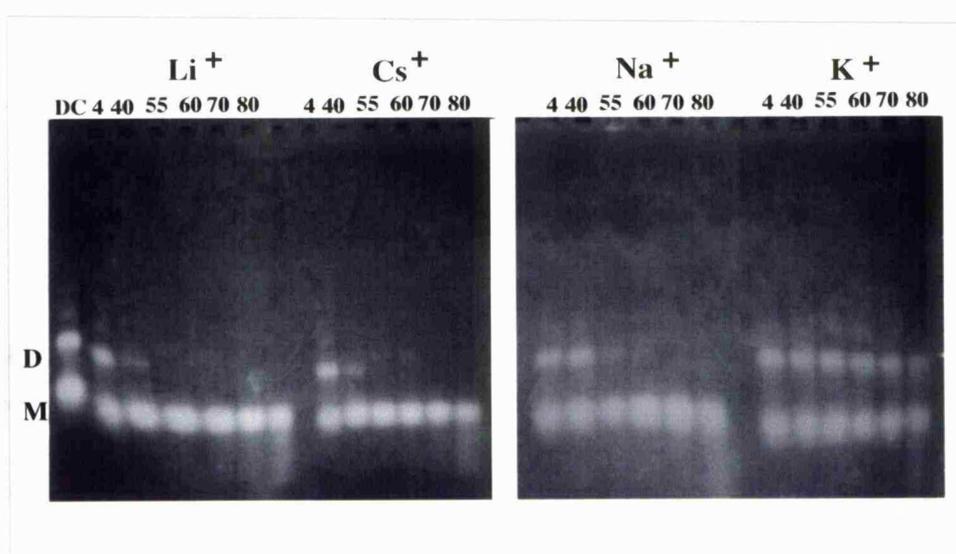


Figure 3.7. Cation Dependent Thermal Stability of RNA 11. RNA 11 dimers were incubated for 5 min at the temperatures shown in buffers containing Li<sup>+</sup>, Cs<sup>+</sup>, Na<sup>+</sup> and K<sup>+</sup>, then prepared for electrophoresis. Samples were electrophoresed on 3% NuSieve GTG gels in TBE (K<sup>+</sup>).

RNA 11 dimers had reduced melting temperatures of 40<sup>0</sup>C when Li<sup>+</sup> and Cs<sup>+</sup> were the solution cations. This was increased with Na<sup>+</sup> upto 55<sup>0</sup>C, with the maximum resistance to thermal denaturation occurring with K<sup>+</sup> at 80<sup>0</sup>C. DMS modification was not performed on this RNA as it began with two guanine residues, and it was possible that these could contribute to tetrad formation in this RNA dimer. Therefore a slightly longer RNA 2 (708-858) was constructed, figure 3.1 and 3.2.

### **3.6. RNA 2 Dimerised by Dual Guanine Tetrad Formation.**

To determine which tetrad sets of guanines were used in an RNA where the 732GGG734 was internalised and the RNA did not begin with a guanine rich sequence, RNA 2 (708-858) was studied in detail. This RNA began with the GAA start sequence and should not form a tetrad with 817GGGG820, like RNA 12. In RNA 2 the 732GGG734 sequence was part of a four guanine tract, 731GGGG734 in this longer RNA.

RNA 2 formed a dimer on incubation in dimerisation buffer I, fig 3.3, lanes 10 to 14 and figure 3.13, panel A. However it was necessary to show that the upper band which formed by RNA 2 with time was a two-stranded RNA dimer.

#### **3.6.1. RNA 2 Formed a Two-stranded RNA Dimer.**

Dimerisation has to be analysed under non-denaturing conditions. Therefore following a dimerisation reaction, the presence of a band with reduced electrophoretic mobility could be due to dimeric RNA or a monomeric RNA conformer. To prove that RNA 2 formed a two-strand RNA dimer, its gel-retardation pattern with an oligonucleotide was analysed.

The method is described schematically in figure 3.8, along with the results for RNA 2. This RNA was dimerised in the presence of increasing concentrations of an oligonucleotide, SH175. This anneals to the RNA at position 838-858 but does not inhibit dimerisation. In the dimerisation reaction without any oligonucleotide (- Oligo), the monomer (M) and dimer band (D) are shown. With increasing SH175 concentration (left to right) there was a reduction in the mobility of the monomer (shift), due to the oligonucleotide annealing to a single position on the RNA. There was also a single shift in the dimeric RNA, corresponding to one oligonucleotide bound to one strand of the RNA dimer (point A). At

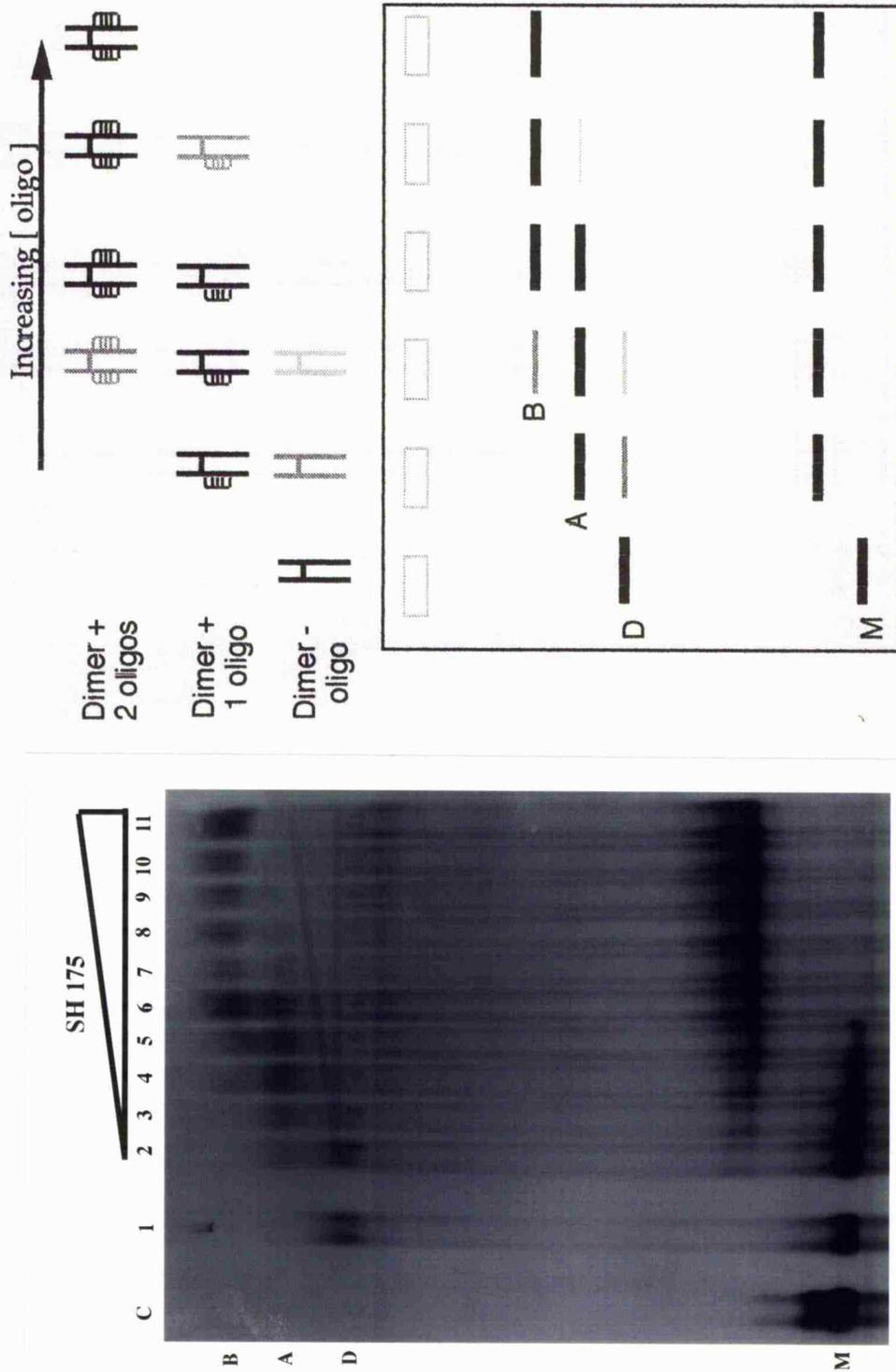


Figure 3.8. RNA 2 forms a dimer. Increasing concentrations of SH175 (left to right) was incubated with RNA 2 during dimerisation, lanes 2 to 11. Lane C was the denatured RNA, lane 1 was the RNA dimer formed in the absence of oligo. A single shift in the mobility of the dimer, point A, and a second shift at higher oligo concentration, point B, demonstrated the two strand nature of the RNA dimer. Samples were electrophoresed through a 6% polyacrylamide TBE gel.

higher SH175 concentrations, a second dimer shift occurred, consistent with a second oligonucleotide bound to the other strand of the dimer (point B). No additional shifts in the dimer occurred. This may have been expected if the upper band was a trimer or tetramer. Based on this evidence, RNA 2 formed a dimer consisting of two strands, figure 3.8.

Oligonucleotide SH175 annealed to positions 838 to 858 on the RNA and did not prevent dimerisation. This suggested that the  ${}_{836}\text{GGG}_{838}$  guanine tract identified by Awang and Sen (Awang and Sen, 1993), was not essential for the dimerisation of RNA 2. One residue ( $\text{G}_{838}$ ) was annealed to the oligonucleotide making it unavailable for tetrad formation; also the structure of the RNA may have been constrained in this region.

To determine if a tetrad was part of the RNA 2 dimer (as expected), the cation dependent thermal stability of this RNA dimer was determined.

### **3.6.2. Cation Dependent Thermal Stability of the RNA 2 Dimer.**

RNA 2 dimers were formed by incubation in dimerisation buffer I for 60min at 37°C, precipitated and resuspended in buffers containing 100mM  $\text{Li}^+$ ,  $\text{Cs}^+$ ,  $\text{Na}^+$ , and  $\text{K}^+$  ions. The thermal stability was determined by incubation of the dimers at the temperatures indicated for 5 min, followed by electrophoresis, figure 3.9. The dimers showed an enhanced thermal stability in the presence of potassium ions, with a melting temperature of about 75°C. The thermal stabilities of the dimers were reduced with the other solution cations, 55°C for  $\text{Na}^+$ , and 40°C for  $\text{Li}^+$  and  $\text{Cs}^+$ .

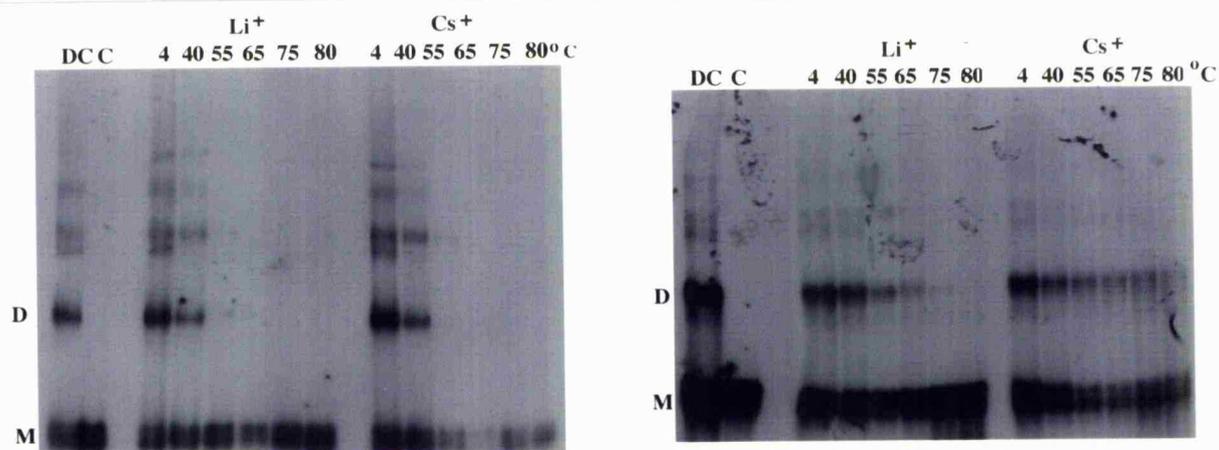


Figure 3.9. Cation dependent thermal stability of RNA 2 dimers. RNA 2 dimers were incubated at the temperatures shown for 5 min in the four cation buffers Li<sup>+</sup>, Cs<sup>+</sup>, Na<sup>+</sup> and K<sup>+</sup>, then electrophoresed on a 4% polyacrylamide TBE(K<sup>+</sup>) gel. Lane DC was the RNA 2 dimerisation control and lane C was the denatured RNA.

This thermal stability profile was consistent with a tetrad present in the RNA 2 dimer containing guanines. Also, higher ordered RNA structures were present. These may be trimers or tetramers, but demonstrate the same cation dependent thermal stability as the dimer. The tetrad model can account for higher ordered RNAs, eg. a tetramer could be formed where each RNA strand contributes one tract of guanines to the structure.

To identify which guanine tracts were involved in RNA 2 dimer formation, the guanine tracts, 731GGGG734, 817GGGG820 and 836GGG838 were analysed by DMS modification protection of RNA 2 dimers.

### **3.6.3. Identification of the Guanine Tracts Responsible for the Dimerisation of RNA 2.**

DMS probing of the RNA 2 dimers showed that the 817GGGGG821 sequence was fully reactive in the monomer under native conditions

(figure 3.10 (a), lanes 5 to 7), whereas  $_{817}\text{GGG}_{819}$  became highly protected and  $\text{G}_{820}$  became partially protected on dimerisation (figure 3.10 (a), lanes 13-15). Similarly as shown in figure 3.10 (b),  $_{731}\text{GGG}_{733}$  was specifically protected from DMS modification on dimerisation ( $\text{G}_{734}$  appears unprotected) compare lanes 5 to 7 with 13 to 15, ie. monomer and dimer under native conditions.

Since terminal guanines in a quadruplex are not always well protected (Williamson et al. 1989), it was possible that the partially protected  $\text{G}_{820}$  and the unprotected  $\text{G}_{734}$  were also involved in forming the quadruple helix, ie. a four quartet  $_{731}\text{GGGG}_{734}/_{817}\text{GGGG}_{820}$  quadruple helix was formed.  $\text{G}_{821}$  was again unprotected and presumably not involved in G-quartet formation.

Unexpectedly some protection of the  $_{836}\text{GGG}_{838}$  sequence was evident in the RNA 2 dimer under native conditions (figure 3.10 (a), compare lanes 5 to 7 with 13 to 15 ie. monomer with dimer). No protection of this sequence was evident in the RNA 1 dimer. (Figure 3.4 compare lanes 4 to 6 with 11 to 13). Partial protection of the  $_{836}\text{GGG}_{838}$  sequence may be due to the purified RNA 2 dimer consisting of a mixture of  $_{731}\text{GGGG}_{734}/_{817}\text{GGGG}_{820}$  (major product) and  $_{817}\text{GGGG}_{820}/_{836}\text{GGG}_{838}$  (minor product). Therefore both types of pairing were involved in dimer formation for RNA 2, but the  $_{731}\text{GGGG}_{734}/_{817}\text{GGGG}_{820}$  dimer predominated under these experimental conditions.



### **3.7. RNA Dimers can be Composed of Several Guanine Tetrads.**

The  $817\text{GGGG}_{820}$  guanine tract appeared to be essential for the dimerisation of all tetrad RNAs (Sundquist and Heaphy, 1993; Awang and Sen, 1993; and figure 3.2). In contrast, the second tract was variable.

RNA 1 used  $732\text{GGG}_{734}$ , although  $836\text{GGG}_{838}$  could substitute when  $732\text{GGG}_{734}$  was mutated to  $732\text{GAA}_{734}$  in RNA 12 (figure 3.6). However like RNA 1, RNA 13 (708-825) dimerised without using the  $836\text{GGG}_{838}$  sequence (figure 3.3, lanes 2 to 6) as it was absent from this RNA. Dimerisation of RNA 13 appeared to be less efficient than comparable RNAs where the  $836\text{GGG}_{838}$  tract was present, compare lanes 2 to 6 for RNA 13 with 7 to 9 for RNA 12, and 10 to 14 for RNA 2. This feature has previously been reported (Sundquist and Heaphy, 1993).

In RNA 2 containing the  $731\text{GGGG}_{734}$ ,  $817\text{GGGG}_{820}$ , and  $836\text{GGG}_{838}$  sequences, the guanine tetrad formed between  $731\text{GGGG}_{734}/817\text{GGGG}_{820}$  with a minor contribution from  $836\text{GGG}_{838}/817\text{GGGG}_{820}$ . Therefore in HIV-1 genomic RNA any of these tracts may be required for tetrad formation, but  $817\text{GGGG}_{820}$  appears to be essential.

The DLS 112 nt RNA analysed by Awang et al., lacks the  $731\text{GGGG}_{734}$  sequence, so a tetrad formed with  $836\text{GGG}_{838}$  and  $817\text{GGGG}_{820}$ , a result consistent with those observed here for RNA 12. The arrangement of the guanines in the tetrad for RNA 2 and DLS 112 could be as shown in figure 3.11.

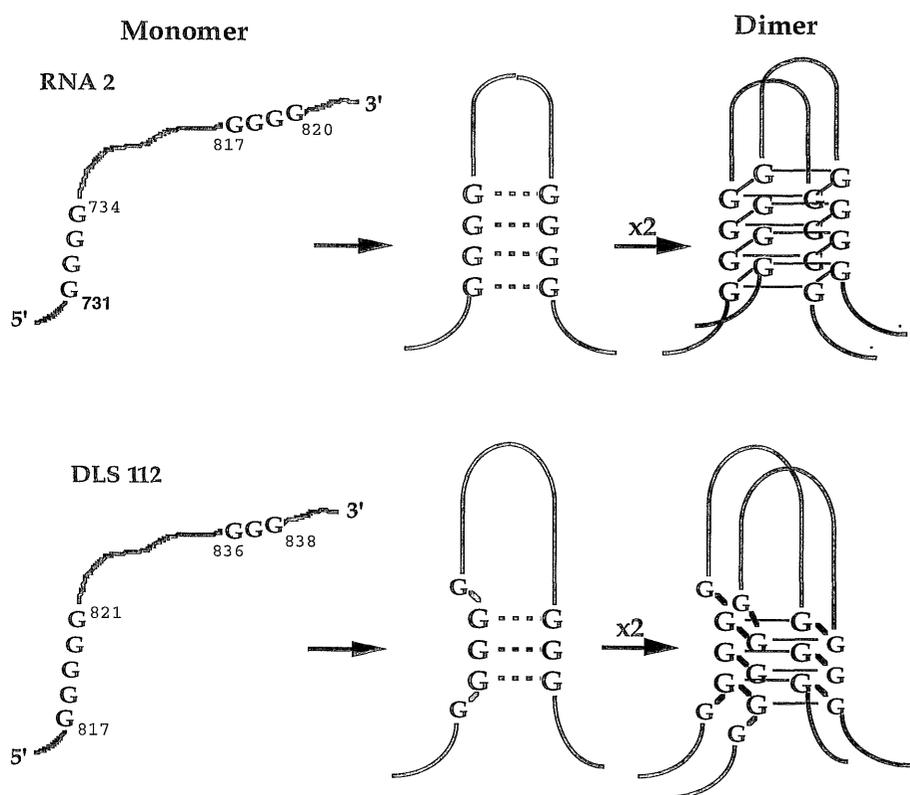


Figure 3.11. Possible guanine tetrad conformations responsible for the dimerisation of RNA 2 and DLS 112.

For the DLS 112 RNA, sequences 836GGG838 can form a tetrad with 817GGGG821 perhaps with three planes of guanine residues (Awang and Sen, 1993). However, the tetrad structure may be such that any of the 817GGGG821 guanines could be used, as they were equally protected in DLS 112 (Awang and Sen, 1993). An arrangement of guanines like this may have formed in the RNA 1 and 12 dimers, where 732GGG734 and 836GGG838 were important. However for RNA 2, four planes of guanines may be present in the tetrad involving 731GGGG734 and 817GGGG820, as each guanine 731-733 and 817-819 was unreactive to DMS on analysis. The

partial reactivity of G<sub>820</sub> and lack of modification of G<sub>734</sub> may have been due to their location on the outer edges of the four guanine tetrad (Williamson et al., 1989). However these observations do not rule out the possibility that a three plane guanine tetrad may be present in RNA 2.

The ability of the RNAs to adopt alternative tetrad conformations was sequence and context dependent. The functional redundancy of purine quartets containing guanines, or adenines and guanines, has also been observed in the dimerisation of a 615 nucleotide HIV-1<sub>LAI</sub> RNA (Paillart et al., 1994). It remains possible that mixed guanine/adenine quartets could form, as initially proposed by Marquet et al. (Marquet et al., 1991). However such structures were not detected in the RNAs analysed here. It is unknown whether cation co-ordination will occur with the selectivity shown by guanine quartets (and these dimerisation competent RNAs) in a mixed purine tetrad.

Guanine-mediated quartets are polymorphic, for example under different conditions, a telomeric DNA oligomer is able to form either a parallel stranded tetramer, or an antiparallel strand dimer (Kang et al., 1992; Smith and Feigon, 1992). The dimerisation behaviour of RNAs 1, 2, 12 and DLS 112 suggested that there was considerable flexibility in which the DLS region, or RNA fragments from it, can dimerise *in vitro*. However the 817GGGG<sub>820</sub> tract appeared to be involved in all tetrad mediated RNA dimerisations. This was tested by mutagenesis.

### **3.8. Mutagenesis of an Essential Guanine Residue Abolished Tetrad Mediated Dimerisation.**

Further confirmation that RNA 2 dimers contained a guanine tetrad was obtained by mutagenesis of the essential 817GGGG<sub>820</sub> sequence. The G<sub>819</sub> to

U mutation was introduced into RNA 2 by RT-PCR mutagenesis producing RNA 3.

The G<sub>819</sub> to T mutation was selected because this will disrupt or prevent the formation of the tetrad in the dimer, due to the incorporation of two uridine residues one from each RNA monomer. This change was in the middle of the tetrad, making the use of neighbouring guanine residues difficult. The G<sub>819</sub> residue is part of the codon triplet for the second glycine residue in the HIV-1<sub>NL4-3</sub> gag protein, figure 3.12. The mutation changes the wild type GGG triplet (occurs 64 times in the wild type genome) to GGT (occurs 30 times).

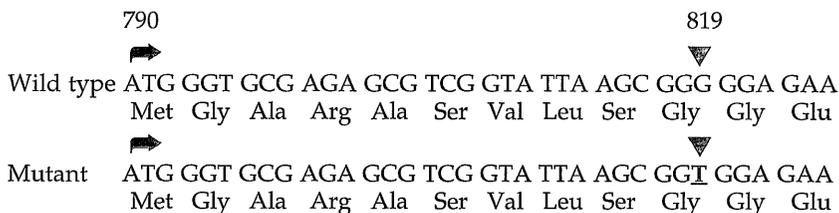


Figure 3.12. Gag coding sequence from the initiation codon at position 790 of HIV-1<sub>NL4-3</sub> gag. The G<sub>819</sub> to T mutation is shown.

Although other bases could be present in the wobble position and code for glycine, uridine was chosen because this was predicted to disrupt the guanine tetrad hydrogen bonding profile to the maximum extent. For example if adenine had been used then guanine-adenine Hoogsteen hydrogen bonding may have allowed the tetrad to remain (Guo et al., 1992).

### 3.8.1. The Dimerisation of RNA 3

A dimerisation time course of RNA 3 containing the G<sub>819</sub> to U mutation was performed in parallel with the wild type sequence RNA 2, as shown in figure 3.13.

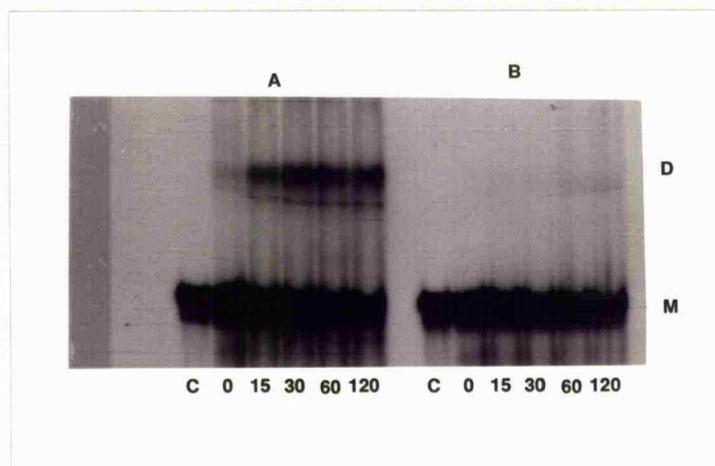


Figure 3.13. Dimerisation time course for RNA 2 panel A (wild type) and RNA 3 panel B (mutant). Lane C was the denatured RNA control. RNA 2 formed a dimer with time (min) as shown. RNA 3 containing the mutation dimerised inefficiently if at all. RNAs were analysed by electrophoresis on a 4% polyacrylamide gel.

RNA 3 dimerised very inefficiently if at all. Compared to the control dimerisation was drastically reduced, presumably due to the disruption of the tetrad requiring sequences  $731\text{GGGG}_{734}$  and  $817\text{GGGG}_{820}$ . Alternatively the mutation may have introduced other unknown structural changes, which may have indirectly prevented dimerisation.

There was a small fraction of dimeric RNA present in RNA 3 after a two hour incubation. The plasmid DNA for RNA 3 was sequenced showing that there was no wild type RNA present, or mutations other than expected. More likely as previously shown for RNAs 2 and 12, alternative guanine tracts may have been used to dimerise RNA 3. It remained possible that a tetrad had formed between  $731\text{GGGG}_{734}$  and  $817\text{GGUGG}_{821}$  in RNA 3, if the U residue was looped out from the structure allowing some G-quartet formation to occur.

Guanine tetrad formation was responsible for the dimerisation of several HIV-1 RNAs from sequences 708 to 858. If a guanine tetrad was the mechanism of dimerisation for HIV-1 RNA, then this feature should be present in all dimerisation competent RNAs around the DLS region, and perhaps by direct extrapolation, to genomic RNA. The 708-858 region is an extended DLS compared to that proposed earlier (Darlix et al., 1990) which was confined to sequences downstream from the splice donor site (position 746 in HIV-1<sub>NL4-3</sub>). The guanine tract at position 731-734 was important for *in vitro* RNA dimerisation and is located upstream of the splice donor site. Similarly the importance of upstream sequences was recognised by Marquet et al. (Marquet et al., 1994). Therefore the contribution of sequences upstream of nucleotide 708 to *in vitro* RNA dimerisation was investigated. This involved determining if a guanine tetrad was present in RNAs containing additional 5' sequences.

### **3.9. 5' Extended RNAs Failed to Dimerise by a Guanine Tetrad Mechanism.**

A comparison of RNA dimers with and without sequences upstream of the splice donor showed differences in cation dependence, rate of formation and thermal stability (Marquet et al., 1994). For example, dimerisation of a 707 nucleotide HIV-1 RNA was facilitated by lithium ions and not preferentially stabilised by potassium ions. This indicated that dimerisation of these RNAs was different to the tetrad RNAs 1, 2 and 12. The dimerisation of RNAs which contained sequences upstream of position 708, for example RNA 6 (649-858) was investigated.

#### **3.9.1. Dimerisation of RNA 6**

RNA 6 corresponds to sequences 649 to 858 ie. from the primer binding site extending into the gag gene. This RNA incorporated the tetrad

guanine tracts 731GGGG<sub>734</sub>, 817GGGG<sub>820</sub> and 836GGG<sub>838</sub>, figures 3.1 and 3.14.

RNA SEQUENCE	DIMER
4 649-813 <u>GC GCGC</u>	+
5 649-813 <u>AAACGC</u>	-
6 649-858 <u>GC GCGC</u> GGGG      GGGGG	+
7 649-858 <u>AAACGC</u> GGGG      GGGGG	-
8 649-825 <u>GC GCGC</u> GGGG      GGGGG	+

Figure 3.14. RNAs containing sequences upstream from position 708. The previously defined guanine tetrad tracts are shown. The kissing loop sequences (GCGCGC) are also indicated. Dimerisation ability of the RNAs is indicated by a + or -.

RNA 6 formed dimers on incubation under the same buffer I conditions as the tetrad RNAs 1, 2 and 12 (figure 3.15).

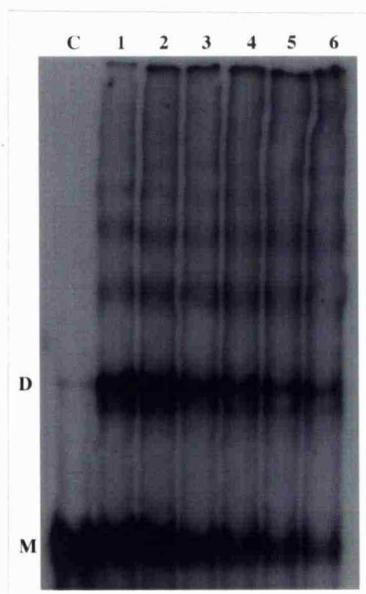


Figure 3.15. Dimerisation time course of RNA 6. RNA 6 formed dimers rapidly in buffer condition I. Lane C was denatured RNA control, lanes 1 to 6 correspond to times 5, 10, 20, 40, 60 and 120 min. RNAs were analysed on a 4% polyacrylamide TBE(K<sup>+</sup>) gel.

The formation of RNA 6 dimers was unexpectedly rapid, much faster than tetrad mediated RNA dimerisation. The amount of RNA monomer and dimer was reduced in lanes 5 and 6, probably due to RNase contamination. On this overexposed autoradiograph higher ordered RNA bands were also observed, often a feature of tetrad mediated RNA dimerisation.

The two strand nature of the RNA 6 dimer was demonstrated by heterodimerisation with RNA 4, these experiments are described later in section 3.21, figure 3.27. To determine if a guanine tetrad was involved in RNA 6 dimerisation, the thermal stability profile of the RNA 6 dimer was assessed in four different monovalent cation buffers,  $\text{Li}^+$ ,  $\text{Cs}^+$ ,  $\text{Na}^+$  and  $\text{K}^+$ , figure 3.16.

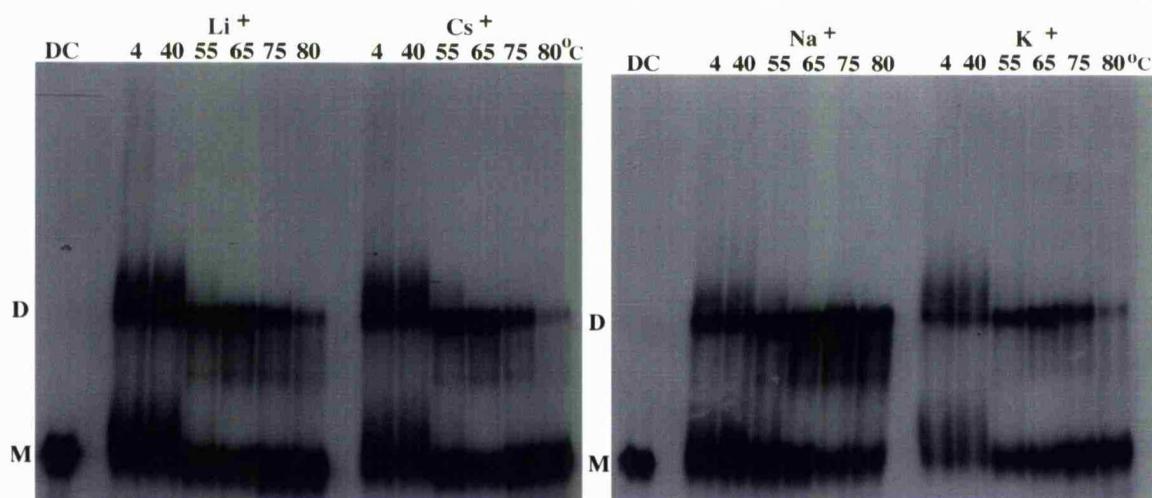


Figure 3.16. Cation dependent thermal stability of RNA 6. RNA 6 dimers were resuspended in buffers containing  $\text{Li}^+$ ,  $\text{Cs}^+$ ,  $\text{Na}^+$  and  $\text{K}^+$  ions, and incubated for 5 min at the temperatures indicated. Lane C was the denatured RNA control. Samples were analysed on a 4% polyacrylamide TBE ( $\text{K}^+$ ) gel.

No cation dependent thermal stability, or preferential stability in the presence of potassium ions was observed for the RNA 6 dimers. Similarly the reduced melting temperature of the dimers when Li<sup>+</sup> and Cs<sup>+</sup> were the solution cations was not observed. Dimer bands were more defined at temperatures around 55<sup>o</sup>C and the melting temperature in each buffer was 75 to 80<sup>o</sup>C. This suggested that RNA 6 dimer formation was not due to a guanine tetrad.

### **3.9.2. Dimerisation of RNA 8.**

To confirm that it was the sequences 649-708 (RNA 6 compared to RNA 2) that made an important contribution to RNA 6 dimerisation, RNA 8 was analysed. This RNA began at the same position as RNA 6 but terminated at position 825. RNA 8 dimerised, but did not show any cation dependent thermal stability (not shown). This demonstrated that sequences beyond residue 825 were not involved in the dimerisation of RNAs 6 and 8. Also a guanine tetrad had probably not formed in RNA 8 dimers, just like RNA 6. The sequences 649-708 were probably responsible for the new dimerisation features of these RNAs. DMS modification protection was performed on RNA 6 and 8 dimers (not shown). No protection of any previously identified guanine tract was observed, further suggesting that these RNAs dimerised by a non-tetrad mechanism.

To confirm that RNAs 6 and 8 could not dimerise by any previously defined guanine tetrad, RNA 4 was produced. This terminated at residue 813 so this RNA did not have the <sub>817</sub>GGGG<sub>820</sub> guanine tract which was essential for dimerisation of all tetrad RNAs.

### 3.9.3. Dimerisation of RNA 4.

RNA 4 (649-813) was constructed which terminated at residue 813, to completely remove the  $_{817}\text{GGGG}_{820}$  tract, figure 3.14. RNA 4 also dimerised shown in figure 3.17, and did not show any cation dependent thermal stability (not shown). This demonstrated that the mechanism of dimerisation did not require  $_{817}\text{GGGG}_{820}$  for RNA 4. Presumably this sequence was also not involved in the dimerisation of RNAs 6 and 8.

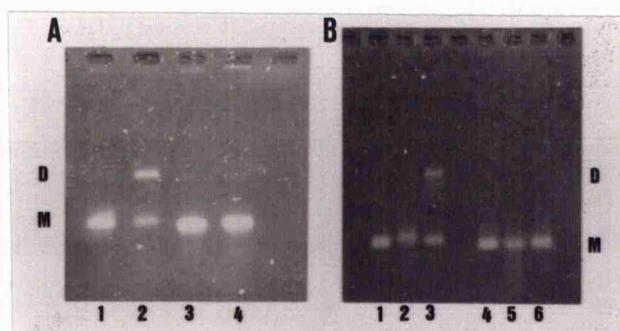


Figure 3.17. Panel A, RNA 4 dimerises, lane 2, compared to denatured control, lane 1. RNA 5 containing the kissing loop mutation fails to dimerise lane 4, lane 3 was the denatured RNA 5 control (described later).

Panel B, RNA 6 was able to dimerise on phenol extraction, lane 3, more rapidly than on incubation only for 5 min at 37°C in buffer II. Lane 1 was the denatured RNA 6 control. RNA 7 with the  $_{711}\text{GCG}$  to  $_{711}\text{AAA}$  mutation, lanes 5 to 6 was unable to dimerise, under any conditions (described later). Lane 4 was the denatured control. RNAs were analysed on a 3% NuSieve GTG agarose gel.

These observations were consistent with the lack of a guanine tetrad in the mechanism of dimerisation for RNAs containing sequences upstream of nucleotide 708.

### 3.10. Sequences Required for the Alternative Mechanism of RNA Dimerisation.

Sequence alignment was performed for RNAs 1, 2, 4, and 6 (figure 3.18), to compare the sequence requirements responsible for the dimerisation of these RNAs.

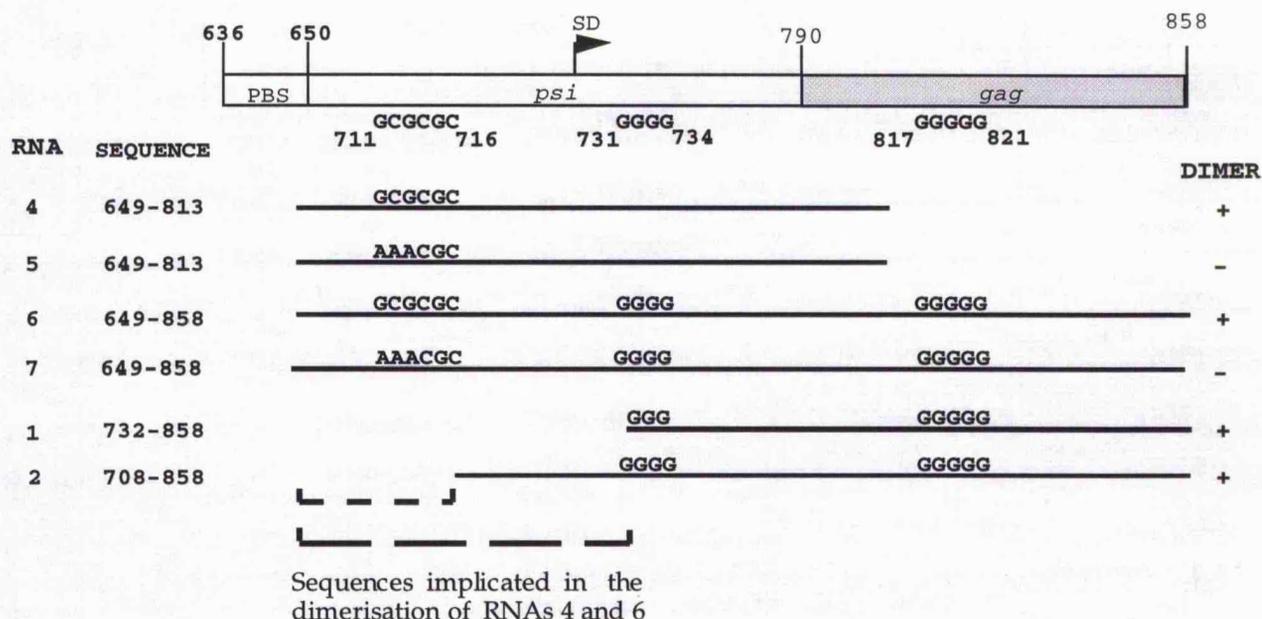


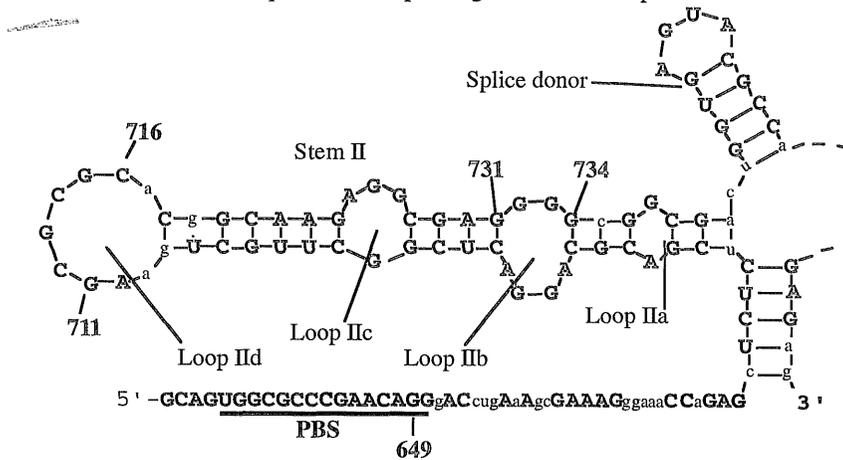
Figure 3.18. Sequence alignment of tetrad RNAs 1 and 2 and non-tetrad RNAs 4, and, 6. RNAs 5 and 7 are described later.

Comparing RNAs 4 and 6 with 1 and 2 suggested sequences from 649 to 708 or 732 must be implicated in the dimerisation of RNAs 4 and 6. The additional upstream 57 nucleotides (RNA 6 compared to RNA 2) had introduced markedly different dimerisation behaviour for example, a lack of cation dependent thermal stability, figure 3.16.

The sequences from 649 to 732 constitute Stem II and Loop II of the packaging signal secondary structure (Harrison et al., 1992), figure 3.19.

Adapted from: Harrison & Lever. 1992 J. Virol 66 4144-4153

RNAs 4 and 6 contain sequences corresponding to Stem and loop II



RNA 2, beginning at position 708 cannot form the stem loop II structure

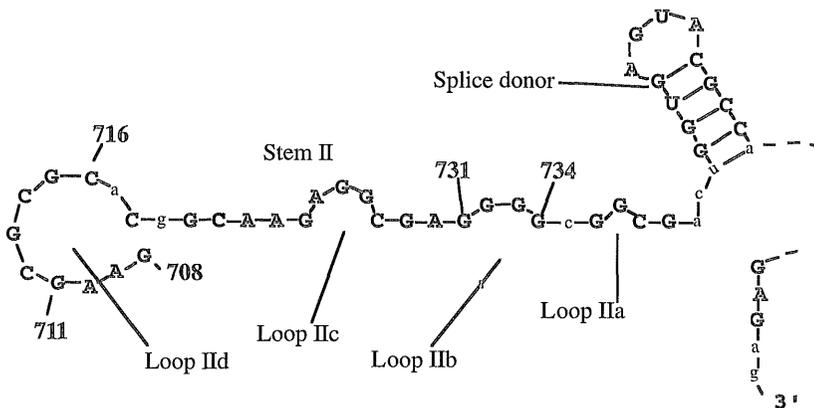


Figure 3.19. Sequences from position 649 are essential to form the Stem II and Loop II structure. These sequences were involved in dimer formation by a non-tetrad mechanism.

RNAs 4 and 6 may contain the stem II and loop II structure. For RNA 2, although it extended from 708-858 and contained the 711GCGCGC716 sequence, the stem loop could not form in this RNA, and was of course

absent from RNA 1 (732-858). The stem II and loop II structure was considered to be important for the dimerisation of RNAs 4 and 6, by a mechanism which was unidentified at this stage. However stem loop integrity was essential, and the loops probably interacted with one another to dimerise the RNA.

Shortly after this observation, the kissing loop model of RNA dimerisation was proposed by Skripkin et al (Skripkin et al., 1994) for HIV-1<sub>MAL</sub> RNA transcripts. The sequences and stem loop II region identified in the dimerisation of RNAs 4, 6 and 8 are consistent with this mechanism.

### **3.11. The Kissing Loop Model of HIV-1 RNA Dimer Formation.**

A mechanism to account for the dimerisation of a 707 nucleotide HIV-1<sub>MAL</sub> RNA encompassing the DLS and upstream sequences was proposed by Skripkin et al (Skripkin et al., 1994), and subsequently confirmed (Laughrea et al., 1994; Paillart et al., 1994; and Muriaux et al., 1995; 1996). Chemical modification of a 707 nucleotide RNA at the palindromic sequence GUGCAC prevented dimerisation (Skripkin et al., 1994). The kissing loop model was proposed in which a loop:loop interaction between two monomers was responsible for dimer formation. Dimerisation was initiated by the complementarity of the GUGCAC residues exposed at the end of a stem loop structure in each RNA monomer. Intra-strand stem base pairs then dissociated, and reformed in the dimer as inter-strand base pairs. Mutagenesis of the GUGCAC sequence to AAACAC prevented RNA dimerisation, presumably due to destruction of the complementarity required to initiate dimer formation. An alternative interaction has been proposed between the two double stranded helices. These may stack in a continuous helix, with the only inter-strand base pairs between the complementary loop sequences (Marino et al., 1995; Laughrea and Jette, 1996).

The kissing loop mechanism was a candidate for the dimerisation of RNAs 4 and 6. This was tested by mutagenesis.

### **3.12. Kissing Loop Mutagenesis Abolished RNA 4 and RNA 6 Dimerisation.**

To determine whether RNAs 4 and 6 dimerised by the formation of a kissing loop structure, mutagenesis was performed. The complementary loop sequences of HIV-1<sub>NL4-3</sub> are 711GCGCGC<sub>716</sub> essential for the initiation of dimerisation. This sequence was mutated to 711AAACGC<sub>716</sub> in RNAs 5 and 7 using RNAs 4 and 6 as a template by the RT-PCR approach described in materials and methods. This mutation abolished the complementarity which was required for dimer initiation (Skripkin et al. 1994).

The dimerisation ability of the wild type and mutated RNAs was compared. The RNAs were incubated in dimerisation buffer II at 37°C for 15 min then phenol extracted, and analysed by electrophoresis on native agarose gels. Wild type RNAs 4 and 6 dimerised, but the mutated RNAs 5 and 7 did not, figure 3.17 (Section 3.9.3).

The disruption of dimerisation by the mutations in RNAs 5 and 7, was good evidence that the dimerisation of RNAs 4 and 6 was due to the formation of a kissing loop structure requiring residues 711GCGCGC<sub>716</sub>. Additionally introducing the 711TTTCGC<sub>716</sub> mutation also prevented RNA dimerisation (not shown). This sequence was contained within the 649 to 732 sequences essential to the integrity of the stem II and kissing loop structure, figure 3.19.

In RNA 6, which contained both the kissing loop 711GCGCGC<sub>716</sub> and tetrad 731GGGG<sub>734</sub> and 817GGGG<sub>820</sub> sequences, the determinant for dimerisation was the kissing loop as RNA 7 failed to dimerise. This does

not exclude subsequent tetrad formation, however it was not observed under these experimental conditions.

### **3.13. Summary.**

Two possible mechanisms for HIV-1 RNA dimerisation from sequences 649-858 have been characterised. This region corresponds to the dimer linkage site (DLS, Darlix et al., 1990) and the extended version containing a dimer initiation site (DIS, Marquet et al., 1994 and Skripkin et al., 1994).

RNAs 1, 2, 10, 11 and 12 were able to dimerise *in vitro* by the formation of a guanine tetrad structure based on evidence from mutagenesis, cation dependent thermal stability, enhanced thermal stability in the presence of potassium, and DMS modification protection profiles. Guanine tracts 731GGGG<sub>734</sub>, 817GGGG<sub>820</sub> and 836GGG<sub>838</sub> had important roles in the formation of the RNA dimers. This mechanism of dimerisation was disrupted by the introduction of a single point mutation, G<sub>819</sub> to U in RNA 3. The evidence for a tetrad which formed *in vitro* was good, however the functional redundancy of the guanine tracts was unexpected. Two populations of tetrad dimers were indicated for RNA 2 involving sequences 731GGGG<sub>734</sub>/817GGGG<sub>820</sub> and 817GGGG<sub>820</sub>/836GGG<sub>838</sub>.

Retroviral RNA dimerisation by the formation of a guanine tetrad has been questioned because *in vitro* transcribed HIV-2 and BLV RNAs dimerised in the absence of any PuGGAPuA consensus (Berkhout et al., 1993; Katoh et al., 1993). Also no cation dependent thermal stability was observed for isolated HIV-1 genomic RNA (Fu and Rein, 1994) and RSV RNA (Lear et al., 1995) suggesting that a guanine tetrad was not part of virion RNA. However no specific mutations designed to disrupt tetrad formation have been introduced into infectious viruses. This is addressed here in Chapter 4, using the G<sub>819</sub> to U mutation.

RNAs 4, 6, and 8 dimerised by an alternative mechanism, that required sequences from position 649. This region was essential to the formation of a stable stem loop structure, with the exposed sequences  ${}_{711}\text{GCGCGC}_{716}$  at the apical tip of the loop. These RNAs dimerised by the formation of a kissing loop structure. The dimers could form on incubation in buffer II and on phenol extraction. Dimerisation was abolished by the  ${}_{711}\text{GCG}$  to  ${}_{711}\text{AAA}$  mutation (RNAs 5 and 7) consistent with the requirement of the self complementarity of this region to initiate dimer formation.

Evidence for the kissing loop structure *in vitro* has been compelling. Mutagenesis of the complementarity in the loop abolished dimerisation, which could be restored by the compensatory changes. (Skripkin et al., 1994, Muriaux et al., 1995). The integrity of the stem loop was similarly analysed by mutagenesis. Heterodimer formation between RNAs of different lengths, or between RNA mutants which could not efficiently form homodimers, further validated this mechanism of RNA dimerisation (Muriaux et al., 1995). To determine if the kissing loop mutation  ${}_{711}\text{GCG}$  to  ${}_{711}\text{AAA}$  had any effects on virion RNA dimerisation, this was introduced into an infectious HIV-1 virus, described in Chapter 4.

The presence of two putative dimerisation structures, the kissing loop (upstream) and the purine tetrad (downstream), lead to the suggestion that the kissing loop acts as a dimerisation initiation site (DIS) which is essential for dimer formation (Skripkin et al., 1994). The tetrad if it occurs within the dimer linkage site (DLS), may have a role in stabilising the structure. Paillart et al. described the 'bi-partite' dimerisation domain of a 615 nucleotide HIV-1<sub>MAL</sub> RNA (Paillart et al., 1994). Deletion of two purine tracts ( ${}_{792}\text{GGG}_{794}$  and  ${}_{798}\text{GAGAG}_{802}$ ) around the initiation codon of the gag gene, significantly decreased the thermal stability of the HIV-1 RNA dimer. Only a modest decrease in the kinetics of dimer formation

was observed when the  $_{817}\text{GGGG}_{820}$  important for the dimerisation of the tetrad RNAs analysed here was removed. There was also no effect on dimer stability. The  $_{731}\text{GGGG}_{734}$  tract was not analysed as it is located in the DIS region, upstream of the splice donor and DLS. Conversely, Muriaux et al. (Muriaux et al., 1995) suggested that purine quartets were not involved in the dimer structure of a 178nt HIV-1<sub>LAI</sub> RNA. This RNA corresponds to sequences 673 to 855 of HIV-1<sub>NL4-3</sub>, and was a similar result to that reported for RNAs 4, 6 and 8. Laughrea and Jette proposed that the kissing loop structure was the primary, or core dimerisation domain (Laughrea and Jette, 1996). An HIV-1<sub>LAI</sub> RNA which terminated 8 residues 3' from the complementary loop sequences dimerised. Therefore no sequences from the DLS were required. To investigate whether there was a bi-partite dimerisation domain *in vivo*, the tetrad G819 to U and the kissing loop  $_{711}\text{GCG}$  to  $_{711}\text{AAA}$  mutations were both introduced into HIV-1 virus, as described in Chapter 4.

The dimerisation behaviour of RNAs *in vitro* in the absence of any nucleocapsid protein has been analysed. The dimerisation of the tetrad and kissing loop RNAs was also investigated when a peptide, corresponding to the functional domains of the HIV-1 NC protein, was included in dimerisation reactions.

#### **3.14. Peptide Assisted Dimerisation of *In vitro* Transcribed HIV-1 RNAs.**

*In vitro* transcribed RNAs from the DIS/DLS region of HIV-1 can dimerise in the absence of any viral proteins. The ability to dimerise is an intrinsic property of the RNA. However dimerisation can be assisted by the viral nucleocapsid protein, or peptides corresponding to it.

### **3.15. Introduction**

A 104 nucleotide RNA from the DLS of HIV-1<sub>MAL</sub> could dimerise in the absence of any proteins *in vitro* (Darlix et al., 1990). However the addition of nucleocapsid protein extracted and purified from HIV-1 virions, greatly facilitated the dimerisation of this RNA (Darlix et al., 1990). A recombinant NCp15:GST fusion protein overexpressed in and purified from *E.coli*, also assisted dimerisation of a 93 nucleotide HIV-1<sub>LAV</sub> RNA (Weiss et al., 1992). The effect was due to the NC protein, and not GST. Sakaguchi et al. showed the requirement of the nucleocapsid protein for the dimerisation of a 44 nucleotide HIV-1 RNA, and proposed the A-U base pairing model for HIV-1 RNA dimerisation (Sakaguchi et al., 1993). Further studies (DeRocquigny et al., 1992) demonstrated that a synthetic peptide, NCp7 corresponding to the p7 nucleocapsid protein of HIV-1<sub>LAV</sub>, greatly assisted dimerisation of, and tRNA annealing to *in vitro* transcribed HIV-1 RNA. Deletion analysis of the NCp7 peptide demonstrated that the functional residues were the basic amino acids (13VK and 29RAPRKKG35) surrounding the zinc finger motifs. A similar result was obtained using NCp10 peptides corresponding to the nucleocapsid protein of MoMuLV (DeRocquigny et al., 1993). The role of the  $\Delta$ NCp7 peptide corresponding to the NC protein of HIV-1, where the zinc fingers have been removed and replaced by glycine linkers, was investigated in the dimerisation of tetrad and kissing loop RNAs.

### **3.16. The Peptide $\Delta$ NCp7.**

The peptide  $\Delta$ NCp7 (ADP 7015) was synthesised by Peptide & Protein Research Consultants, University of Exeter. It was judged to be 98% pure by HPLC analysis, and had a molecular weight of 3772.4 determined from mass spectroscopy which agrees well with the expected 3771.4.



The two zinc finger motifs (residues 16-28 and 36 to 50) in the wild type nucleocapsid protein are absent, and replaced by glycine linkers G-G joining residues 15 to 29 and 35 to 51. This peptide facilitated the annealing of the replication primer tRNA<sup>Lys,3</sup> to and dimerisation of, a 415nt HIV-1<sub>LAI</sub> RNA *in vitro* (DeRocquigny et al., 1992). The effect of the intact peptide was not replicated when separate peptides 13 to 35 and 51 to 72 were included in dimerisation reactions, so the correct positioning of the basic amino acids <sub>13</sub>VK and <sub>29</sub>RAPRKKG<sub>35</sub> was critical for activity. The role of the peptide to potentially facilitate dimerisation of the tetrad and kissing loop RNAs was investigated.

**3.17. The effect of the ΔNCp7 Peptide on the Dimerisation of RNAs 1, 2, 6, 8, 10, 11 and 13.**

RNAs 1, 2, 10, 11 and 13 dimerised by the formation of a guanine tetrad. The sequence tract <sub>817</sub>GGGG<sub>820</sub> was essential for the dimerisation of these RNAs. However the other guanine tract was variable, <sub>731</sub>GGGG<sub>734</sub> for RNA 2 (and presumably RNAs 11 and 13), and <sub>732</sub>GGG<sub>734</sub> for RNAs 1 and 10. Some contribution from sequence <sub>836</sub>GGG<sub>838</sub> was also identified in the dimerisation of RNA 2. The kissing loop RNAs, 6 and 8, dimerised by interstrand hydrogen bond formation involving a kissing loop interaction, for which residues <sub>711</sub>GCGCGC<sub>716</sub> were essential, figure 3.20.

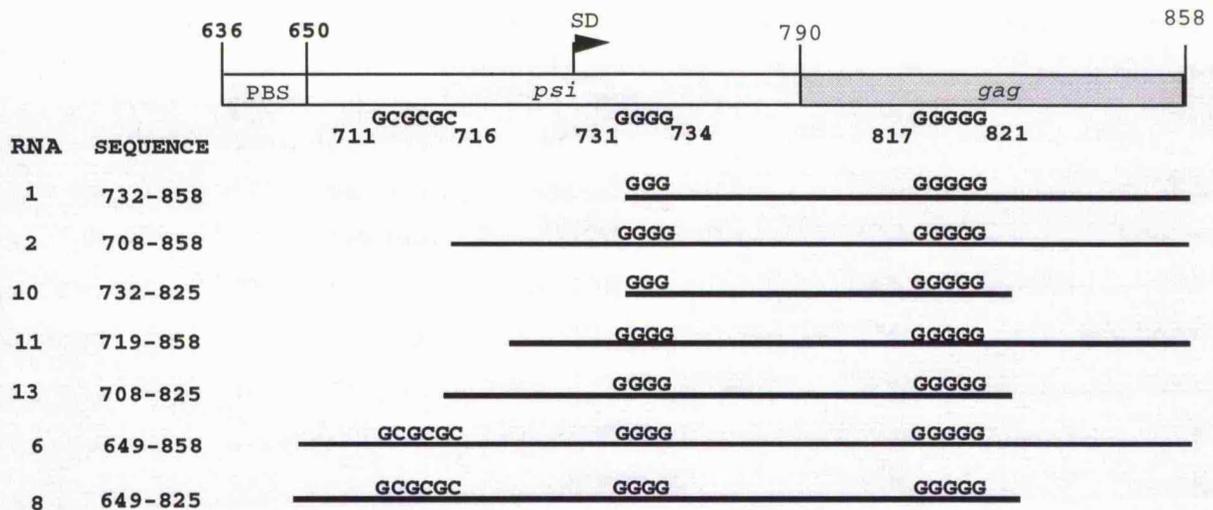


Figure 3.20. Tetrad and Kissing Loop RNAs used for the analysis of the  $\Delta$ NCp7 peptide:RNA interaction.

To determine if dimerisation of RNAs 1, 2, 6, 8, 10, 11 and 13 was assisted by the  $\Delta$ NCp7 peptide, samples of each RNA were incubated with and without  $\Delta$ NCp7 peptide. The RNAs at  $2\mu\text{M}$  concentration were incubated at  $37^\circ\text{C}$  in dimerisation buffer II for 15 min, with and without  $1\mu\text{g}$  ( $88\mu\text{M}$ ) of  $\Delta$ NCp7 peptide, figure 3.21. ProteinaseK was then added to a final concentration of  $25\mu\text{g/ml}$  to those samples containing peptide, incubated for an additional 15min at  $37^\circ\text{C}$ , then phenol and phenol/chloroform extracted as described (DeRocquigny et al., 1992). Loading buffer was added, and the RNAs electrophoresed on TBE agarose gels.

In lanes 1 to 5 (tetrad RNAs) without peptide there were no detectable higher ordered RNA structures. In lanes 6 and 7 (kissing loop RNAs) there was an upper band perhaps corresponding to RNA 6 and 8 dimers. This was expected for the kissing loop RNAs 6 and 8 which could dimerise in buffer II. However the tetrad RNA dimers did not form rapidly in this buffer.

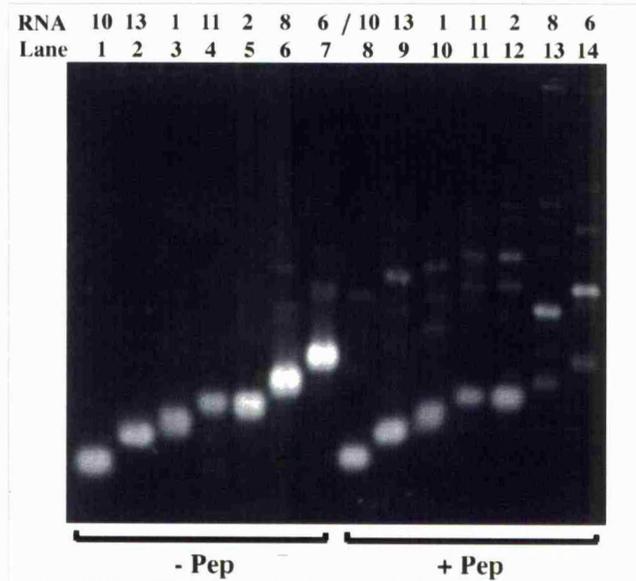


Figure 3.21. Peptide:RNA interaction. RNAs 1, 2, 6, 8, 10, 11 and 13 were incubated in dimerisation buffer for 15 min with (lanes 8 to 14) and without (lanes 1 to 7)  $\Delta$ NCp7 peptide. Proteinase K and organic extractions were performed on the RNAs +peptide prior to electrophoresis on 3% GTG agarose. Lanes 1 and 8, RNA 10, lanes 2 and 9, RNA 13, lanes 3 and 10, RNA 1, lanes 4 and 11, RNA 11, lanes 5 and 12, RNA 2, lanes 6 and 13, RNA 8 and lanes 7 and 14, RNA 6.

For those samples with peptide followed by organic extraction, higher molecular weight RNA bands were present in all lanes 8 to 14. In lanes 13 and 14 there was an increase in the amount of RNA 6 and 8 dimers, along with other higher ordered RNAs. These lower mobility bands (lanes 8 to 12) had an apparent molecular weight greater than their corresponding RNA dimer. These bands may be multimers of RNA or correspond to band shifts where the peptide remains bound to the RNA, perhaps due to incomplete digestion of the peptide with ProteinaseK .

This result suggested that the peptide had promoted the dimerisation of RNAs 6 and 8. The RNA:peptide interaction was analysed in more detail for these RNAs.

### **3.18. Interaction of the $\Delta$ NCp7 Peptide with Kissing Loop RNAs 6 and 8.**

There appeared to be an increase in dimer formation of the kissing loop RNAs 6 and 8 due to the  $\Delta$ NCp7 peptide. Focussing in detail on the effect of the peptide on RNA 8, the RNA was incubated at 2 $\mu$ M with 1 $\mu$ g of peptide in dimerisation buffer II at 37 $^{\circ}$ C, figure 3.22. Samples were removed at time t=0, 5, 10, 15 and 20 min. Peptide was digested with ProteinaseK at an elevated level of 100 $\mu$ g/ml at 37 $^{\circ}$ C for 15 min, followed by a phenol then phenol/chloroform extraction.

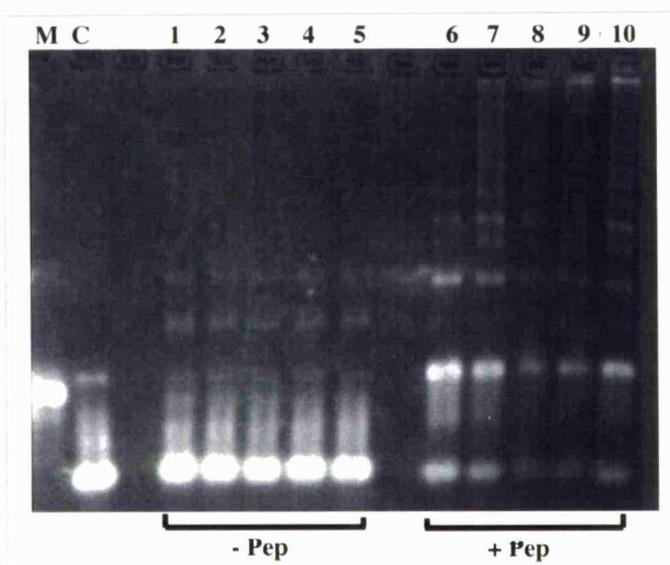


Figure 3.22. Effect of  $\Delta$ NCp7 peptide on the dimerisation of RNA 8. Aliquots of RNA were removed at times t=0, 5, 10, 15 and 20 min (lanes 1 to 5 without peptide and lanes 6 to 10 with peptide). Peptide was digested and removed from lanes 6 to 10 by ProteinaseK and organic extraction, followed by electrophoresis on 3% NuSieve GTG agarose gel. M was a 351nt RNA marker.

Comparing lanes 1 to 5 (without peptide) and 6 to 10 (with peptide) at time t=0, there was an increase in the dimer band for RNA 8 (lane 6 compared to lane 1). There was also an increase in those bands which trace back to the wells in lanes 7 to 10 which correspond to RNA samples removed at later times 5, 10, 15 and 20 min. The dimerisation could be due to the

peptide which assisted the dimerisation of a 104 nucleotide HIV-1 RNA within three minutes (Darlix et al., 1990). There were no higher ordered RNAs present in lanes 1 to 5 which did not contain peptide. Due to probable RNase degradation in lanes 8 and 9, and to detect any bands not readily visible from ethidium bromide stained agarose gels, this analysis was repeated using radiolabelled RNA 6.

RNA 6 containing additional downstream sequences compared to RNA 8 was incubated at  $2\mu\text{M}$ ,  $37^\circ\text{C}$  in buffer II for time  $t=0, 5$  and  $15$  min with and without  $1\mu\text{g}$  of  $\Delta\text{NCp7}$  peptide, figure 3.23.

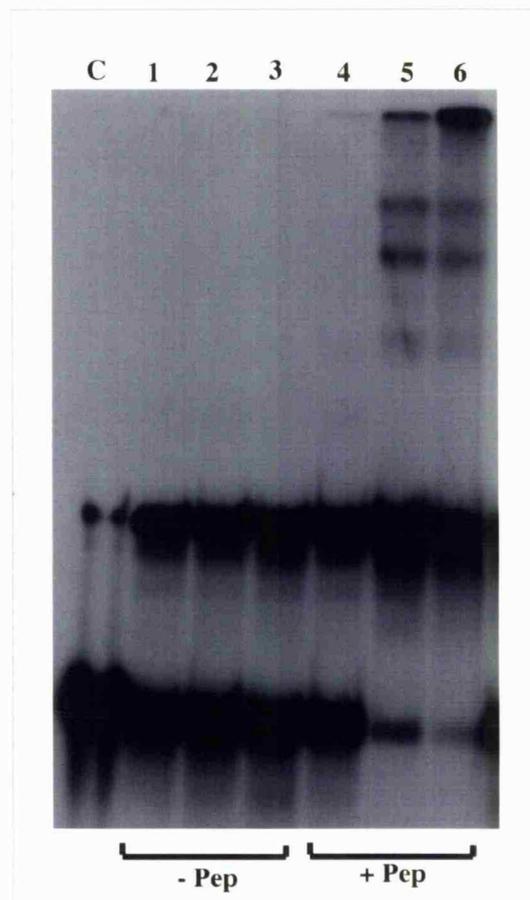


Figure 3.23. RNA 6 and  $\Delta\text{NCp7}$  peptide assisted dimerisation. Lane C is the denatured RNA marker, lanes 1 to 3 are without peptide, 4 to 6 are with. Samples were removed at times  $t=0, 5$  and  $15$  min (lanes 1-3 and 4-6), then prepared for electrophoresis and run on a 4% polyacrylamide TBE gel.

Following ProteinaseK and organic extraction of the samples, in lanes 4 to 6 (those with peptide) there was band shifting back to the well, and an apparent reduction in the amount of RNA monomer present. Without peptide (lanes 1 to 3) there was no band shifting, the only upper band was the dimer RNA. The amount of RNA dimer appeared to be similar in lanes with and without peptide, compare lanes 1 to 3 with 4 to 6. The presence of a small amount of dimeric RNA in lane C may have been due to incomplete denaturation of this sample.

The production of slower migrating RNA bands after coincubation with the peptide, appeared to promote both dimerisation and/or multimerisation of the RNA. Alternatively, despite the harsh protein digestion and extraction procedures, the possibility remained that the upper bands corresponded to a complex pattern of band shifts of RNA and peptide. The specificity of peptide binding to the RNA was investigated.

### **3.19. Specificity of the Peptide:RNA interaction.**

To determine whether specific binding between the peptide and RNA had occurred, yeast tRNA acting as a competitor was included in the dimerisation reactions of RNA and  $\Delta$ NCp7 peptide. If peptide binding was non-specific the peptide binds competitor effectively removing itself from the RNA dimerisation reaction. Alternatively for specific binding the presence of the competitor (at the concentrations examined) is irrelevant and RNA bands which trace back to the wells should persist. The lack of the zinc finger residues and the high basicity of the peptide, would indicate that a non-specific interaction was more likely.

RNA 8 was incubated at 2 $\mu$ M in buffer II at 37°C, with and without peptide (1 $\mu$ g). Yeast tRNA was included in the reactions at 0, 2, 8, 16, 32

and 64 $\mu$ M concentration. The samples containing peptide were again ProteinaseK treated (100 $\mu$ g/ml) at 37 $^{\circ}$ C for 15 min then phenol and phenol/chloroform extracted, figure 3.24.

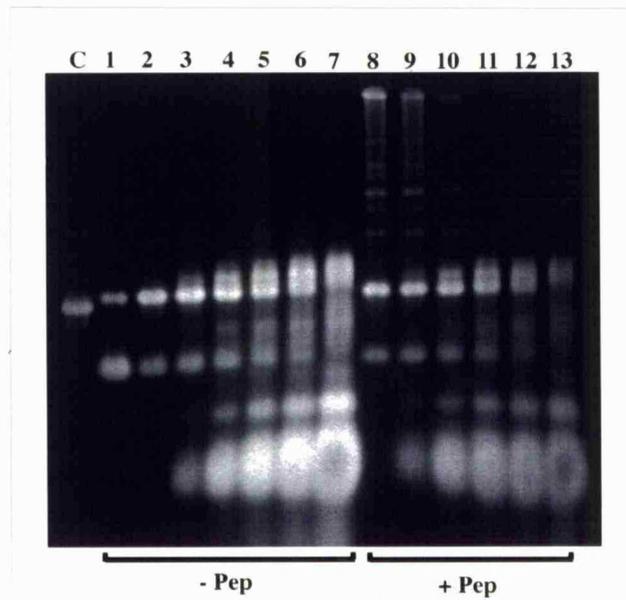


Figure 3.24. Specificity of  $\Delta$ NCp7 peptide:RNA binding. RNA 8 only (lanes 1 to 7) at 2 $\mu$ M and the RNA: $\Delta$ NCp7 peptide at 1 $\mu$ g (in lanes 8 to 13) were incubated in the presence of competitor yeast tRNA at 0, 2, 8, 16, 32 and 64 $\mu$ M (lanes 2 to 7 and 8 to 13). Peptide was removed from the reactions in lanes 8 to 13. Lane C was a denatured RNA marker of 351nt, lane 1 was RNA 8 electrophoresed immediately following transcription. Samples were analysed on a 4% NuSieve GTG agarose, TBE gel.

Lanes 2 and 8 corresponded to RNA incubated without yeast tRNA, again there were bands tracing back to the wells in the presence of the  $\Delta$ NCp7 peptide (lane 8), absent in lane 2 (without peptide). As the tRNA concentration increases (from left to right), this banding reduced (lanes 8 to 13) leaving only the RNA 8 monomer and dimer bands, lanes 11 to 13. This suggested that there was no specific interaction between the RNA and  $\Delta$ NCp7 peptide.

Non-specific binding between the RNA and peptide can explain the band shifts back to the wells if the Proteinase K and organic extraction techniques failed to remove all of the peptide. If the bands are RNA multimers, then the peptide: $\Delta$ NCp7 binding may be followed by NC:NC interactions which together multimerise RNA strands. However the presence of dimeric RNA cannot easily be explained in terms of the peptide. There are equivalent amounts of the RNA 8 dimer present irrespective of the  $\Delta$ NCp7 peptide, also seen for RNA 6 in figure 3.23.

Along with the addition of peptide, the RNA dimerisation reactions were Proteinase K digested then phenol and phenol/chloroform extracted. The presence of  $\Delta$ NCp7 peptide appeared to be irrelevant for dimer formation, so the extraction procedures may be contributing to the RNA dimerisation.

### **3.20. The Effect of Organic Extraction Procedures on RNA Multimerisation.**

Increased dimer RNA was present in those samples that had been incubated with the  $\Delta$ NCp7 peptide and subsequently ProteinaseK and phenol/chloroform extracted. Phenol extraction has been performed in other RNA dimerisation studies. Darlix et al. extracted the 104 nucleotide HIV-1<sub>MAL</sub> RNA with phenol in the absence of nucleocapsid protein this procedure did not promote dimerisation (Darlix et al., 1990). Similarly DeRocquigny et al. (DeRocquigny et al. 1992, and 1993) performed the same control. Two RNAs in their study failed to dimerise, demonstrating that organic extraction was not responsible for the dimerisation of all RNAs. To determine the effects of organic extraction on the kissing loop RNA oligomerisation, an experiment was performed using radiolabelled RNA 6. This RNA was incubated at 2 $\mu$ M concentration, 37°C in buffer II with

and without peptide (1 $\mu$ g) for 5 min. The samples were then ProteinaseK treated and phenol extracted, figure 3.25.

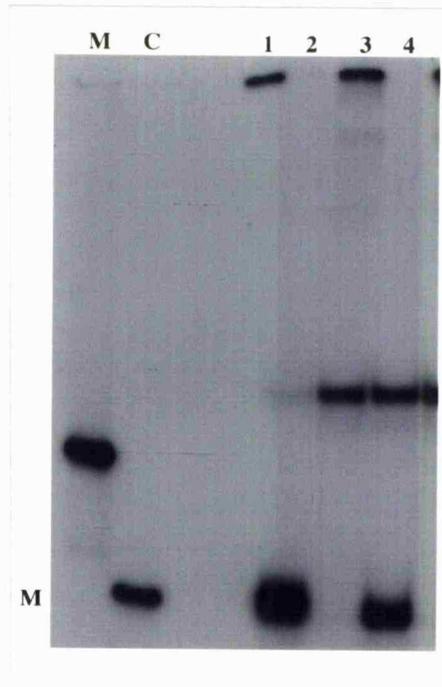


Figure 3.25. Effect of organic extraction on RNA 6 dimerisation. Lane M is the 351nt RNA marker, lane C the denatured RNA 6 monomer. Lanes 1 and 3 have peptide, 2 and 4 without. Lanes 1 and 2 were the RNA electrophoresed without any extraction procedures, in lanes 3 and 4 the samples were proteinaseK, phenol then phenol/chloroform extracted. RNAs were analysed on a 4% polyacrylamide gel.

In lane 1 there was a complete band shift of the RNA into the well. Lane 2 corresponds to the RNA only incubated for 15 min, there was a small amount of RNA dimer present. In lane 3 band shifts were present due to the peptide and there was also an increased amount of RNA dimer compared to lane 2. Lane 4 corresponds to RNA without peptide which was Proteinase K then phenol and phenol chloroform extracted. No band shifts have occurred, although there was a significant increase in RNA 8 dimer compared to lane 2. This suggested that the ProteinaseK or more likely the organic extraction procedure was promoting dimerisation.

Curiously there was no monomeric RNA present after extraction in the sample that contained peptide in lane 3, although the RNA monomer was clearly present after organic extraction only, lane 4. This indicated that the peptide remained complexed to the RNA after extraction at the origin of the gel.

To confirm this result, an experiment was performed where RNA 6 was incubated at 2 $\mu$ M for 15min at 37 $^{\circ}$ C with and without  $\Delta$ NCp7 peptide (1 $\mu$ g), figure 3.26. The reactions were then extracted with phenol and phenol/chloroform in the combinations shown.

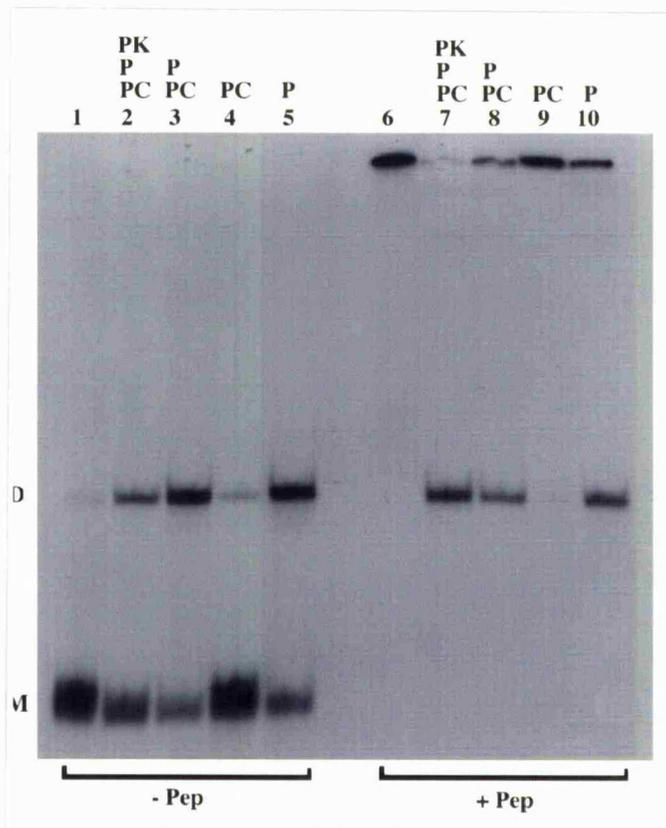


Figure 3.26. Effect of organic extraction procedures on RNA 6 dimer formation. Lanes 1 to 5 are without peptide, lanes 6 to 10 are with. RNA was electrophoresed without any extraction procedure (lanes 1 and 6), in lanes 2 and 7, ProteinaseK (PK), phenol (P) and phenol/chloroform (PC) extractions were performed. Lanes 3 and 8 were phenol and phenol/chloroform treated, lanes 4 and 9 were phenol/chloroform extracted, and lanes 5 and 10 were just extracted with phenol. RNAs were analysed using a 4% polyacrylamide TBE gel.

Lane 1 shows the RNA after incubation for 5 min at 37°C. This was completely complexed in the well when peptide was present, lane 6. Following ProteinaseK then organic extraction, there was a single band corresponding to the RNA dimer (lane 7), presumably due to complete peptide removal from the RNA. In lane 2, these techniques have produced a similar dimer band (much increased compared to lane 1), in the absence of any peptide. Monomer RNA was again present in lane 2 (without peptide), and absent in lane 7 (with peptide). Lanes 3 to 5 and 8 to 10 demonstrated that dimer formation was greatly assisted by phenol extraction. Phenol/chloroform treatment only appeared to produce reduced amounts of RNA dimer (lanes 4 and 9), than a combination of phenol and phenol/chloroform (lanes 3 and 8) or phenol only (lanes 5 and 10). This effect occurred in the presence or absence of peptide and showed that the extraction procedures were responsible for dimerisation of this RNA. Monomeric RNA was only present in those samples which were not treated with peptide, it was presumably band shifted at the origin of the gel in those incubations that contained  $\Delta$ NCp7 peptide. The effect of phenol was a surprising result. It had not dimerised the tetrad RNAs, and was apparently specific for the kissing loop structure.

### **3.21. Phenol Extraction Facilitated Kissing Loop Dimer Formation.**

To determine whether phenol specifically enhanced dimerisation by the formation of interstrand contacts according to the kissing loop mechanism, a heterodimer formation experiment was performed between kissing loop RNAs 4 and 6. These two dimerisation competent RNAs of different lengths, and therefore electrophoretic mobilities, were incubated together and analysed by electrophoresis following phenol extraction.

### **3.21.1. Phenol Induced Heterodimerisation of RNAs 4 and 6.**

RNA 4 and RNA 6 dimerised when incubated in dimerisation buffer II followed by phenol extraction. This produced an RNA 4 dimer band consisting of two RNA 4 molecules, and similarly RNA 6 formed its homodimer. The position of these dimers on electrophoresis was different due to the sizes of the RNAs. RNA 4 (164 nucleotide) had a 328 nucleotide dimer which migrated ahead of the RNA 6 (209 nucleotide) 418 nucleotide dimer, figure 3.27.

When RNA 4 and RNA 6 were incubated together at opposing and different concentrations a single heterodimeric RNA was formed. This band had an intermediate electrophoretic mobility between the RNA 4 and RNA 6 homodimers. This heterodimeric band consists of one RNA 4 and one RNA 6 strand ( $164+209=373$  nucleotide).

A single heterodimeric band (373 nucleotide) between the RNA 4 (328 nucleotide) and RNA 6 (418 nucleotide) homodimer positions was good evidence for the two-strand RNA nature of the RNA dimers. Heterodimer formation demonstrated that the sequences/structures required for dimerisation were derived from each monomeric RNA. Intermolecular contact was made using the complementary loop sequences  $7_{11}GCGCGC_{716}$ . Phenol extraction of these RNAs promoted genuine RNA dimer formation.

The effect of phenol appeared to be specific for the kissing loop structure. Tetrad RNAs or those in which the kissing loop residues were mutated eg. RNAs 5 and 7, did not dimerise with phenol. Further evidence that a kissing loop structure was required for phenol assisted dimerisation, was from the analysis of an RRE 33 RNA.

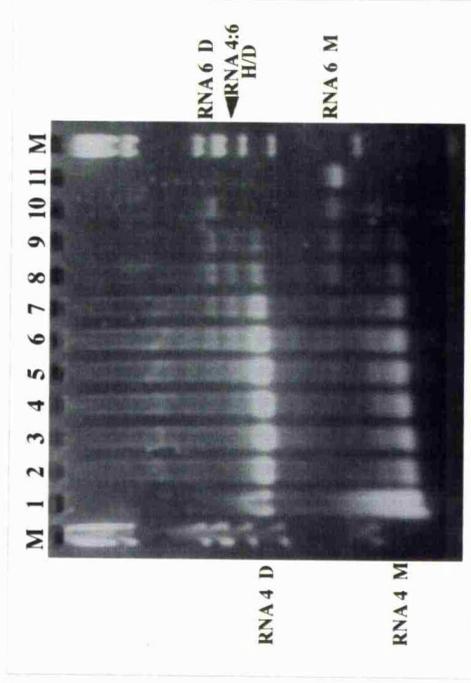
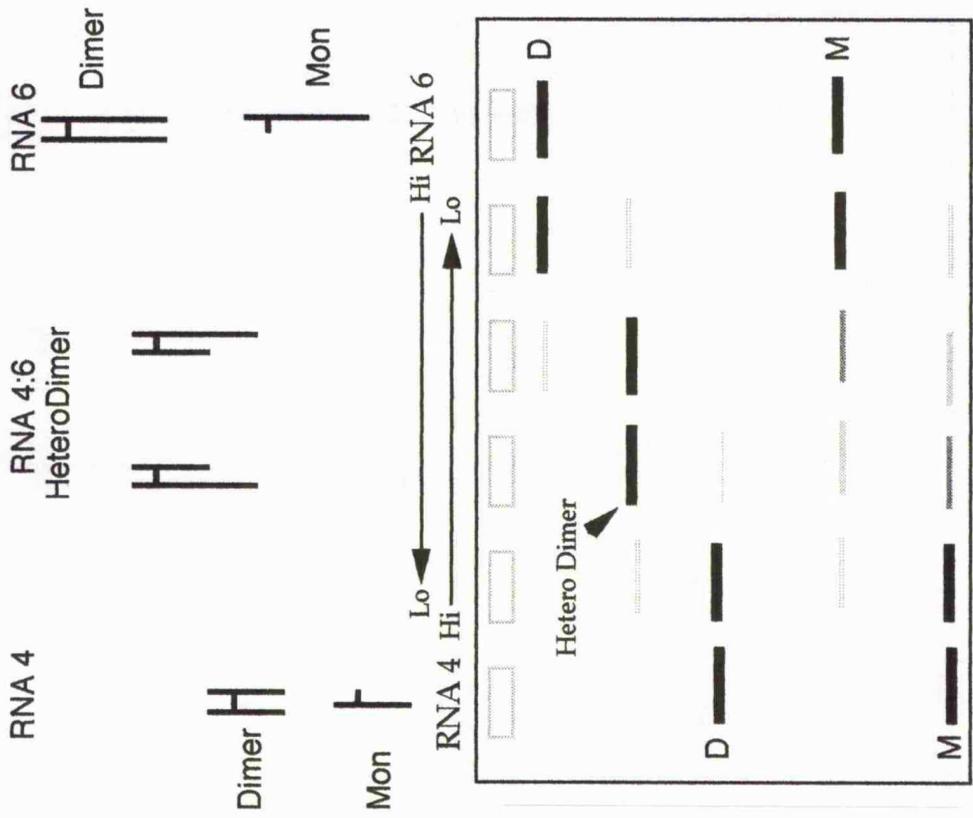


Figure 3.27. Phenol induced heterodimerisation. RNAs 4 and 6 were incubated together at opposing concentrations ranging from 0.5 to 5 $\mu$ M in lanes 3 to 9 and were phenol extracted. On electrophoresis a single heterodimeric band (arrow) is visible in lanes 8 to 5 showing that phenol extraction does promote RNA dimerisation. Lanes 1 and 11 were denatured RNAs 4 and 6 respectively, lanes 2 and 10 were the RNA dimerisation controls. RNAs were electrophoresed through a 4% NuSieve GTG agarose TBE gel.

### 3.21.2. Phenol Extraction Failed to Dimerise RRE 33 RNA.

RRE 33 RNA (Heaphy et al., 1991) corresponds to part of the Rev Response Element for HIV-1. This small RNA has a stem loop structure, but does not contain a kissing loop. Therefore RRE 33 RNA should not dimerise with phenol, figure 3.28.

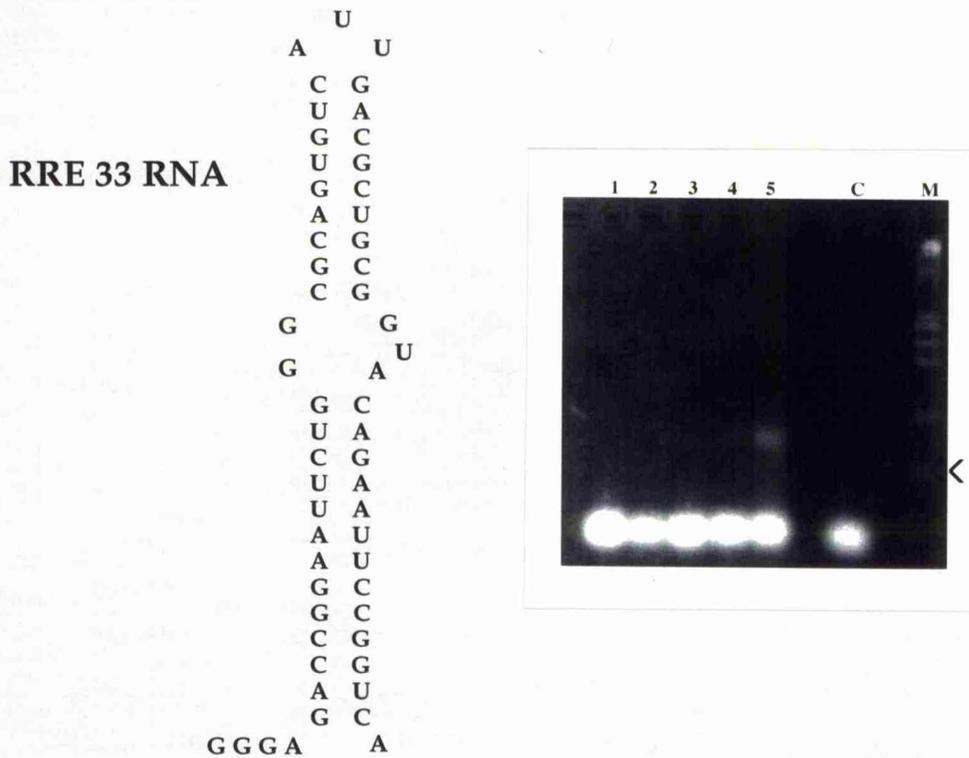


Figure 3.28. Dimerisation of RRE 33 RNA. The structure of the RNA is shown. Lane C was the denatured RNA. Lanes 1 and 2 correspond to the RNA after incubation and phenol extraction, lane 3 was after incubation only in buffer II. Lane 4 was the RNA analysed under tetrad buffer I conditions, lane 5 was the RNA analysed after precipitation from the transcription reaction. Samples were run on a 4.5% NuSieve GTG agarose TBE gel.

The RRE 33 RNA failed to dimerise under any of the incubation conditions which allowed the dimerisation of tetrad or kissing loop RNAs. This demonstrated that phenol induced dimerisation required a kissing loop structure. No tetrad RNAs were dimerised with phenol, as these RNAs lack the stem II and loop II sequences. The presence of a

higher ordered RRE 33 band following transcription was probably due to the presence of spermidine in the transcription buffer, lane 5. Spermidine is positively charged and may easily allow aberrant RNA:RNA associations by masking the negative charges between the RNAs.

### **3.22. Mechanism of Action of Phenol on RNA Dimerisation.**

Phenol extraction promoted the dimerisation of RNAs which dimerised by the formation of a kissing loop structure. This effect was specific as related RNAs which dimerised by the formation of a guanine tetrad, and an unrelated RRE 33 RNA sequence, were not dimerised on phenol extraction. Heterodimer formation between two RNAs of different lengths (RNAs 4 and 6) demonstrated that genuine interstrand contacts were made between the two monomers.

The mechanism by which phenol assisted dimerisation was unknown, but several possibilities exist. Phenol extraction may partially denature the RNA, allowing the interstrand base-pairing required by the kissing loop model to more easily occur, generating a more stable RNA dimer. It could be that phenol extraction acting like a denaturant, dissociated regions of the RNA structure shifting the monomer:dimer equilibrium towards dimer formation.

Alternatively phenol may indirectly promote RNA dimerisation. During extraction, the organic phenol and aqueous solution are emulsified. This phase disruption may effectively increase RNA concentration. It has been reported that Rous sarcoma virus RNA transcripts were dimerised by ethanol precipitation, this procedure may in a similar way increase the effective RNA concentration (Lear et al., 1995).

Whatever the mechanism, these techniques are clearly not part of any biological process. However it could be possible that RNAs are forced together effectively increasing their concentration, when the viral core

undergoes maturation. Additionally the nucleocapsid protein binding to the RNA, may drive dimer formation due to protein:protein interactions effectively condensing the RNA into a more compact chromatin like structure.

### **3.23. Summary.**

Assisted dimerisation due to a zinc finger less peptide  $\Delta\text{NCp7}$  corresponding to the nucleocapsid protein of HIV-1<sub>LAL</sub>, was investigated for RNAs that dimerised by the formation of a tetrad or a kissing loop structure. These RNAs contained sequences from the PBS, DIS/DLS/PSI and gag regions. All RNAs bound the peptide as judged from RNA band shifts tracing back to the wells in native gels. This interaction was non-specific, the inclusion of competitor yeast tRNA abolished the RNA: $\Delta\text{NCp7}$  association. Presumably the non-specific interaction was between positively charged peptide, and the negatively charged RNA backbone. The zinc finger motifs have been implicated for RNA recognition during packaging (Meric et al., 1989), these were absent from the  $\Delta\text{NCp7}$  peptide probably responsible for the non-specific binding.

No peptide assisted dimerisation of the tetrad RNAs was observed. However dimerisation of the kissing loop RNAs was apparently favoured in reactions containing the peptide. In fact this was not due to the peptide. Rather organic extraction procedures designed to eliminate peptide from the reaction prior to electrophoresis, assisted dimer formation. There was no evidence of facilitated RNA dimerisation due to the peptide, in contrast to other observations (DeRocquigny et al., 1992; 1993). For the kissing loop RNAs 6 and 8, phenol extraction greatly facilitated RNA dimerisation. The inability to dimerise tetrad RNAs with phenol probably relates to the absence of the kissing loop structure from these RNAs. Phenol induced dimerisation was specific, as RRE 33 RNA failed to

dimerise. Phenol promoted the complementary loop residue base pairing between RNAs 4 and 6, but failed to dimerise RNAs 5 and 7 where the loop complementarity had been disrupted by mutagenesis.

No further experiments were performed to investigate assisted RNA dimerisation by Nucleocapsid protein or peptides. This was because the organic extraction procedures necessary to remove these proteins, assisted dimer formation of the kissing loop RNAs. The organic extraction or the effect of the nucleocapsid protein could not easily be distinguished for their contributory role(s) in RNA dimerisation.

## Chapter 4: Testing *in vitro* Models of HIV-1 RNA Dimerisation *in vivo* using Virus Like Particles (VLPs) and Mutant HIV-1 Viruses.

### **4.1. Introduction.**

The mechanism of RNA dimerisation *in vivo* has not been identified for any retrovirus. However from an analysis of *in vitro* transcribed RNA dimerisation, the guanine tetrad and the kissing loop models have been proposed.

From the *in vitro* analysis of HIV-1 RNAs around the DIS/DLS region, two mutations were identified which significantly reduced dimerisation. The G<sub>819</sub> to U mutation introduced into RNA 2, abolished dimerisation by disruption of the guanine tetrad structure requiring residues <sub>817</sub>GGGG<sub>820</sub>. The <sub>711</sub>GCG to <sub>711</sub>AAA mutation introduced into RNA 6, abolished dimer formation by disrupting sequence complementarity essential for the kissing loop interaction. RNA dimers formed *in vitro* and *in vivo* are reported to have similar morphologies (Prats et al. 1990), thermal stabilities (Darlix et al. 1992) and chemical modification patterns (Tounekti et al 1992; Alford et al. 1991). Therefore having defined the disruptive effect of the mutations on RNA dimerisation *in vitro*, they were introduced *in vivo* to identify whether the nucleotides G<sub>819</sub> or <sub>711</sub>GCG were important for HIV-1<sub>NL4-3</sub> genomic RNA dimerisation.

Conservation of the tetrad and kissing loop sequences suggested that they have some important function (described later in the discussion). It was conceivable that the introduction of the mutations into these conserved regions of HIV-1 virus may prevent virus replication, RNA packaging, or RNA dimerisation. These viruses would be unable to replicate (null phenotype), and analysis of the RNA content would be impossible. For

these reasons, a Virus Like Particle (VLP) system was developed, with the intention of introducing the dimerisation mutations. In this system, viral assembly was decoupled from replication, so RNA packaging and dimerisation should occur without requiring viral replication.

#### **4.2. Virus Like Particles.**

Vector systems based on complementation of genetic deletions have been used to generate non-infectious HIV-1 particles (eg. Chen et al., 1992), however these require the use of infectious HIV-1 helper virus in permissive cells. Transient particle production systems can also be used and have advantages over stable expression in cell lines due to stability of viral protein expression (Krausslich et al., 1993). Therefore a transient particle production system was constructed using mammalian cells, so that native HIV-1 proteins with authentic modifications should be produced and released.

The construct pBCCX CSF X4 described in materials and methods (section 2.6.2, figure 2.5), has the strong CMV IE promoter which drives the expression of the HIV-1<sub>JRC<sub>SF</sub></sub> insert. This contains all the HIV-1 genes except the LTRs and the 3' exon of the non essential *nef* gene. As a result the vector directs the expression of all virus proteins (except *nef*) which can assemble into complete, but non- infectious particles. There is an 84 nucleotide CMV encoded RNA leader before the HIV-1 sequence begins at position 762 within the PBS (figure 4.1). This means that the 'genomic' RNA (approx 8500 nucleotides) of the VLP contained non-native nucleotides as well as the sequences thought to be involved in RNA packaging and dimerisation including the tetrad and kissing loop tracts.

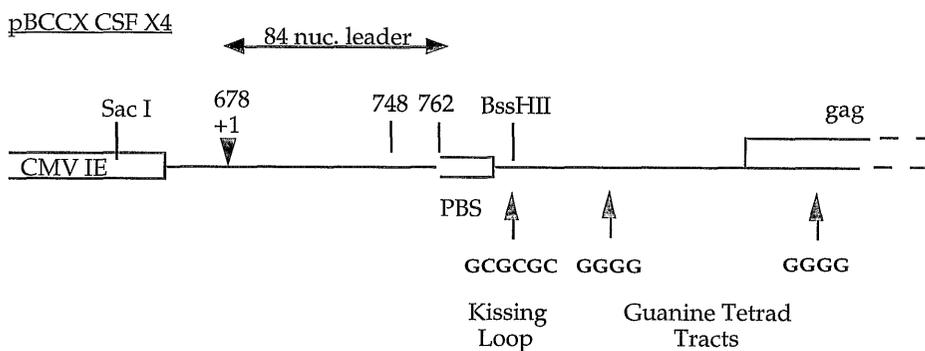


Figure 4.1. CMV:HIV DNA Junction in pBCCX CSF X4. The first nucleotide of the RNA is at +1. The HIV-1 sequence began at position 762 within the incomplete PBS.

Restriction digestion was performed on pBCCX CSF X4 to confirm that the structure of the plasmid was as expected, figure 4.2.

Digestions with EcoRI and XhoI linearised the vector as these were unique sites, lanes 3 and 4 respectively. BamHI and EcoRI double digestion, lane 2, produced 9,727 and 1,888 bp bands. HindIII restriction, lane 1, produced four bands, of 5,673, 3,874, 1,615 and 453 bp (arrowed). The results of the restriction digestion analysis were consistent with the expected structure of the plasmid.

The VLP system was originally designed to study particle morphogenesis, so it was necessary to demonstrate whether particles were produced and were capable of packaging and dimerising VLP 'genomic' RNA.

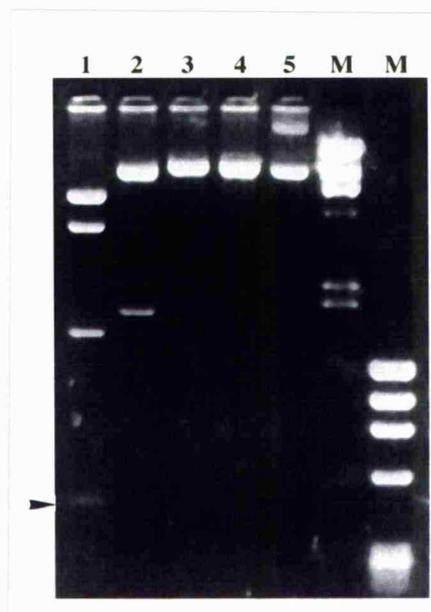
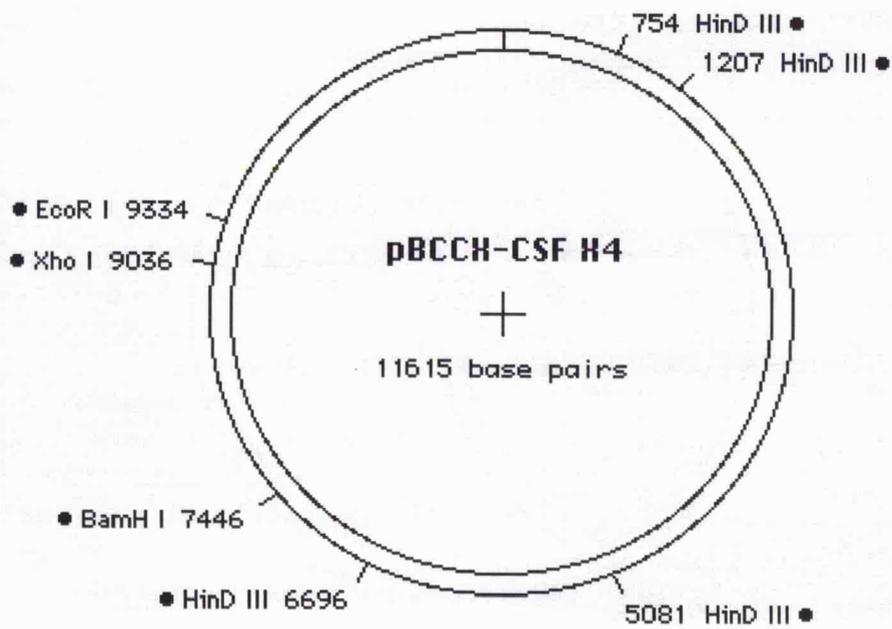


Figure 4.2. Restriction digestion of pBCCX CSF X4. 0.5ug of the plasmid was digested with HindIII, lane 1, BamHI and EcoRI lane 2, EcoRI lane 3, and XhoI lane 4. The products were analysed on a 0.8% agarose TBE gel, alongside the uncut plasmid (lane 5) λ HindIII, and φX174HaeIII markers.

### 4.3. Electroporation of pBCCX CSF X4 into COS Cells.

25µg of pBCCX CSF X4 plasmid DNA was electroporated into  $2.5 \times 10^6$  COS cells as described in materials and methods, and compared with the electroporation of an infectious molecular HIV-1 clone. Samples of culture supernatant were removed, filtered through a 0.45 µm Acrodisk, and analysed for p24 viral antigen. The construct gave rise to large amount of virus like particles as judged from the p24 titre, from about 48-72 hours after electroporation, when compared to proviral DNA clones. A representative plot is shown in the figure 4.3.

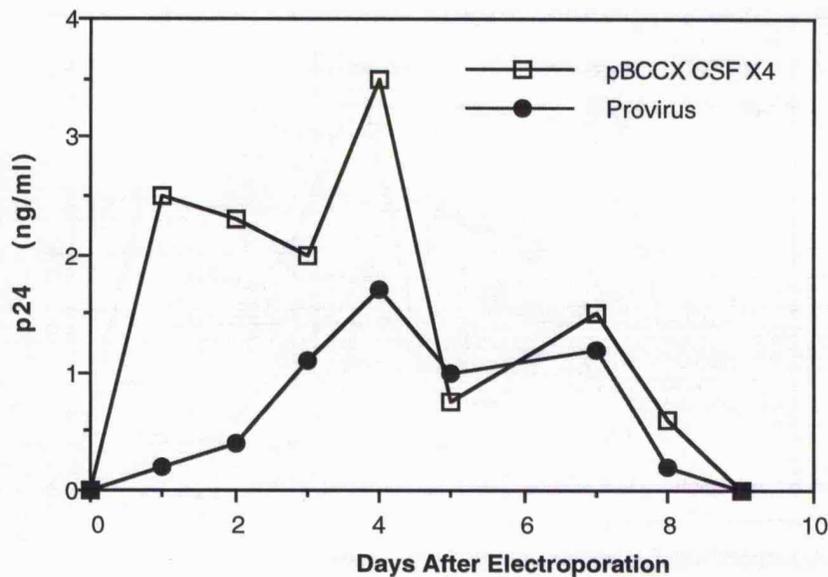


Figure 4.3. Electroporation of pBCCX CSF X4 into COS cells. pBCCX CSF X4 produced large amounts of VLPs from COS cells.

The increase in p24 protein was consistent with the production of VLPs from cells transfected with the pBCCX CSF X4 vector. Also similar results were obtained when the viral reverse transcriptase marker was assayed, figure 4.4, (performed by A. Lau).

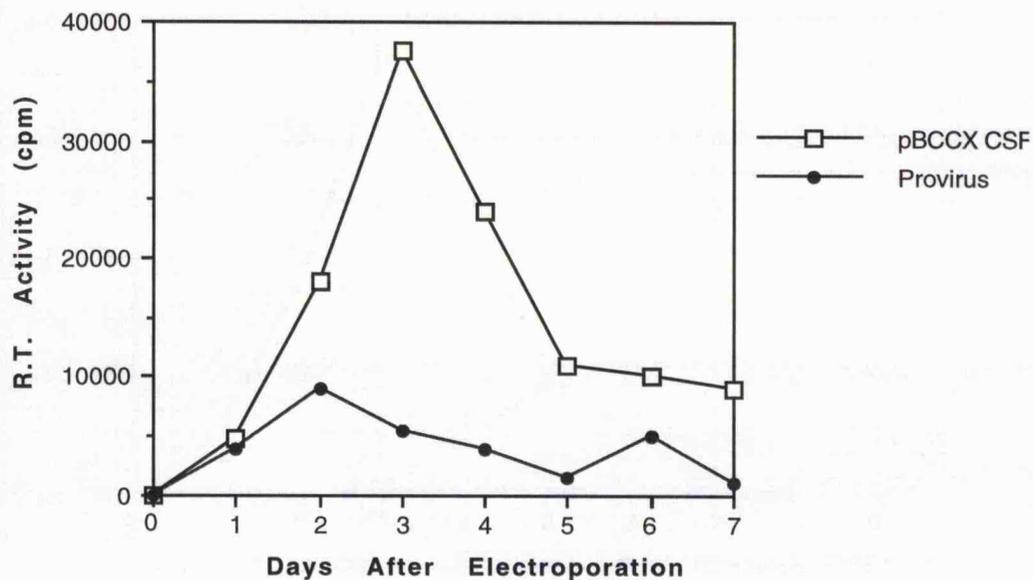


Figure 4.4. RT activity from the culture supernatant from cells electroporated with pBCCX CSF X4 and a HIV-1<sub>NL43</sub> proviral clone.

The signal due to reverse transcriptase was further evidence that VLPs were being produced. Reverse transcriptase is incorporated into assembling particles as part of the Pr 160<sup>gag-pol</sup> fusion protein (Wills and Craven, 1991), and is independent of genomic RNA packaging (Linial et al., 1978). Following protease activation, reverse transcriptase associates with the nucleocapsid structure and tRNA primer molecules (Barat et al., 1989).

#### **4.4. Physical Analysis of the VLPs.**

Electron microscopy and sucrose gradient centrifugation of the VLPs was performed to determine whether genuine particulate VLPs were produced from the electroporated cells.

#### **4.4.1. Electronmicrographs of the VLPs.**

To visualise VLPs from electroporated COS cells, the cells were fixed with glutaraldehyde, embedded in epoxy resin, sectioned and negatively stained with uranyl acetate. This analysis was performed by the staff of the Electronmicroscopy Unit, University of Leicester.



Figure 4.5 VLP Electromicrograph.

Particles were observed which appeared to be virus like, and could be present in several stages of development.

(1). At the surface of the COS cell a characteristic patch of dense material, perhaps polyprotein gag, can be seen aggregated underneath the cell membrane.

(2). These VLPs are at a later stage in the budding process. They appear immature, with an electron-lucent core, and may have just budded from the host cell membrane.

(3). This VLP has a conical, electron dense core suggesting that it was mature and could have completed morphogenesis.

Localisation of the VLPs to the plasma membrane was as expected. VLP assembly should occur at the underside of the plasma membrane, followed by budding and release from the cell. The appearance of particles which could be in several stages of development suggested that the VLPs produced could undergo correct morphogenesis, as expected.

VLPs from concentrated culture supernatant were also analysed by electronmicroscopy (M. Nermut, personal communication). In contrast to the pictures obtained from the transfected cells, few VLPs were visible. The inability to observe VLPs from concentrated supernatant was disappointing but possibly reflects the greater technical challenge required to obtain these samples in a suitable form for electronmicroscopy. Further electronmicroscopy was not pursued, however the particulate nature of the VLPs was also examined using sucrose density gradient centrifugation.

#### **4.4.2. Sucrose Density Gradient Centrifugation of VLPs.**

Electronmicroscopy, the production of p24 and reverse transcriptase from transfected cells indicated that particulate VLPs were produced. To confirm this the migration of VLPs through a sucrose density gradient was analysed. If particulate the VLPs should band to a single narrow position within the gradient, with a similar density to that reported for HIV-1 virus.

To determine the buoyant density of the VLPs, culture supernatant containing VLPs was pelleted through a 20 % sucrose TNE solution. The

VLPs were carefully resuspended in PBS and spun through a stepwise gradient from 60/50/40/30/ to 20 % sucrose in PBS for 3 hours at 166,000 g. The gradient was fractionated from the top into 150 ul aliquots, and p24 titres were determined. A sucrose gradient without VLPs was also spun in parallel. It was fractionated, and the density estimated by weighing 100ul from the 150ul aliquots on a three figure balance. A representative plot is shown below in figure 4.6:

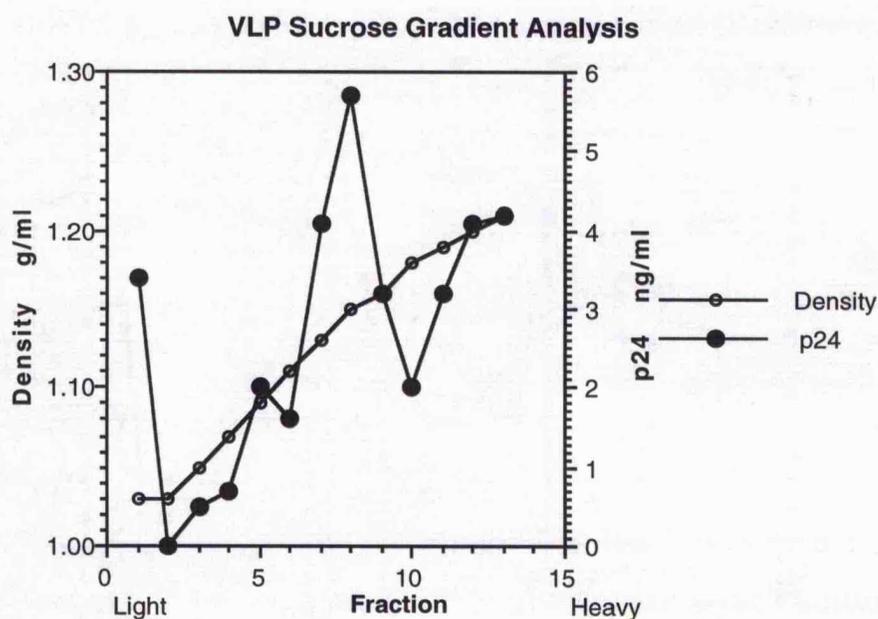


Figure 4.6. Sucrose density gradient analysis of VLPs.

Across the gradient (left to right) the p24 profile probably corresponds to free p24 protein, perhaps from unassembled or degraded VLPs. Then a p24 peak, fractions 7 to 9, was consistent with the particulate nature of the VLPs that had a density of about 1.15 g/ml. This was a similar value to that reported for HIV-1 virus (1.16 g/ml from Zhang et al., 1995). The presence of a p24 signal towards the heavier end of the gradient, could be due to higher molecular weight p24 protein aggregates or irregular VLPs. Sucrose gradient analysis was independently performed by M.Boyd (Chester Beatty

Laboratory, London; personal communication) with similar results to that obtained above, confirming that p24 protein was assembled into particles.

#### **4.5. RNA Content of the VLPs.**

To use the VLP system for an analysis of RNA dimerisation, it was necessary to determine if the VLPs were able to recognise and specifically package HIV RNA. The RNA content of the VLPs was initially investigated using RT-PCR. This cannot determine if the RNA was dimeric so in addition native Northern analysis was then performed.

#### **4.6. RT-PCR analysis of VLPs.**

RT-PCR was performed on concentrated VLPs. RNA was extracted from the VLPs by the RNazol method, to ensure removal of contaminating plasmid DNA which would interfere with the RT-PCR assay.

For RT-PCR detections appropriate controls were essential. Plasmid DNA was included to ensure that the PCR stage of the process worked. To avoid false positive results, samples containing water only were treated with and without reverse transcriptase to confirm the reagents were not contaminated. The reverse transcriptase stage was also assessed by including samples containing *in vitro* transcribed and gel purified RNA. The assay was based on the ability to detect a signal in those samples with RT (+RT) but not without RT (-RT). This shows that RNA had been detected. If DNA was present then a signal would be present in the samples treated both with and without RT.

##### **4.6.1. HIV-1 RNA was detected with the VLPs by RT-PCR.**

To determine if the VLPs contained HIV RNA, RT-PCR was performed using the primer sets:

MH6/SH175 amplifies a 245 bp fragment from the gag gene.

A3/A4 amplifies a 312 bp fragment from the pol gene.

PCR5/PCR20 amplifies a 440 bp fragment from the env gene.

Results obtained with the gag primers (MH6/SH175) are shown in figure 4.7

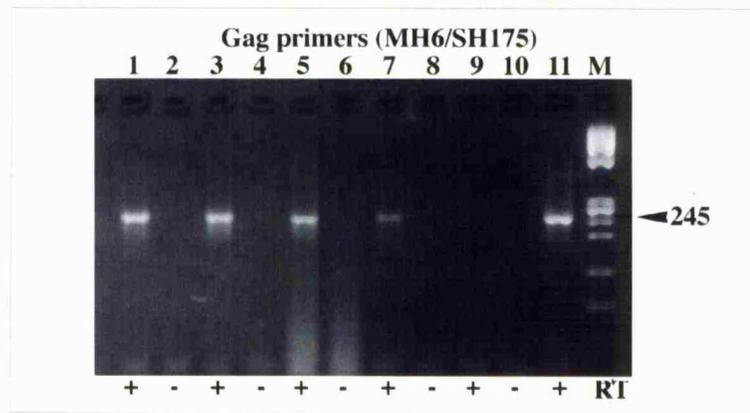


Figure 4.7 HIV-1 Gag RNA was detected with the VLPs. Odd numbered lanes are + RT, even -RT. The 245 bp band resulted from the detection of RNA in NL43 virus (lanes 1 and 2), VLPs (lanes 3 and 4), and cellular RNA from VLP producing cells (lanes 5 and 6). Control lanes correspond to RNA 6 (lanes 7 and 8), water blank (lanes 9 and 10) and PCR positive (plasmid DNA) in lane 11. M is HaeIII digested  $\phi$ X174 DNA. 3% Nusieve GTG agarose gel.

Lanes 7 to 11 were controls confirming that the PCR and RT stages worked, and there was no contamination from the reagents. HIV-1 Gag RNA was detected in lane 1 from wild type HIV-1<sub>NL43</sub> virus, as expected. This RNA was also present in total cellular RNA from the COS cells electroporated with pBCCX CSF X4 (lane 5), and associated with the RNA extracted from the VLPs (lane 3). A similar result was observed for pol and env primers, figure 4.8.

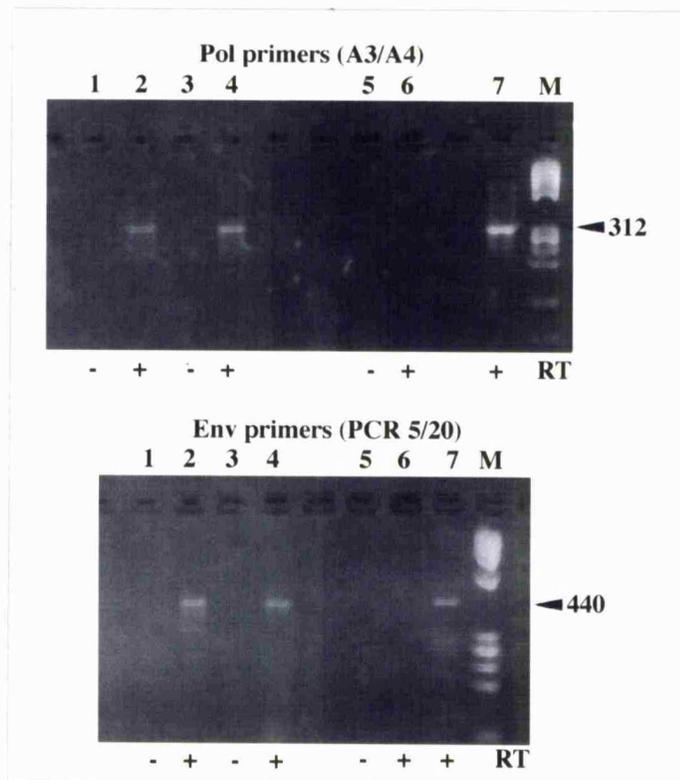


Figure 4.8. HIV-1 pol and env RNA was detected with the VLPs. Lanes 2, 4, 6 and 7 are with RT, 1, 3, 5 are without. The 312 bp band from the pol RNA was detected in the VLPs (lanes 1 and 2), and wild type HIV-1 (lanes 3 and 4). The 440 bp band corresponding to env was present for the VLPs (lanes 1 and 2), and wild type virus (lanes 3 and 4). Controls were water blank (lanes 5 and 6 ) and PCR positive lane 7. M is HaeIII digested  $\phi$ X174 DNA. 3% Nusieve GTG agarose gel.

Pol RNA was detected with the VLPs, lanes 1 and 2, and wild type HIV-1 virus, lanes 3 and 4. Similarly env RNA was associated with the VLPs in lanes 1 and 2, and for HIV-1 virus, lanes 3 and 4.

The amplified region using the MH6/SH175 primers contained sequences important for the incorporation of RNA ie. the packaging signal  $\Psi$  (Lever et al., 1989). The detection of this region with a signal from the pol and env genes indicated that the VLPs were capable of recognising and packaging HIV-1 RNA. Furthermore if the  $\Psi$  sequence is the primary

determinant for RNA packaging in HIV-1, then it follows that genomic RNA had been packaged. The pol and env genes do not have their own packaging signals, so these sequences must have been detected on the same molecule that had the packaging signal i.e. genomic RNA. However it remained formally possible that the RNA was particle associated in some way, or was co-purified with the VLPs, and not packaged within a particle. This was addressed using detergent disruption of the VLPs.

#### **4.6.2. HIV-1 RNA was Packaged Within a Virus Like Particle.**

It was necessary to show that HIV-1 RNA was contained within a particle, and not just co-purified with the VLPs. To do this concentrated VLPs were centrifuged through a 20% sucrose gradient in the presence and absence of 0.1 % SDS detergent. After centrifugation, samples of the supernatant and the pellet fraction were removed and treated with RNazol.

The basis of the assay was disruption of the particle during centrifugation by the detergent SDS. With SDS, the VLP should be disrupted and any packaged RNA released into the supernatant. Without SDS the RNA should remain within the VLP and be pelleted with it during centrifugation. The results using the env primer pair PCR5/PCR20 are shown in figure 4.9.

In the presence of SDS, the RNA was detected in the supernatant fraction after centrifugation in lanes 7 and 8. No signal was detected with the pellet fraction in lanes 11 and 12. This was consistent with the detergent disrupting the VLP, causing the RNA to be released into the supernatant. The RNA could not be pelleted by these centrifugation conditions and so it was detected in the supernatant fraction only.

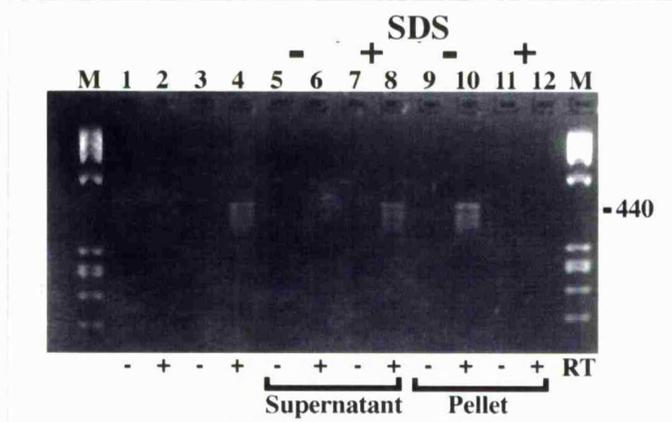


Figure 4.9 VLPs were disrupted by treatment with SDS. Even numbered lanes were +RT, odd -RT. RNA was detected in this batch of VLPs (lanes 3,4) prior to centrifugation on the sucrose gradient. For the supernatant (SN) fraction (lanes 5 to 8) in the absence of SDS (lanes 5 and 6) no RNA was detected after centrifugation. However with SDS there was a signal from the RNA (lanes 7 and 8) that had been released from the particles into the SN where it could not be pelleted. The pellet fractions (lanes 9 to 12) show that RNA was detected with the pellets (P) in the absence of SDS (lanes 9 and 10) and not when SDS was present (lanes 11 and 12). Water blank was lanes 1 and 2, M was HaeIII digested  $\phi$ X174 DNA marker. 3% Nusieve GTG gel.

In the absence of SDS the results showed that the RNA signal was only associated with the VLP pellet in lanes 9 and 10. This indicated that RNA was contained within the VLPs that had been pelleted. There was no signal present in the supernatant fraction after centrifugation, lanes 5 and 6. If the RNA was just co-purified with the VLPs then a signal would have been observed in the supernatant fraction (lanes 5 and 6) without SDS.

This was good evidence that the RNA was contained within a VLP and was not detected as a result of co-purification of the RNA and VLPs. Multiple bands were produced after RT-PCR in this assay using the primer set PCR5/20, the reason for this is unknown. However the results showed

that the VLPs were able to package RNA. The next step was to determine the specificity of RNA incorporation.

#### **4.6.3. Specificity of RNA Packaging into the VLP.**

In infected cells less than 1% of total cellular RNA is virion retroviral RNA, yet this is the predominant RNA found in retroviral particles (Linial and Miller, 1990). The recognition of the retroviral RNA packaging signal (*cis*) by the viral gag polyprotein (*trans*) is responsible for selective incorporation of genomic RNA. To determine whether the VLPs were able to specifically recognise and package HIV-1 RNA, compared to an abundant cellular message, RT-PCR was performed using primers to detect HIV-1 gag and human Actin RNA.

RNA was extracted from Centricon concentrated VLPs, and wild type HIV-1 virus. Total cellular RNA from transfected COS cells, and infected C8166 cells was obtained. RT-PCR was performed using HIV-1 gag and  $\beta$ -actin primers. The Act-1 (588-606) and Act-2 (1522-1504) oligonucleotides detect spliced actin mRNA and not genomic DNA, producing a 934 bp band. The actin gene is conserved, so the Act primers can detect simian actin (from COS cells) and human actin (from C8166 cells). The results are shown in figure 4.10.

A signal corresponding to actin mRNA was present for all of the cellular RNA samples; VLP producing cells (panel A, lane 3), mock transfected cells (panel A, lane 7) and for the infected C8166 cells (panel A, lane 14). However for both the VLPs and HIV-1 virus, actin RNA was not detected in the particles (panel A, lanes 1, 2, 12 and 13) indicating that actin RNA was not incorporated.

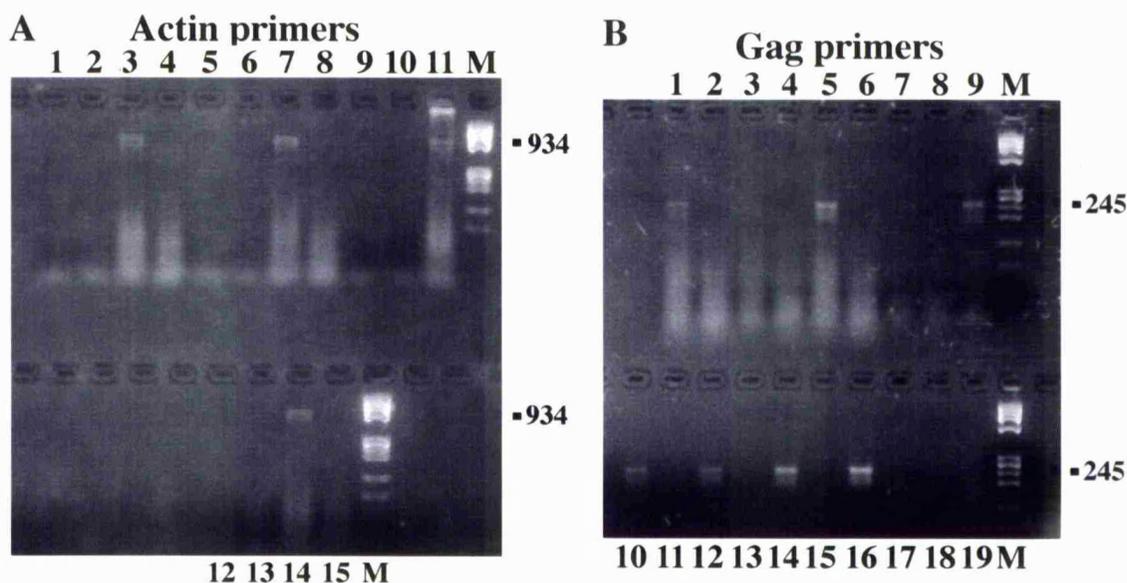


Figure 4.10. Specificity of RNA packaging in VLPs.

Panel A. Actin primers. PCR positive control (lane 11) and negative water blank (lanes 9 and 10). Lanes +RT are 1, 3, 5, 7, 9, 12 and 14; others are -RT. Actin RNA was detected in cellular RNA from mock transfected COS cells (lanes 7 and 8), pBCCX CSF X4 transfected COS cells (lanes 3 and 4), and HIV-1 infected C8166 cells (lanes 14 and 15), as expected. Actin RNA was not detected in the VLPs (lanes 1 and 2), supernatant from mock transfected cells (lanes 5 and 6) or NL4-3 virus (lanes 12 and 13).

Panel B. Gag primers. PCR positive control (lane 9) and negative water blank (lanes 7 and 8). Lanes +RT are 1, 3, 5, 7, 10, 12, 14, 16, and 18, others are -RT. HIV-1 gag RNA was present within the VLPs (lanes 14 to 17), and NL4-3 virus (lanes 10 to 13). Gag RNA was part of the cellular RNAs from VLP, lanes 5 and 6, and HIV producing cells, lanes 1 and 2. No HIV-1 gag RNA was detected in the mock transfected cellular RNA (lanes 3 and 4) or culture supernatant (lanes 18 and 19).

Therefore a specific incorporation of HIV-1 RNA had occurred into the VLPs. Actin RNA had been detected in the cells from which virus or VLPs had budded, but was absent from the viroplasm.

To ensure RNA was detectable within the VLPs and virus, RT-PCR was performed using the gag primers, panel B. Gag RNA was detected in total cellular RNAs, lane 1 for cells infected with NL4-3 virus, and lane 5 for

transfected and VLP producing cells. No gag RNA was detected in lane 3 for the mock transfected COS cells. For the duplicate virus and VLP samples in lanes 10, 12, 14 and 16 gag RNA was present as expected. No signal for the supernatant from the mock transfected cells, lanes 18 and 19 was observed.

These results indicated some degree of specificity in packaging HIV-1 RNA compared to spliced actin mRNA, an abundant cellular message. Identical results were obtained for HIV-1 virus and the VLPs. At the level of a 25 cycle RT-PCR, the signal produced from the HIV-1 RNA was much greater than that produced for actin RNA. The actin signal was not visible on ethidium bromide stained agarose gels when the HIV-1 signal was clearly present. On this basis, the VLPs specifically incorporated HIV-1 RNA.

To determine if genomic RNA was present within the VLPs as a monomer or dimer, Northern analysis was performed.

#### **4.7. Native Northern Analysis of VLP RNA.**

The RT-PCR analysis used here cannot determine if the VLP RNA was a monomer or a dimer. Therefore native Northern analysis was attempted to visualise the genomic VLP RNA. VLPs were centrifuged through a 20% sucrose TNE cushion and RNA was extracted by ProteinaseK/SDS digestion followed by phenol and phenol/chloroform extraction, this method is the best for virion RNA extraction (Clever et al., 1995). Proteinase K was selected in preference to RNazol used for the RT-PCR analysis, because RNazol may not inactivate all nucleases. To visualise genomic VLP RNA it was important that degradation of the RNA was

minimised. VLP RNA samples were analysed under native and denaturing conditions by electrophoresis on 0.8% agarose gels.

#### **4.7.1. VLP RNA was observed under Denaturing Conditions.**

VLPs were harvested and the RNA analysed under denaturing and non-denaturing conditions. Three bands were produced from the denatured VLP RNA samples, see figure 4.11.

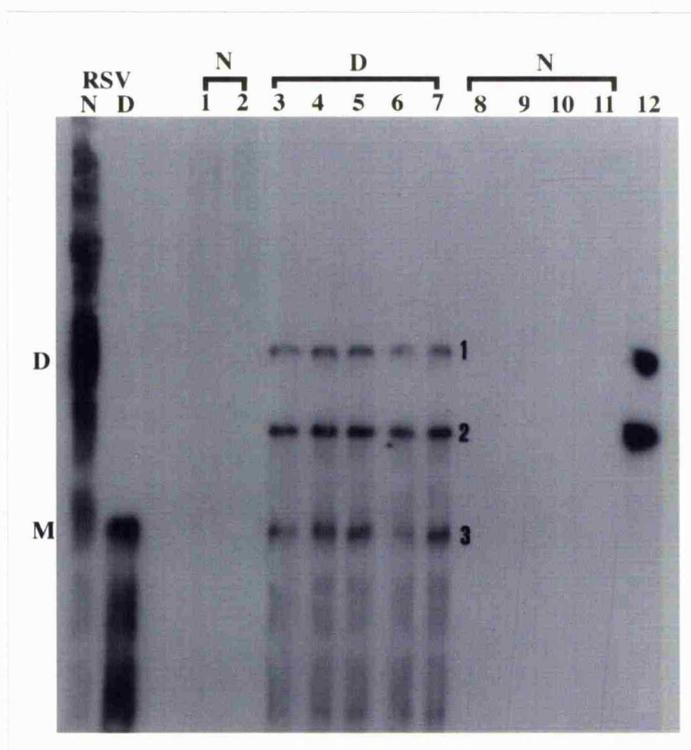


Figure 4.11. Northern analysis of VLP RNA. VLP RNA was analysed under native (lanes 1, 2, 8 to 11) and denaturing conditions (lanes 3 to 7). Lanes 3 to 7 had formamide dye added to 50% (v/v), and were heated at 90°C for 2 min (lanes 3 to 5), 100°C for 2 min (lane 6), or 65°C for 10 min (lane 7). Lane 12 was 500pg of pBCCX CSF X4 DNA. As a molecular weight marker, RSV genomic RNA was included under native (N) and denaturing (D) conditions.

Bands labelled 1, 2 and 3, were only present after the VLP RNA samples had been denatured, by the addition of formamide loading buffer and heating. Bands 1 and 2 had similar mobilities to the parental vector,

pBCCX CSF X4 DNA, and probably correspond to this plasmid carried over from the VLP harvest. Band 3 however could be the full length 'genomic' RNA from the VLP. It had a similar mobility to the denatured Rous sarcoma virus monomer RNA, used as a marker. Also like RSV, RNA smearing was seen migrating ahead of the VLP monomer band on denaturation. Dimeric VLP RNA would not be detected in the denatured samples. However no bands at all were observed in lanes 1, 2, and 8 to 11 when samples were analysed under non-denaturing conditions.

To observe any dimers VLP RNA must be analysed under non-denaturing conditions. It was unfortunate that the pBCCX CSF X4 DNA was running in a position where the VLP RNA dimers may be expected. Therefore it was necessary to identify conditions to remove the contaminating plasmid DNA and allow RNA dimers if present to be detected.

#### **4.7.2. VLP RNA Samples were treated with DNase.**

To identify the nature of the nucleic acid bands associated with the VLPs, samples were digested with enzymes prior to electrophoresis. Duplicate VLP RNA samples were treated with DNase, RNase, HindIII digested and denatured with 50% formamide, figure 4.12. After incubation samples containing enzymes were extracted once with phenol/chloroform and loading buffer added prior to electrophoresis.

The VLP bands 1, 2 and 3 were again present on denaturation, lanes 3 and 6. Lanes 1 and 4 showed that on DNase digestion bands 1 and 2 disappeared, suggesting that these were DNA. Also bands 1 and 2 had identical mobilities to the pBCCX CSF X4 plasmid DNA. Band 3, was resistant to DNase in lanes 1 and 4, suggesting that this could be RNA.

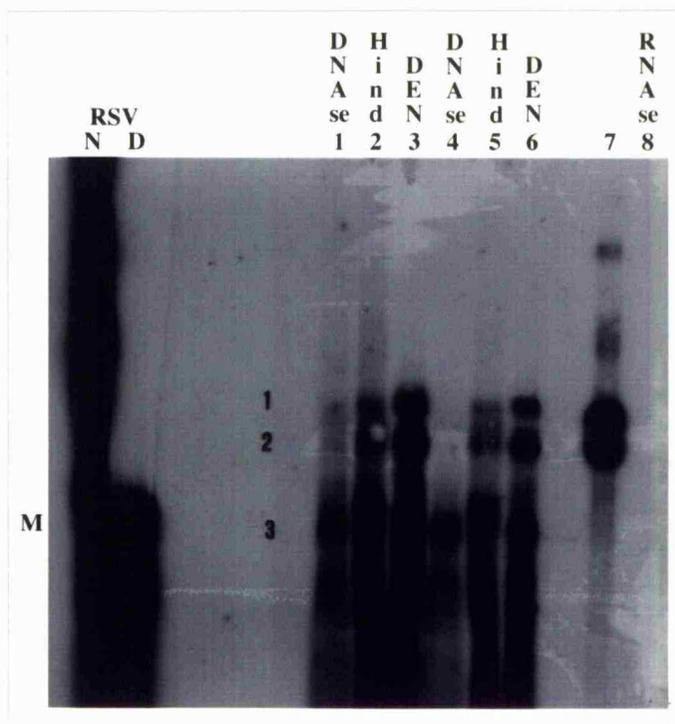


Figure 4.12. Nature of the VLP RNA. RNA samples (duplicates) were DNAsed (1U) at 37°C for 15 min, (lanes 1 and 4); HindIII digested (10U) at 37°C for 30 min (lanes 2 and 5) and RNAsed (0.5ug) at 37°C for 5 min (lane 8), all loaded under native conditions. Lanes 3 and 6 are denatured VLP RNA (50 % formamide, 90°C for 2 min) and lane 7, 500pg of pBCCX CSF X4 DNA.

It formed a discrete band similar in mobility to denatured RSV RNA, although these VLP samples had not been heat denatured. In the expected dimeric RNA position, there was no band present after DNase treatment suggesting that the VLPs did not contain dimeric RNA, or that it had been degraded. It was reasonable to suppose that the RNA dimer would have been detected if it was present, as monomer RNA was visible under these conditions.

In lanes 2 and 5, after Hind III digestion, there was a reduction in the signal for the DNA bands 1 and 2 but they were not completely removed. This was probably due to an inhibitor of digestion, or insufficient time for restriction of the plasmid. The candidate RNA band 3, was unaffected by Hind III treatment.

The absence of any bands after RNase digestion, lane 8, was misleading. It was probably due to an impure RNase, as plasmid DNA had been removed along with the monomer VLP RNA. A DNase free RNase was obtained and a similar experiment performed.

#### **4.7.3. VLP RNA samples were treated with RNase.**

DNase free RNase ONE (Promega) was used to treat the VLP samples before electrophoresis. Samples were also both DNase and RNase digested. The results are shown in the figure 4.13.

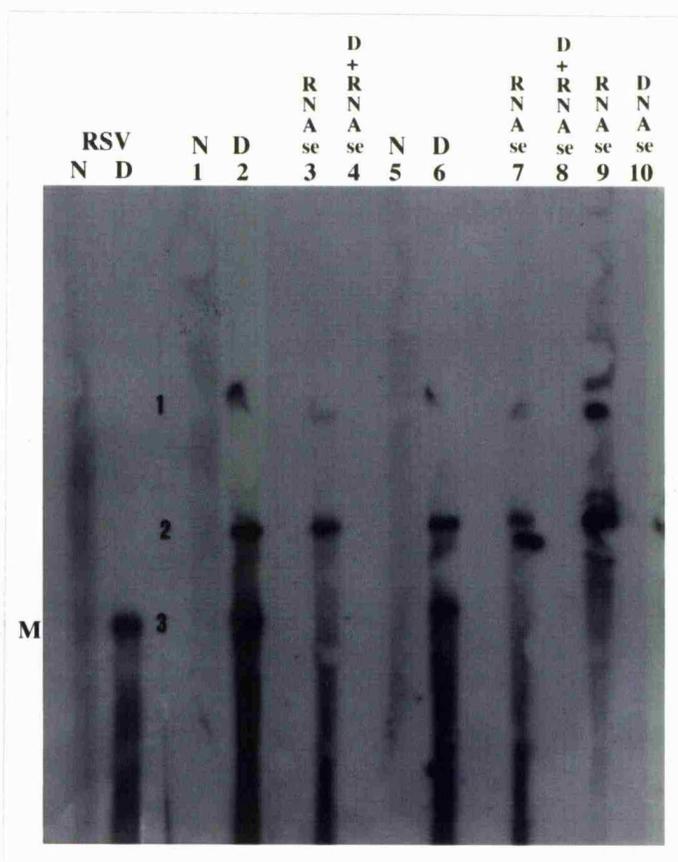


Figure 4.13. Nature of the VLP RNA. RSV RNA was used as a marker, native (N) and denatured (D). Lanes 1 and 5 were VLP RNA analysed under native conditions, lanes 2 and 6 denatured (50% (v/v) formamide, 90°C 2 min). Lanes 3 and 7 were RNase treated samples (0.5U RNase ONE, 37°C, 15 min). Lanes 4 and 8 were DNase and RNase treated (RNase ONE as lanes 3 and 7, then 1 U of RQ1 RNase free DNase, 37°C, 15 min). Lane 9 was 500 pg of pBCCX CSF X4 DNA, RNase treated, and lane 10 was this plasmid, DNase treated.

To confirm that the DNase and RNase enzymes were pure, they were used to digest pBCCX CSF X4 DNA. Lane 9 was the plasmid after RNase treatment and in lane 10 after DNase digestion. As expected the plasmid DNA signal remained on RNase treatment (lane 9), but was absent after DNase digestion (lane 10).

Digestion with RNase has specifically eliminated band 3 in lanes 3 and 7, however the DNA bands 1 and 2 have remained. This was consistent with band 3 corresponding to VLP 'genomic' RNA. A combination of DNase and RNase has eliminated all signal in lanes 4 and 8, as expected.

For the VLP samples under non-denaturing conditions lanes 1 and 5, no discrete bands were seen, only smearing back to the wells. The denatured VLP samples again gave rise to bands 1, 2 and 3, in lanes 2 and 6.

This result suggested that VLP 'genomic' RNA was present in the VLPs, but also showed the same contaminating DNA problem as encountered before. There was no RNase sensitive band where the dimeric VLP genomic RNA would be expected. The conclusion must be that the VLPs did not contain a genomic RNA dimer.

#### **4.8. Summary of VLP Analysis.**

Rising p24 viral antigen and reverse transcriptase, electronmicrographs, and the peak of p24 protein on sucrose gradients, indicated that the pBCCX CSF X4 construct produced virus like particles (VLPs) consisting of assembled p24 molecules. The RNA content of the VLPs was investigated using RT-PCR and Northern blotting. RT-PCR detected HIV RNA sequences corresponding to the gag, pol and env genes, but not cellular actin mRNA, suggesting that full length genomic VLP RNA was present. This also suggested that there was some specificity for viral RNA incorporation. The RNA was released from the VLPs by detergent

treatment, so it was completely packaged within a particle. Northern analysis showed that monomeric RNA was probably present in the VLPs, but dimeric RNA was never observed. Contaminating plasmid DNA from the electroporation further complicated the Northern analysis. Conditions were never established in which the plasmid could be removed, or dimeric HIV RNA could be identified under non-denaturing conditions, despite many attempts.

The inability of the VLP RNA to dimerise may have been due to the presence of the vector encoded 84 nucleotide RNA leader. In the VLP system dimerisation did not appear to be a prerequisite for RNA packaging. Therefore monomer RNA may be recognised and incorporated into a VLP or virus which may then subsequently dimerise.

The lack of dimeric RNA was probably due to its absence, as the experimental conditions allowed the RNase sensitive VLP monomer band to be detected. Denaturation of the VLP samples consistently produced better signals and discrete bands. Under non-denaturing conditions, either no signal or high molecular weight smearing was present despite variation of the extraction procedures. Perhaps the RNA was inaccessible, complexed to protein and lipids which were only removed by denaturing with formamide and heating prior to electrophoresis. RNA dimerisation may be associated with particle maturation (Fu and Rein, 1993 and Fu et al., 1994). If the VLP particles were immature, then this may explain why RNA dimerisation has not occurred. The inability to easily detect RNA dimers, or their absence, indicated that the VLP system was unsuitable for mutational analysis of RNA dimerisation.

#### **4.9. Testing *in vitro* models of HIV-1 RNA Dimerisation Directly in Mutated HIV-1 Viruses.**

One of the original reasons for using VLPs was that mutations which disrupted dimerisation or packaging, may also prevent viral replication. So a system where viral assembly was decoupled from replication would allow potentially lethal mutations to be introduced and examined. *In vivo* it is unknown if dimerisation is essential for viral replication, packaging and assembly. From the VLP analysis RNA packaging and assembly of the particles did not appear to require dimeric RNA. The possibility also remained that the mutations which disrupted dimerisation *in vitro* may have absolutely nothing to do with virion RNA dimerisation. For these reasons mutagenesis was performed directly onto wild type HIV-1<sub>NL4-3</sub> viruses via the plasmid pUCANL.

#### **4.10. Mutagenesis of Wild Type HIV-1<sub>NL4-3</sub>.**

From the *in vitro* RNA dimerisation studies, two mutations were identified which prevented RNA dimerisation of small transcripts from the DLS and DIS of HIV-1. The G<sub>819</sub> to U change abolished tetrad mediated dimerisation, and the <sub>711</sub>GCG to AAA mutation disrupted the kissing loop structure. To determine whether these mutations were involved in virion RNA dimerisation, they were introduced singly and collectively into an infectious molecular clone of HIV-1.

##### **4.10.1. The HIV-1<sub>NL43</sub> infectious molecular clone, pUCANL**

pUCANL plasmid is a circularly permuted clone of HIV-1<sub>NL4-3</sub>, as described in materials and methods (Section 2.6.3, figure 2.6). EcoRI digestion, ligation, and electroporation of the DNA concatemer into T-cells produced wild type HIV-1<sub>NL4-3</sub> viruses (M.Johnson and A.J. Cann, unpublished observations). Restriction digestion of the plasmid was

performed to ensure that the SpeI and BssHIII sites were unique and therefore suitable for LP-USE mutagenesis. The restriction digestion analysis of the plasmid is shown in figure 4.14.

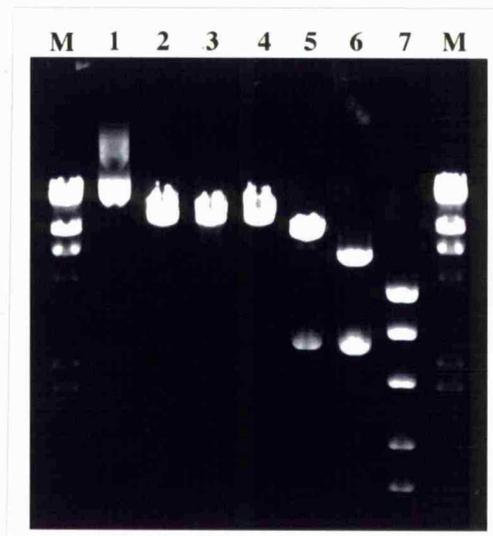
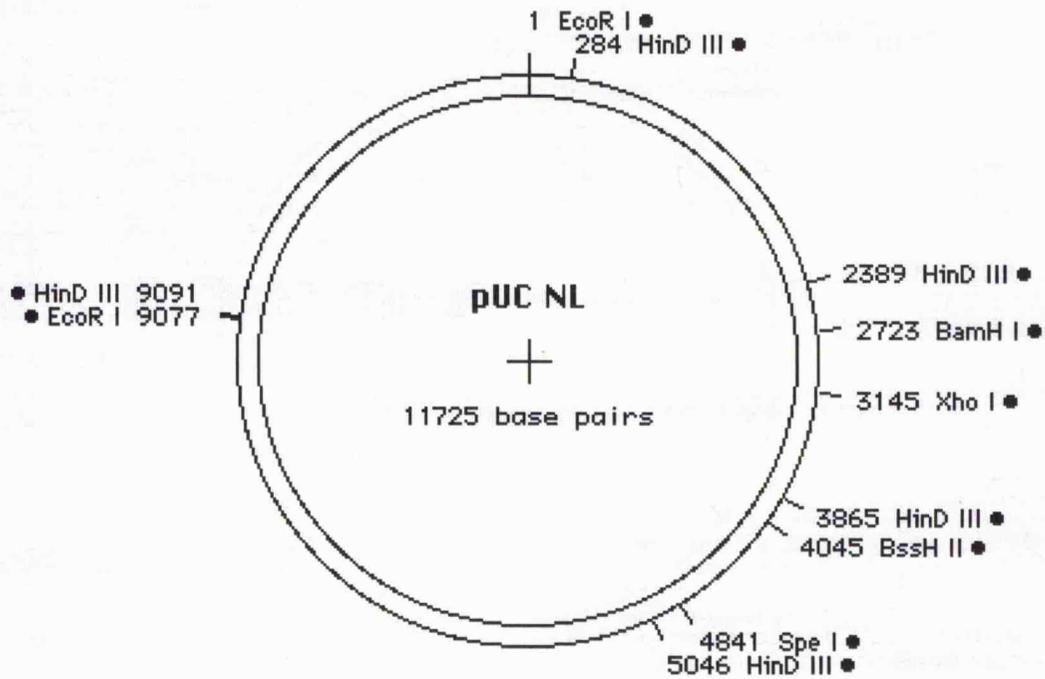


Figure 4.14. Restriction digestion of pUC $\Delta$ NL. 0.5 $\mu$ g of this plasmid in lane 1, was digested with SpeI (lane 2), Xho I (lane 3), BssHIII (lane 4), EcoRI (lane 5), BamHI and EcoRI (lane 6) and HindIII (lane 7). After heat inactivation of the enzymes, the samples were electrophoresed on a 0.8% agarose TBE gel alongside  $\lambda$  HindIII DNA markers (lane M).

The plasmid was linearised by digestion with SpeI, BssHIII and Xho I, at these unique sites (lanes 2, 3 and 4). EcoRI digestion released the HIV fragment from the pUC18 vector, producing bands of 9,076 and 2,649 bp respectively (lane 5). BamHI and EcoRI digestion produced the 6,354; 2,722 and 2,649 fragments (lane 6). HindIII digestion was performed in lane 7, producing the expected 4,045; 2,918; 2,105; 1,476 and 1,181 bp fragments. This analysis showed that the construct was as expected. No rearrangement of the vector had occurred, often a problem for this type of plasmid (Yamada et al. 1995, Cann 1990).

#### **4.10.2. Mutagenesis of pUCANL**

The dimerisation mutations defined *in vitro* were introduced into pUCANL by the LP-USE procedure as described in materials and methods. The construction of the mutants is summarised in figure 4.15.

### HIV Mutant Plasmids

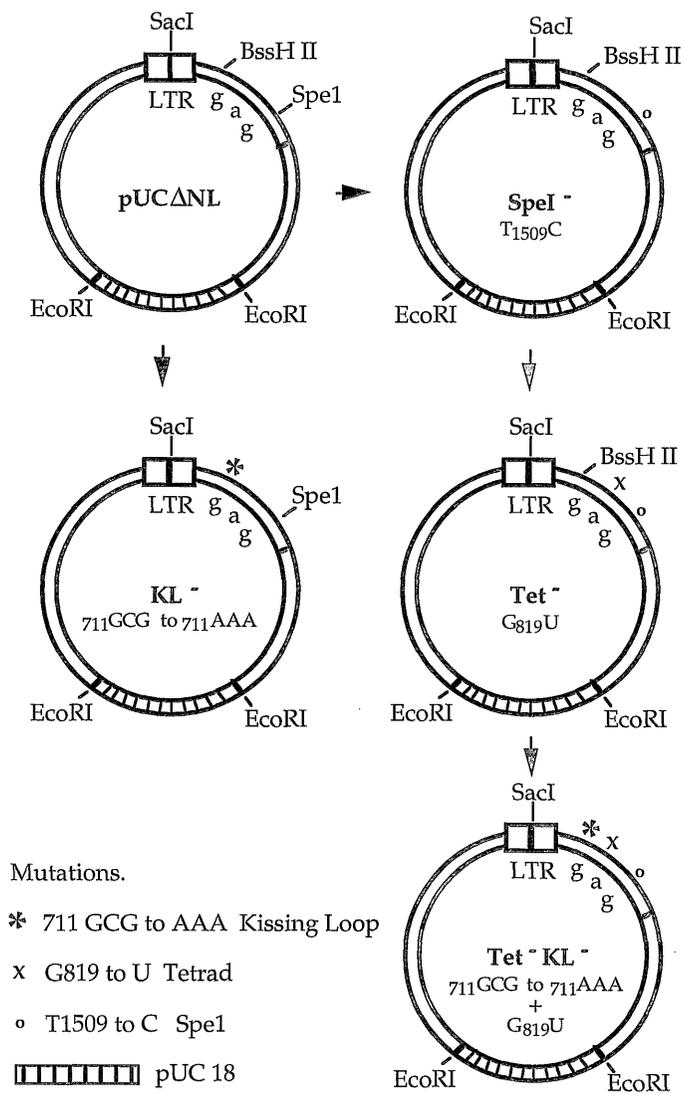


Figure 4.15. Mutagenesis summary of the DNA clones producing the mutant HIV viruses, SpeI<sup>-</sup>, Tet<sup>-</sup>, KL<sup>-</sup>, Tet-KL<sup>-</sup> and wild type NL43.

The plasmids constructed were:

pMH 95.010 and 95.011. (HIV Tet<sup>-</sup>)

Primers MH22 and MH23 and pUCΔNL plasmid template generated a PCR product of 717bp containing the G<sub>819</sub> to U and the T<sub>1509</sub> to C change. This was used as the primer for LP-USE mutagenesis of pUCΔNL. Mutant plasmids were selected for by digestion with SpeI. The G<sub>819</sub> to U mutation does not change gag coding, but it does alter the RNA sequence. Viruses produced are designated HIV Tet<sup>-</sup>.

pMH 95.016 and 95.019. (HIV KL<sup>-</sup>)

MH24 was annealed to pUCΔNL, extended and the LP-USE mutagenesis procedure followed. This introduced the 7<sub>11</sub>GCG to 7<sub>11</sub>AAA kissing loop mutation which also disrupted the BssHII site. This enzyme was used for selection of the mutant plasmids. The mutated sequence is part of non coding RNA, within the DIS. Viruses produced are designated HIV KL<sup>-</sup>.

pMH 95.027 (HIV Tet-KL<sup>-</sup>)

The double mutant was obtained by annealing the PCR oligo MH24 onto pMH95.010 (HIV Tet<sup>-</sup>), following the mutagenesis procedure and selecting with BssHII. Viruses produced are HIV Tet<sup>-</sup>KL<sup>-</sup>.

pMH 95.021 (HIV SpeI<sup>-</sup>)

This clone was a necessary control, to ensure that the single T<sub>1509</sub> to C SpeI mutation, introduced as a consequence of the mutagenesis, did not affect HIV RNA dimerisation. This mutation is within gag coding but it was a non-coding change. pMH95.021 was obtained during screening for HIV Tet<sup>-</sup> mutants. pMH 95.021 did not have the G<sub>819</sub> to U change, only the T<sub>1509</sub> to C SpeI<sup>-</sup> mutation. Viruses produced are HIV SpeI<sup>-</sup>.

To identify mutants, plasmid miniprep DNA was recovered from colonies after the USE mutagenesis. Sequencing and restriction digestion was performed to identify mutants. Maxiprep DNA was obtained, further

restriction digestions and sequencing of the entire amplified fragment was performed to produce the plasmids described above.

As an example, for clones pMH95.016 and pMH95.019 miniprep DNA was prepared from eight colonies after mutagenesis. The DNA was digested with BssHIII and XhoI. For the wild type vector, a 900bp band results (arrowed). This was absent from the mutants because the BssHIII site has been removed by the kissing loop mutation. Two clones (lanes 3 and 8) were candidate mutant plasmids, shown in figure 4.16, panel A.

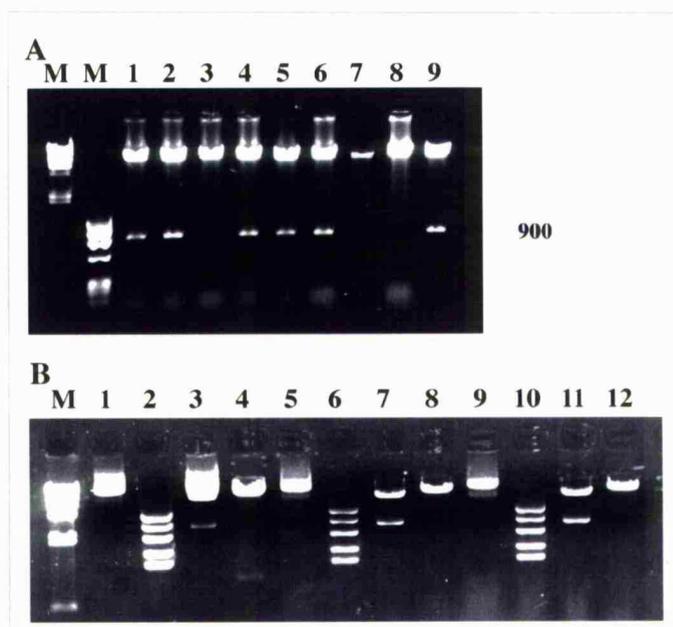


Figure 4.16. Production of clones pMH95.016 and 019.

Panel A, eight miniprep DNAs were digested with BssHIII and XhoI. A 900 bp band was present in the wild type vector (lane 9), and in lanes 1, 2, 4, 5, 6, and 7. This was absent in lanes 3 and 8 due to the kissing loop mutation. M was  $\lambda$  HindIII DNA, and  $\phi$ X174 HaeIII DNA markers.

Panel B, maxiprep DNA for the two kissing loop mutants (lanes 5 to 12) was digested and compared with the parental pUC $\Delta$ NL vector (lanes 1 to 4). Uncut plasmid lanes 1, 5 and 9. Hind III digested lanes 2, 6 and 10, EcoRI lanes 3, 7 and 11, and BssHIII with XhoI, lanes 4, 8 and 12. The absence of the 900bp band in lanes 8 and 12, present in lane 4, indicates that these clones had the kissing loop mutation. Lane M was a  $\lambda$  DNA HindIII digest.

Maxiprep DNA was prepared for these clones and designated pMH95.016 and pMH95.019. Further restriction digestion was performed with HindIII, EcoRI, and a double BssHII and XhoI digestion. As shown, in figure 4.16 panel B, the 900 bp band was absent from the mutants, and the other digestion profiles were consistent with the wild type vector. Sequencing throughout the mutated region confirmed the expected mutations and the absence of any others. Plasmids pMH95.016 and 95.019 produced HIV KL-viruses.

#### **4.11. Electroporation of Mutant Virus DNA Clones.**

25µg of each plasmid DNA was digested with EcoRI and ligated overnight at 4°C as described in materials and methods. To produce mutant viruses the resulting DNA concatemer was introduced into  $5 \times 10^6$  C8166 cells by electroporation. Cells were maintained at  $0.5 \times 10^6$  per ml and inspected daily. Culture supernatant was removed from the cells aliquotted and stored either in liquid nitrogen for viral stocks, or inactivated for p24 assay.

For the electroporation of clones pMH 95.010, 011(HIV Tet-), 016, 019 (HIV KL-), 021(HIV SpeI-), and 027(HIV Tet-KL-), an increase in p24 production was observed, as shown in figure 4.17.

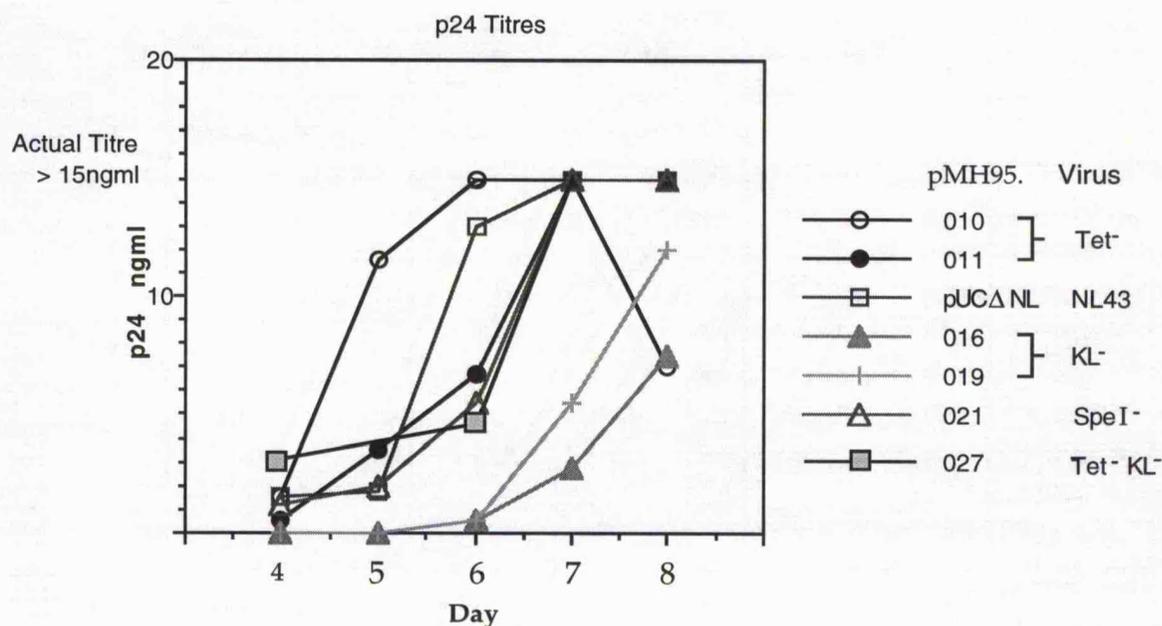


Figure 4.17. p24 assay from electroporated C8166 cells. p24 antigen titre was established for each virus, resulting from electroporation of the appropriate pMH95. plasmid DNA.

There was an increase in p24 produced from the electroporated C8166 cells, along with the development of syncytium in the cultures for all of the viral clones. All produced p24 protein in the culture supernatant and therefore virus. Non of the introduced mutations were lethal, viruses were able to replicate and spread the infection. Therefore an analysis of the virion RNA in the mutant viruses can be performed, the null phenotype which could have resulted from these mutations has not occurred.

#### **4.12. Infection of C8166 T Cells Using the Mutant Viruses.**

Infections using the mutant viruses were performed to allow a comparison of growth rates and to isolate virion RNA. This avoids contaminating DNA, which was a problem during the VLP Northern analysis from electroporated cells. Also higher titres of virus were generated to produce large amounts of virion RNA.

Using virus stocks from the electroporation (figure 4.17), infections were normalised to an infectious dose of 5.0ng of p24. Virus from clones pMH 95.010 (HIV Tet<sup>-</sup>), pMH 95.019 (HIV KL<sup>-</sup>), pMH 95.021 (HIV SpeI<sup>-</sup>), pMH 95.027 (HIV Tet<sup>-</sup>KL<sup>-</sup>) and pUCΔNL (HIV NL4-3) was added to 5ml of fresh C8166 cells (5x10<sup>6</sup> cells), and allowed to infect overnight. The infected cells were resuspended in fresh medium, supplemented with 1x10<sup>6</sup> uninfected cells and maintained during the infection. Samples of supernatant were removed for p24 analysis and further virus stocks. The p24 profile for each virus infection is shown in figure 4.18.

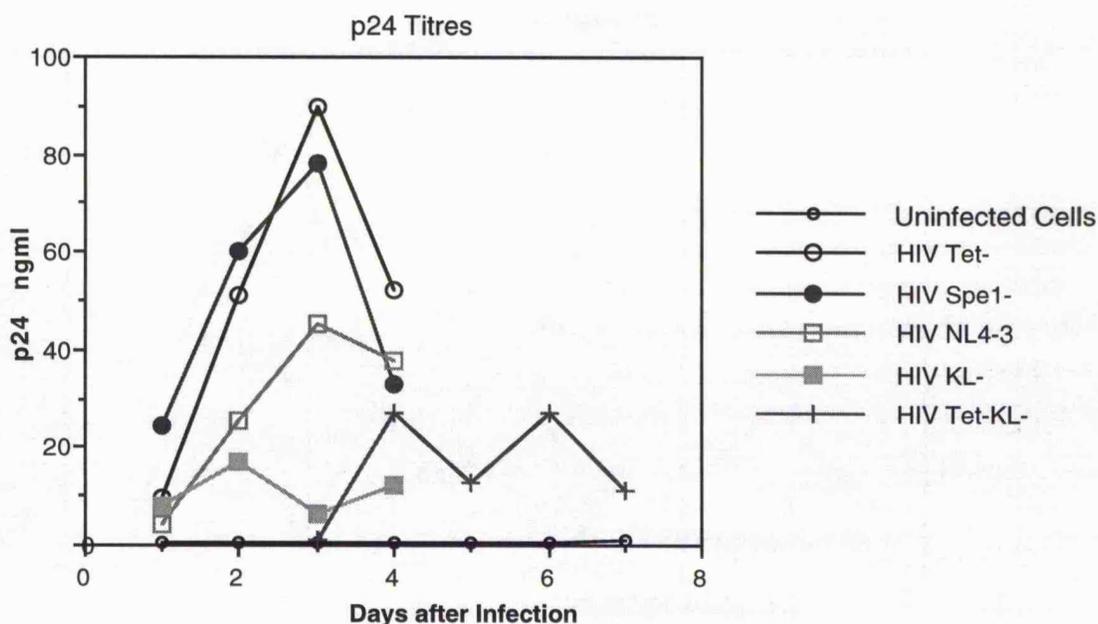


Figure 4.18. A representative p24 profile for the infection of each virus. Culture supernatant samples were removed from infected C8166 cells with the HIV viruses and p24 antigen levels determined: Uninfected cells ○, HIV<sub>NL4-3</sub> □, HIV *SpeI*<sup>-</sup> ●, HIV Tet<sup>-</sup> ○, HIV KL<sup>-</sup> ■ and HIV Tet<sup>-</sup>KL<sup>-</sup> +.

The growth curves indicated that the wild type NL4-3 virus, the tetrad mutant HIV Tet<sup>-</sup>, and the SpeI<sup>-</sup> virus had similar growth rates. However, the viruses that contained the kissing loop mutation, HIV KL<sup>-</sup> and HIV

Tet-KL-, appeared to grow more slowly. These growth characteristics were also reflected in syncytium formation and progression of the infection in the cell culture.

The reduction in the growth rate for the kissing loop mutants was consistently seen on three separate occasions. However a more accurate assay would be required to determine the precise growth rate of each virus.

#### **4.13. Confirmation of Viral Sequences ex Culture.**

The identity of the mutant viruses isolated from cell culture was confirmed by restriction digestion and sequencing analysis to ensure no cross culture contamination had occurred.

##### **4.13.1. RT-PCR Analysis of Mutant Viruses.**

RT-PCR was performed on the viruses to confirm that they packaged RNA, and to determine the sequence of each virus. Primers SH29 and SH175 were used in the RT-PCR reactions. These amplified NL4-3 DNA from position 453 to 858 containing the tetrad and kissing loop sequences, see figure 4.19.

A band of 424 bp was expected from this primer pair. The absence of any signal in the water blank (lanes 2 and 3) and those samples to which reverse transcriptase was omitted (even numbered lanes), indicated that it was RNA that had been amplified. All of the mutant viruses contained RNA, see lanes 5, 7, 9, 11 and 13.

This demonstrated that none of the mutations that abolished dimerisation *in vitro*, had prevented RNA packaging in the mutated viruses *in vivo*.

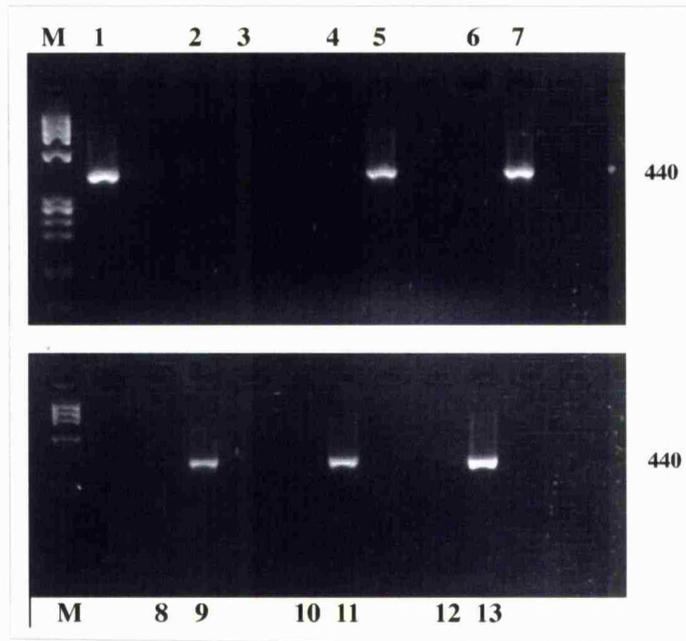


Figure 4.19. RT-PCR analysis of RNA extracted from HIV NL43, SpeI-, Tet-, KL-, and Tet-KL- viruses. Odd numbered lanes were +RT samples, even were -RT. Lane 1 was the PCR positive control, lanes 2 and 3 the water blank. HIV RNA was detected in lanes 5 (NL4-3), 7 (SpeI-), 9 (Tet-), 11 (KL-) and 13 (Tet-KL-) viruses.

#### **4.13.2. Restriction Digestion analysis of the RT-PCR product.**

The SH29/SH175 RT-PCR product was digested with SacI and HindIII to allow cloning of the DNA prior to sequencing, figure 4.20. BssHIII digestion was also performed on the SacI/HindIII RT-PCR fragment from the viruses, to identify the presence or absence of the kissing loop mutation.

RT-PCR product from viral RNA

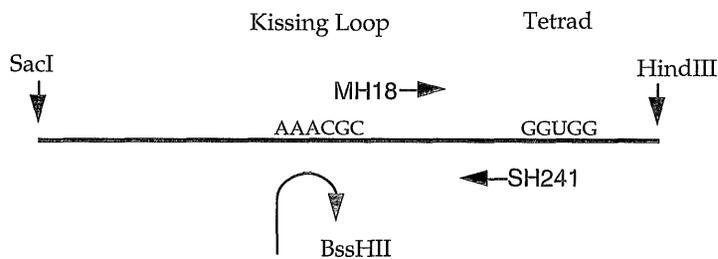


Figure 4.20. Schematic representation of the *SacI*/*HindIII* digested RT-PCR product amplified from viral RNA. The location of the tetrad and kissing loop mutations are indicated along with the sequencing primers MH18 (for G<sub>819</sub> to U) and SH241 (for 7<sub>11</sub>AAA) used to confirm the sequence of each virus.

The kissing loop mutation removes the *BssHII* site, making the fragment resistant to digestion. RT-PCR products from each mutant viruses were digested with *BssHII* and electrophoresed on a 3% GTG agarose gel, figure 4.21.

The Tet<sup>-</sup> and SpeI<sup>-</sup> virus RT-PCR products lane 2 panel A and lane 1 panel B respectively, were digested by *BssHII* producing bands of 180 and 156 base pairs. This indicated that the kissing loop mutation was absent from these viruses, as expected. The KL<sup>-</sup> and Tet-KL<sup>-</sup> RT-PCR products were not digested, lanes 1 and 3, panel A. This showed that the *BssHII* site had been changed and that the identity of the virus from these cultures was as expected. The DNA from each virus behaved as a single species, with no evidence of cross contamination observable at the level of ethidium bromide staining of agarose gels.

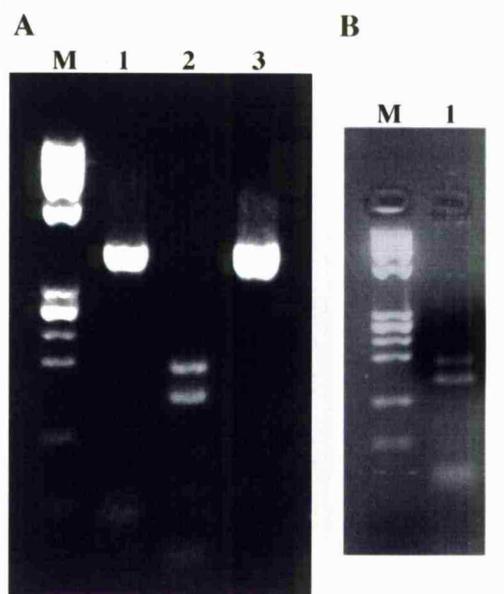


Figure 4.21. BssHIII digestion was performed on the RT-PCR product from mutant Tet<sup>-</sup>, KL<sup>-</sup>, Tet-KL<sup>-</sup> and SpeI<sup>-</sup> viruses. Panel A, the KL<sup>-</sup> and Tet-KL<sup>-</sup> DNAs are not cleaved by BssHIII (lanes 1 and 3). Panel B, the Tet<sup>-</sup> (lane 2) and SpeI<sup>-</sup> DNAs (lane 1 panel B) are cut by this enzyme into 180 and 156 bp bands.

#### **4.13.3. Sequencing of the Mutant Virus RT-PCR Products.**

The BssHIII digestion indicated that no cross contamination of viruses had occurred. To confirm the identity of the viruses recovered from cell culture, sequencing of the RT-PCR product from each virus was performed, see figure 4.22.

#### **Tetrad Mutation, G819 to U.**

For the NL4-3, SpeI<sup>-</sup>, and Tet<sup>-</sup> viruses, the SH29/SH175 RT-PCR product was sequenced directly using primer MH18. This confirmed the G<sub>819</sub> to U change for HIV Tet<sup>-</sup> (panel A), and the wild type G<sub>819</sub> sequence for SpeI<sup>-</sup> and NL4-3 viruses (panels B and C respectively).

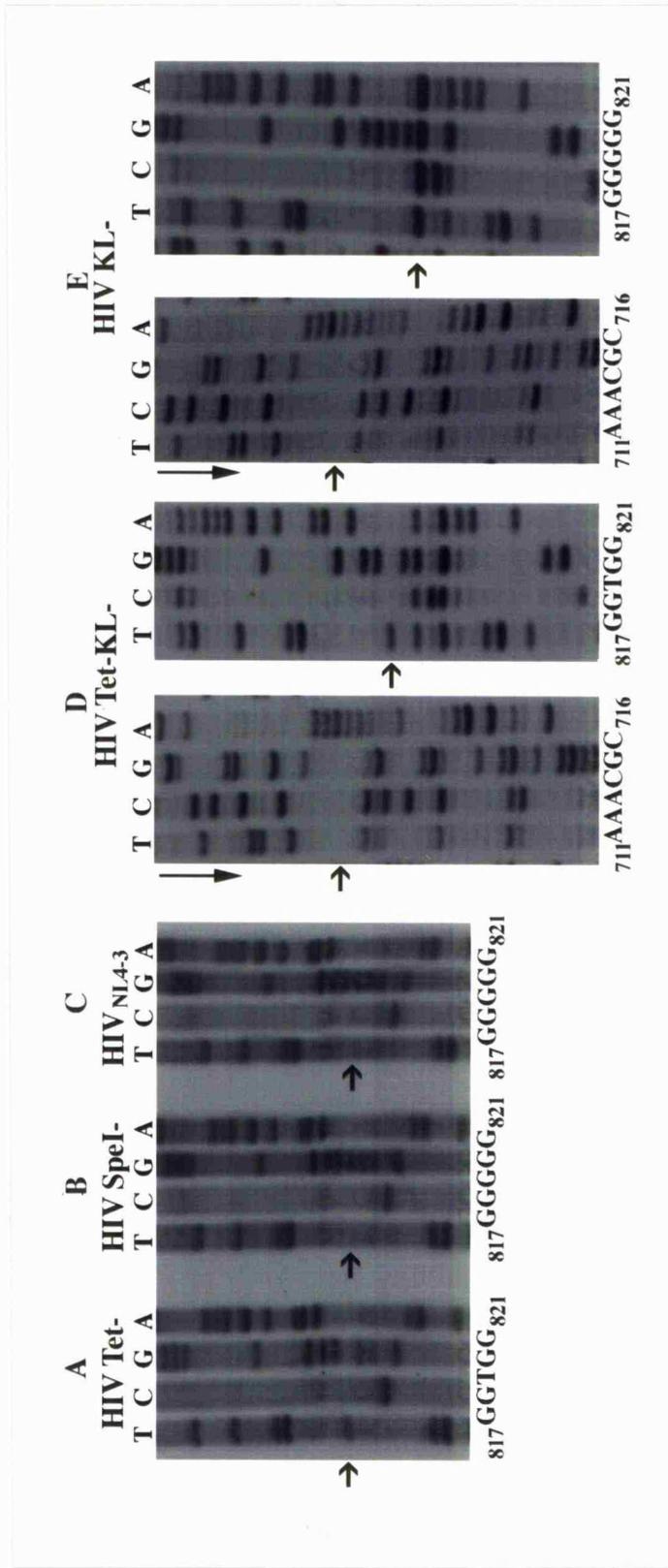


Figure 4.22. Sequencing of recovered mutant viruses. Panel A; The G819 to T mutation was present for the HIV Tet- virus. The wild type G819 sequence was present for the HIV SpeI- virus (Panel B) and wild type HIV-1 NL4-3 virus (Panel C). The double mutant virus HIV Tet-KL- contains the G819 to T change and the 711GCG to 711AAA mutation (Panel D). The HIV KL- virus contained the 711GCG to 711AAA change and the expected G819 wild type sequence. The positions of the mutations or wild type sequences are indicated with an arrow. Sequences are read from bottom to top except for the 711AAA mutations in panels D and E, indicated by the downward pointing arrow.

#### Kissing Loop Mutation, 711GCG to AAA.

To confirm the mutations in the HIV KL<sup>-</sup> and Tet-KL<sup>-</sup> viruses, the SacI/Hind III digested RT-PCR product was cloned into SacI/Hind III cut pUC18. Miniprep DNA was identified from positive clones and sequenced. The presence of the 711AAA change was confirmed in HIV Tet-KL<sup>-</sup> (panel D) and HIV KL<sup>-</sup> (panel E) using sequencing primer SH241. With primer MH18 the G<sub>819</sub> to U mutation was confirmed in the double mutant HIV Tet-KL<sup>-</sup> virus (panel D), and the wild type G<sub>819</sub> was present in this position for HIV KL<sup>-</sup> (panel E).

The mutations were as expected for all of the viruses. This confirmed that no cross contamination of viruses had occurred, and showed that the mutations had not changed, or reverted. Any possibility of a second site mutation beyond the tetrad or kissing loop regions were not explored, as this would entail sequencing the entire genome.

Having confirmed the identity of each virus the effect of the mutations on RNA packaging and dimerisation was analysed by native agarose gel electrophoresis.

#### **4.14. Virion RNA Analysis.**

Northern analysis was performed to determine the effects of the mutations on RNA dimerisation. For virion RNA analysis viruses were harvested from overnight cultures, after the first signs of syncytium formation were observed at about day 2 to 3. Typically 30ml of supernatant was taken from each culture, viruses were harvested by centrifugation and virion RNA was extracted as described in materials and methods.

#### **4.14.1. Native Agarose Gel Electrophoresis of Mutant Viral RNA.**

Both native and denatured viral RNA samples were electrophoresed on 0.8 % agarose 0.5X TBE gels, electroblotted, and probed as described in materials and methods. This analysis was performed many times with essentially similar results, representative samples are shown in figure 4.23(A).

Rous sarcoma virus RNA was electrophoresed alongside the HIV-1 samples, as the Rous dimer RNA has been extensively studied by Northern analysis (Mangel et al., 1974, Meric et al., 1984, and Oertle and Spahr, 1990) and so was a good reference. In the native (N) lanes, there was a band which corresponds to the 70S dimeric RNA, and bands of higher molecular weight (Bader and Ray, 1976). On denaturation (D) there was a band of higher mobility corresponding to the RSV monomer RNA, lane 2. HIV RNA isolated from wild type NL4-3 virions gave a similar result, lanes 3 and 4, with monomer and dimer RNA mobilities similar to RSV.

For the mutant viruses the HIV SpeI<sup>-</sup> virus which contains only the U1509 to C change, also contained monomer and dimer genomic RNA in lanes 5 and 6. The mutation has not affected RNA dimerisation. This was expected as the U1509 to C was a non-coding change within the gag gene, unlikely to be involved in RNA dimerisation. The presence of this mutation in HIV Tet<sup>-</sup> and HIV Tet<sup>-</sup>KL<sup>-</sup> viruses as a consequence of the mutagenesis can therefore be ignored.

Lanes 7 to 12 show that HIV Tet<sup>-</sup>, HIV KL<sup>-</sup> and HIV Tet<sup>-</sup>KL<sup>-</sup> were able to package full length monomeric RNA seen on denaturation. These mutations therefore do not prevent RNA packaging, confirming the RT-PCR results. The native RNA runs more amorphy than the heat denatured RNAs and with a higher molecular weight. There was a band

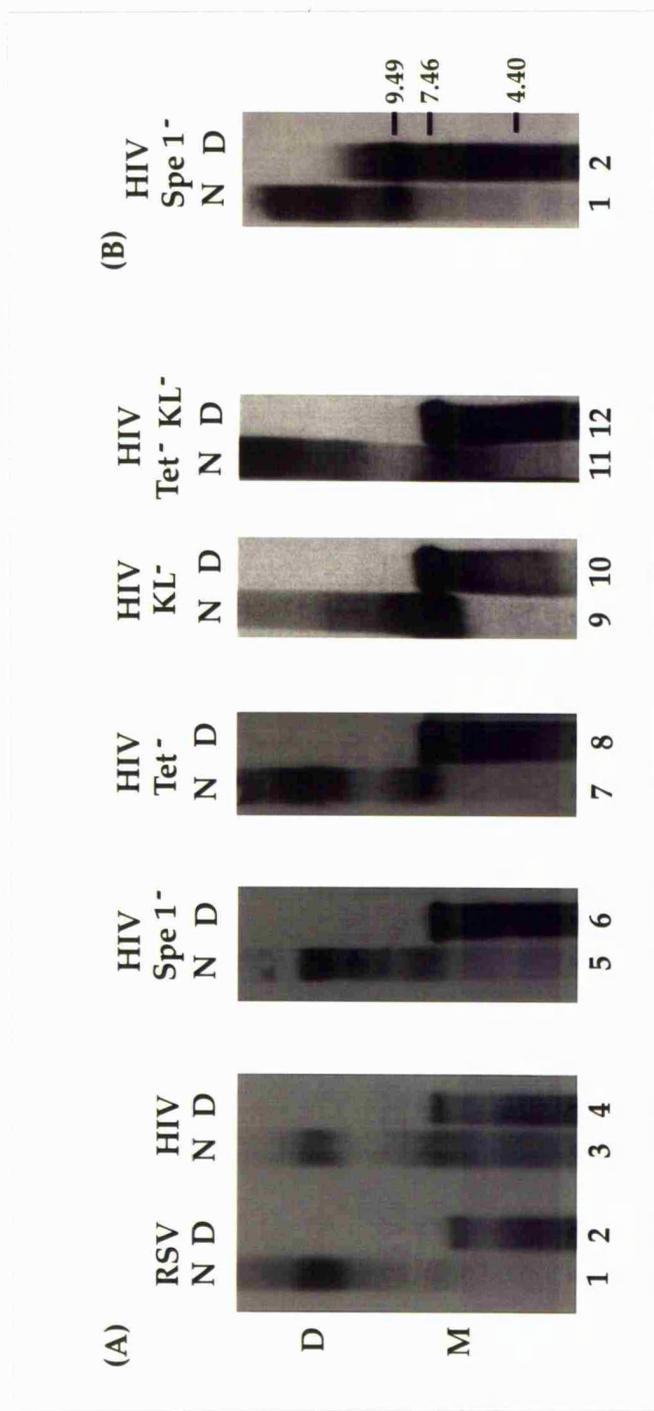


Figure 4.23(A). Northern Analysis of virion RNA. Viral RNA was analysed by non denaturing TBE agarose gel electrophoresis, transfer to a nylon membrane then hybridisation with a 32P labelled probe as described in materials and methods. Rous sarcoma Virus (RSV) RNA is shown in lanes 1 and 2, HIV-1 in lanes 3 and 4. Lanes 5 to 12 show HIV-1 RNA containing the mutations described in figure 4.15. All of the virus samples show a band with a mobility of 9-10 kB band under denaturing (D) conditions lanes 2, 4, 6, 8, 10, and 12. Native (N) RNA samples have a reduced mobility. Dimeric RNA bands can clearly be seen in lanes 1, 3, 5, and 7. In lanes 9 and 11 the dimeric band is present but less apparent. Figure 4.23(B). Analysis of virion RNA without any organic extraction procedures. HIV SpeI- RNA was isolated without phenol extraction or ethanol precipitation and electrophoresed as figure 4.23(A). Monomeric RNA was present in lane 2 after denaturation (D), higher molecular weight RNA species were present in lane 1 under native (N) conditions.

coincident with the Rous 70S dimer RNA, for all mutant viruses, see lanes 1, 3, 5, 7, 9 and 11, indicating that some of the packaged RNA was dimeric. The Tet<sup>-</sup> virus appeared to contain a normal RNA dimer. However the viruses which contained the 711 GCG to 711 AAA kissing loop mutation were different. A poorly defined dimer band in the KL<sup>-</sup> virus, lane 9, and in Tet<sup>-</sup>KL<sup>-</sup> virus, lane 11, was consistently present. These viruses contained an increased amount of monomeric RNA, and a reduced amount of dimeric RNA compared to other KL<sup>+</sup> viruses. This suggested that the dimers formed from KL<sup>-</sup> RNA were different.

Additionally RNA of a molecular weight greater than a dimer for SpeI<sup>-</sup>, Tet<sup>-</sup>, and Tet-KL<sup>-</sup> viruses was present. A greater proportion this RNA was present in the double mutant virus. The nature of this species was unknown, although it has been noticed in other viral RNA extractions (Bader and Ray, 1976). Higher ordered genomic RNA, or a protein:RNA complex may explain this banding.

To interpret the effects of the mutations on RNA dimerisation, it was necessary to first discover if phenol extraction had any effect on virion RNA dimerisation. Phenol extraction was previously shown to assist dimer formation of the kissing loop RNAs *in vitro*.

#### **4.14.2. The Effect of Organic Extraction on virion RNA dimers.**

All of the viral RNA samples were phenol, and phenol chloroform extracted then precipitated with ethanol prior to electrophoresis. In order to exclude the possibility that these procedures might contribute to RNA oligomerisation, HIV SpeI<sup>-</sup> virion RNA was electrophoresed following Proteinase K treatment only of these viruses. No organic extraction or ethanol precipitation was performed on these samples, see figure 4.23(B).

Monomeric RNA was visible following denaturation lane 2. Higher molecular weight RNA was present in the non-denatured sample, lane 1,

presumably corresponding to the RNA dimer. This demonstrated that retroviral RNAs are multimeric when extracted from the virus particle, and not due to subsequent phenol extraction or ethanol precipitation. Phenol has been shown to have no obvious effect on the dimerisation of virion RNA. After dissociation at 85°C, no treatment related to those used for the RNA extraction could restore slower migrating RNA species for MoMLV RNA (Tchenio and Heidmann, 1995). In contrast, the dimerisation behaviour of the *in vitro* transcribed RNAs 4 and 6 was significantly affected by phenol extraction (Section 3.21.1). Therefore in the KL- viruses the mutation was responsible for a reduced amount of RNA dimer, this was not a consequence of phenol extraction. To investigate if this mutation had any effect on dimer stability of the KL- virus RNA compared to the other viruses, thermal stability of the mutant virion RNAs was determined.

#### **4.14.3. Thermal stability of HIV-1 Wild Type and Mutant RNA Dimers.**

To determine whether any of the mutations had an effect on the stability of virion RNAs, thermal stability of the mutant viral RNA dimers was compared with wild type virus. HIV-1<sub>NL43</sub> virus was used to infect C8166 T cells and the virion RNA harvested. Aliquots of the RNA were resuspended in a buffer containing 10mM Tris pH 7.5, 100mM KCl and 5mM MgCl<sub>2</sub>. This buffer composition was selected to represent cytosolic ionic conditions (Alberts et al., 1989). The samples were incubated at the temperatures shown for 5 min then electrophoresed on a 0.8% agarose TBE gel at 4°C, see figure 4.24. For the wild type virus the dimeric RNA band dissociated at around 60°C in this buffer. This result was similar to those of Fu et al., (Fu et al., 1994).

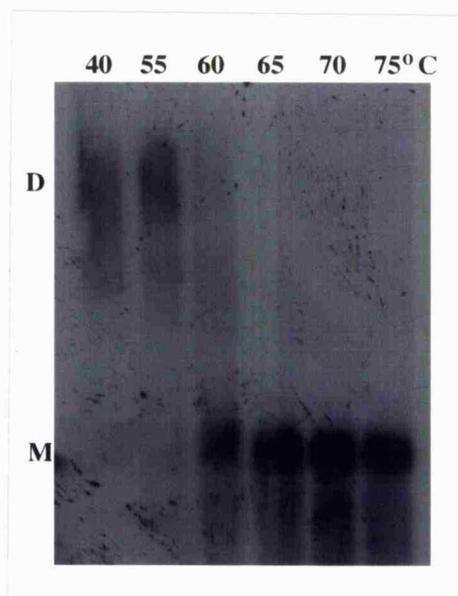


Figure 4.24. Thermal stability of wild type HIV-1 RNA. The RNA was incubated for 5 min at the temperatures indicated.

The effect of heat denaturation on the mutant RNA dimers was also assessed in the same buffer, figure 4.25.

With an initial increase in temperature the dimer bands became more defined in all of the samples, see figure 4.25. Compare for instance HIV Tet<sup>-</sup> RNA at 40°C and at 55°C. It was particularly apparent in the HIV KL<sup>-</sup> and HIV Tet<sup>-</sup>KL<sup>-</sup> RNA samples. The mobility of the RNA dimer also decreased slightly as the temperature increased probably due to partial dissociation and the adoption of a more relaxed structure. This behaviour has been noticed for RSV virion RNA (Lear et al., 1995). Dimeric RNA bands disappeared between 60 to 65 °C for all of the HIV RNA samples. These values were consistent with the behaviour of wild type HIV-1 RNA (figure 4.24).

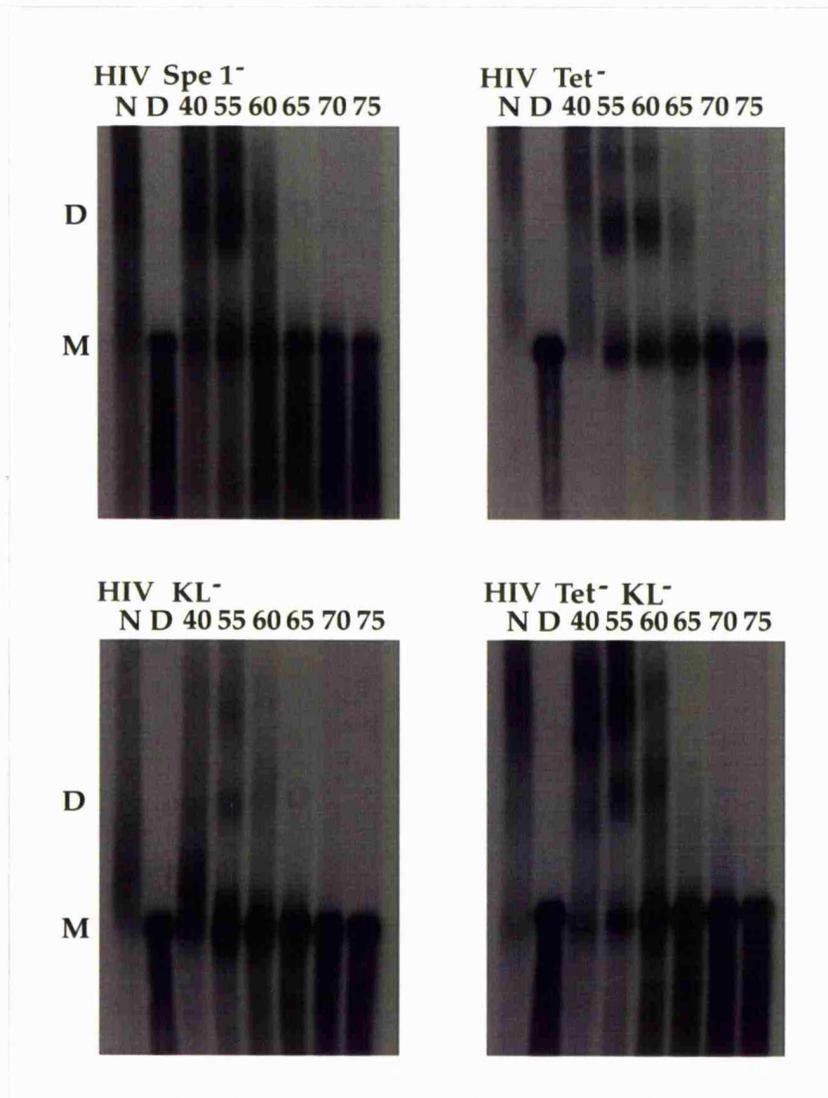


Figure 4.25. Thermal stability of mutant HIV RNA dimers HIV SpeI-, Tet-, KL- and Tet-KL-. The RNA was incubated at the temperatures shown for 5 min. D = dimer, and M= monomer.

Interestingly the bands which had a lower mobility than the dimer RNA, had the same melting temperature as the dimer bands. This may indicate that these upper bands are trimers or tetramers of genomic RNA. Attempts at phosphorimager analysis of the bands to determine a  $T_m$  value (as normally defined), cannot be derived from this experiment because of the complex behaviour of the RNA.

There was no significant difference in the thermal stability of the KL-RNA dimers. This suggested that the reduced amount of dimer within these viruses was probably not connected with the stability of the dimer that was present. Similarly the Tet- virus RNA dimers did not show any reduction in thermostability, indicating that the guanine tetrad defined *in vitro* was not important for virion RNA dimer stabilisation.

#### **4.15. Mutant Virus Summary.**

The tetrad G<sub>819</sub> to U, and the kissing loop 7<sub>11</sub>GCG to 7<sub>11</sub>AAA mutations were introduced into infectious HIV-1<sub>NL43</sub> to produce HIV Tet-, HIV KL- and double mutant HIV Tet-KL- viruses. These mutations did not prevent viral replication or infectivity, although the KL- viruses appeared to have a reduced growth rate. Each virus was able to recognise and package its genomic RNA. The dimerisation ability of each RNA was analysed; the tetrad mutation had not altered the genomic RNA dimer compared to the wild type virus. In contrast, there was a reduced amount of RNA dimer in the HIV KL-, and Tet-KL- viruses, with a corresponding increase in the amount of monomer RNA. The RNA dimer did not have a significant reduction in thermal stability for any of the mutations compared to the wild type virus.

This analysis demonstrated that the sequences 7<sub>11</sub>GCG were important for HIV-1 virion RNA dimerisation. KL- viruses had a reduced growth rate and a reduced level of dimer RNA. The effect of the 7<sub>11</sub>GCG to AAA mutation was consistent with the formation of a kissing loop structure *in vivo*. The kissing loop complementarity was not essential for RNA dimer formation, however its presence increased the probability of normal virion RNA dimer formation.

## Chapter 5: Discussion.

### 5.1 Summary of *in vitro* RNA Dimerisation.

From the *in vitro* RNA dimerisation analysis, two classes of RNAs were identified, those that dimerised by the formation of a guanine tetrad, or a kissing loop structure.

#### Guanine tetrad RNAs

RNAs 1, 2, 10, 11 and 12 dimerised by the formation of a guanine tetrad, figure 3.1 and 3.2. For all of the RNAs the guanine tract  $_{817}\text{GGGG}_{820}$  was essential, in agreement with earlier studies (Sundquist and Heaphy, 1993; Awang and Sen, 1993). For RNA 1 the partner guanine tract was  $_{732}\text{GGG}_{734}$  and for RNA 2 it was  $_{731}\text{GGGG}_{734}$ . However for RNA 12 dimerisation a guanine tetrad formed using sequences  $_{817}\text{GGGG}_{820}$  and  $_{836}\text{GGG}_{838}$ , in accord with previous results (Awang and Sen, 1993). These results indicated the flexibility of tetrad formation. In fact for RNA 2, a mixed population of dimers was probably present using the essential  $_{817}\text{GGGG}_{820}$  tract and both  $_{731}\text{GGGG}_{734}/_{836}\text{GGG}_{838}$  sequences. Mutagenesis of the  $_{817}\text{GGGG}_{820}$  tract to  $_{817}\text{GGUG}_{820}$  in RNA 3 almost completely abolished dimerisation of this RNA reinforcing the importance of this tract, figure 3.13. An increase in thermal stability with potassium ions and cation dependent thermal stability with  $\text{Li}^+$ ,  $\text{Cs}^+$  and  $\text{Na}^+$ , indicated that a tetrad consisting entirely of guanines was involved in these RNA dimerisations (figure 3.9).

#### Kissing Loop RNAs.

RNAs 4, 6 and 8 (figures 3.1 and 3.14) contained additional 5' sequences compared to the tetrad RNAs. For these RNAs the sequence  $_{711}\text{GCGCGC}_{716}$  probably exposed at the apical tip of a stem loop structure

was essential for dimerisation. RNA dimerisation was by a kissing loop mechanism as proposed by Skripkin et al. and subsequently confirmed *in vitro* (Skripkin et al., 1994; Laughrea and Jette, 1994 & 1996; Paillart et al., 1994; Muriaux et al., 1995 & 1996). In agreement with this model mutagenesis of RNAs 4, and 6 producing RNAs 5 and 7 containing the 711AAACGC<sub>716</sub> mutation, almost completely abolished dimerisation of these RNAs (figure 3.17). It appeared that the kissing loop structure was the primary determinant for RNA dimerisation, as RNA 7 also contained the tetrad guanine tracts 731GGGG<sub>734</sub> and 817GGGG<sub>820</sub> but no tetrad formation was observed under the conditions investigated. An unexpected result was the formation of kissing loop dimers on phenol extraction of the RNA (figures 3.26 and 3.27).

## **5.2 Conservation of Tetrad and Kissing Loop sequences.**

The *in vitro* RNA analysis identified the 817GGGG<sub>820</sub> and 711GCGCGC<sub>716</sub> sequences as important for RNA dimerisation. Comparison of HIV-1 strain sequences and the high conservation of the kissing loop and tetrad elements suggested that these sequences are also important *in vivo*, figure 5.1. However it should be appreciated that the HIV-1 sequences in the database may not necessarily originate from replication competent viruses.

Figure 5.1 Kissing Loop 711GCGCGC716 Guanines 731GGGC734

Consensus A:	GAGGTGC . . ACAAC	GGCGAGAG . CGGCGA
HIV U455	----- . -----	----- . -----
Consensus B:	gaAGCGC? ?GCaCg	GGCGAGGGGcGGCGa
HIVLAI	----- . -----	-----A-----
HIVHXB2	-----CC-----	-----
HIVMN	----- . -----	-----
HIVJH31	----- . -----A	-----G-----
HIVJRCSF	----- . -----A	-----
HIVJRFL	----- . -----	-----
HIVOYI	T . ----- . -----	-----
HIVSF2	----- . -----A	-----
HIVNY5	----- . -----	-----
HIVNL43	----- . -----	-----
HIVCAM-1	----- . -----A	-----
HIVCDC41	----- . -----	-----
HIVHAN	----- . -----	-----
HIVD31	----- . -----A	-----
HIVRF	----- . --G--	-----A-----
HIVYU2	----- . -----	-----
Consensus O:	Ga?GtGC . . aCacA	GGCGAGGg . ?CGGcgA
HIVMAL	--G----- . -----	-----A.G-----

Guanines 817GGGGG821

Consensus A:	TaAgtGggggaaaAT	B:	TaAGcGgcGCgagAAT
HIVU455	----C--AA-----	HIVSF2	-----
HIVMAL	----C-----	HIVBZ167	-----
HIVVI59	----C-A-----	HIVPH153	-----GA-----
HIVVI310	----C-----	HIVPH136	-----
HIVVI57	----C-A-----	HIVBZ200	-----
HIVK112	-----	HIVTB132	-----C-----
HIVK88	-----	HIVBZ190	-----A-----
HIVK29	-----	HIVLAI	-----
HIVK124	----C-----	HIVHXB2	-----
HIVK7	----C-----C--	HIVMN	-----
HIVK98	-----	HIVJH3	-----
HIVK89	----C-----	HIVJRCSF	-----
HIVI32	-----	HIVOYI	-----
HIVVI415	-----	HIVNY5CG	-----
HIVCI4	-----	HIVNL43	-----
HIVG141	----C-----	HIVCDC4	-----
HIVLBV23	----C+-----	HIVHAN	-----
HIVTN243	-----	HIVCAM1	-----
HIVTN245	-----	HIVRF	-----C--A-----
HIVTN240	-----	HIVD31	-----
HIVCI20	-G-----	HIVBH102	-----
HIVLBV231	-----	HIVPV22	-----
HIVCI51	-----	HIVJRFL	-----A-----
HIVLBV105	----C-----		
HIVCI32	-----CG--		
HIVC144	----C--A-----		
HIVCM238	----C-----		
HIVUG266	----C-A--G-----		
HIVVI354	----C-----		

Figure 5.1. Sequence conservation for tetrad and kissing loop tracts. (-) indicates identity with the top (reference) sequence, (·) indicates gaps. The consensus sequence for each of the subtypes reflect residues conserved 50% or better (upper case letters indicate 100% conservation) in the sequences below them.

The kissing loop tract  $711\text{GCGCGC}_{716}$  is present in the non-coding RNA leader and so it may be expected to have reduced pressure for sequence conservation. However the sequences are highly conserved. More importantly where sequence identity diverges, the self complementarity of this tract essential for its role in a kissing loop structure, remains. eg.



Figure 5.2 Kissing loop sequences for 3 HIV-1 strains.

The HIV-1<sub>NL4-3</sub> and <sub>MAL</sub> sequences are different but the palindrome remains. Exceptionally, HIV-1<sub>HXB2</sub> contains a dinucleotide -CC- insertion within the kissing loop region. The reason for this is unknown, and it is uncertain whether a kissing loop could form in this RNA. Given that the 711-716 sequence is exposed at the tip of the loop and is apparently not required for formation of the stem structure, the most easily identifiable reason for its conservation could be a role in RNA packaging, or the

kissing loop mechanism of RNA dimerisation *in vivo*. Alternatively there could be other unknown reasons why the autocomplementarity of this sequence is maintained.

Sequence conservation of the tetrad regions is also high. The  $731\text{GGGG}734$  tract is also present in non-coding RNA with no obvious reason for sequence conservation, other than a role in structural integrity or guanine tetrad formation. Where sequence divergence does occur eg. HIV-1<sub>U455</sub> and HIV-1<sub>MAL</sub> the tract remains purine rich.

For the  $817\text{GGGG}820$  tract essential to the dimerisation of tetrad RNAs *in vitro*, sequence conservation is again high. There is some restriction on sequence divergence of this tract because it is part of the *gag* gene. When divergence from the consensus occurs, there are at least 3 contiguous guanine residues present within the 817-821 region, figure 5.3.

HIV-1 <sub>NL4-3</sub>	5'- G G G G G -3'
HIV-1 <sub>U455</sub>	5'- G G G A A -3'
HIV-1 <sub>RF</sub>	5'- G C G G G -3'
HIV-1 <sub>C144</sub>	5'- G G A G G -3'

Figure 5.3. Example  $817\text{GGGG}821$  sequences for four HIV-1 strains.

HIV-1<sub>C144</sub> contains an adenine residue at position G<sub>819</sub>, similar to the G<sub>819</sub> to U mutation which abolished RNA 3 dimerisation *in vitro*. Both changes do not affect *gag* coding. G<sub>819</sub> is the third base in a glycine codon and could equally well be A, C or U for this coding purpose, yet HIV-1<sub>C144</sub> was the only strain to have this change. For the other strains one possibility is that a three plane tetrad may form using any three guanines within the  $817\text{GGGG}821$  tract.

The  $836\text{GGG}_{838}$  sequence (RNA 12 and Awang and Sen, 1993) is also conserved (not shown). This may be constrained because of protein folding requirements, it forms codons for a Trp-Glu sequence and no other coding possibilities are permitted.

Although these arguments are not conclusive for a role of these sequences in virion RNA dimerisation, they are consistent with the kissing loop or guanine tetrad models. Alternatively the high degree of conservation suggests that if these sequences are not involved in RNA dimerisation they are important for some other function.

### **5.3 Specificity of *in vitro* RNA Dimerisation.**

The mechanism of RNA dimerisation *in vitro*, was found to be RNA sequence dependent. RNAs containing sequences primarily downstream from the splice donor site formed tetrad dimers, whereas longer RNAs extending further 5' into the leader region dimerised by the formation of a kissing loop structure. This suggests that specificity for RNA dimerisation comes from RNA structures which form, ie. the sequence tracts identified have to be in the correct configuration to allow a guanine tetrad or a kissing loop to form. In a 615 nucleotide HIV-1 MAL RNA, which dimerised by the formation of a kissing loop, Paillart et al showed that deletion of purine tracts reduced dimer stability (Paillart et al., 1995). In contrast for a 178 nucleotide RNA (the sequence of which was contained within the 615 nucleotide RNA) Muriaux et al., detected no effect on the dimer by mutagenesis of the purine tracts (Muriaux et al., 1995). A similar result was found here, several guanine tetrads were able to form depending on the G-tracts present eg. RNA 2 and RNA 12, or no tetrads formed at all eg. RNA 4. Alternatively the interaction of the nucleocapsid

protein and RNA *in vivo* may dictate the correct conformation for dimerisation.

The *in vitro* studies described here support previous studies on the ability of RNAs from the DLS of HIV-1 to form dimers by a guanine tetrad structure (Marquet et al., 1991; Sundquist and Heaphy, 1993; Awang and Sen, 1993; Weiss et al., 1994), although the involvement of the  ${}_{731}\text{GGGG}_{734}$  sequence has not been previously reported. Similarly the kissing loop mechanism has been extensively investigated (Skripkin et al., 1994; Laughrea and Jette, 1994, 1996; Paillart et al., 1994; Muriaux et al., 1995) and accounts for the results described here for RNAs beginning at position 649.

However caution is required when correlating separate results as slightly different RNA sequences can introduce markedly different dimerisation behaviour eg. RNA 4 (kissing loop 649-813) and RNA 2 (tetrad 708-858). Similarly the presence of non-native nucleotides at the 5' or 3' end of the RNAs may have an effect on the RNA structure. Also different buffer conditions can influence RNA dimerisation eg. HIV-1 RNA dimerisation was reported to proceed better in 0.1mM  $\text{Mg}^{2+}$  than 1mM (Laughrea and Jette 1994 & 1996). The incubation conditions in which RNA dimers were formed also affected their melting temperatures (Fu et al., 1994). An unusual result observed here was that phenol extraction specifically promoted the dimerisation of kissing loop RNAs 4 and 6, demonstrated by heterodimer formation (figure 3.27). Phenol extraction of tetrad RNA dimers eg. RNA 2 did not assist dimerisation (figure 3.21 and Darlix et al., 1990; DeRocquigny et al., 1992). Similarly RNA isolated from virions was not multimerised by extraction procedures (Figure 4.23 (B) and Tchenio and Heidmann, 1995).

RNA dimerisation has only been studied extensively in the context of the 5' leader regions of retroviral RNAs. There is some evidence for specificity ie. antisense RNAs and larger RNA fragments fail to dimerise. On the other hand there are a large number of RNAs which can dimerise and in many different ways. This could reflect a general feature of RNAs or be representative for virion RNAs. It is difficult to be certain because other RNA sequences which are not involved in retroviral RNA dimerisation have not been extensively studied for their ability to dimerise. Therefore it has to remain an open question as to the real value of these *in vitro* studies. Testing *in vitro* models *in vivo* will ultimately determine the validity of sub-genomic RNA dimerisation analysis. Negative results *in vivo* make the situation unclear, whereas positive results eg. a reduction in RNA dimerisation, suggests something of significance.

#### **5.4 *In vivo* RNA Dimerisation for the Mutant HIV-1 Viruses.**

Having defined the effects of the 711GCG to 711AAA and G<sub>819</sub> to U mutations *in vitro*, these sequences were introduced into an infectious molecular clone of HIV-1<sub>NL4-3</sub> producing the single mutant viruses HIV Tet-, HIV KL-, and the double mutant HIV Tet-KL-.

##### **Guanine Tetrad.**

The G<sub>819</sub> to U point mutation that significantly reduced tetrad mediated dimerisation *in vitro*, was introduced into a single but conserved guanine tract. The G<sub>819</sub> to U mutation had no apparent effect on viral replication (figure 4.18), RNA packaging, or dimerisation (figure 4.23) of the HIV Tet-virus genomic RNA. The thermal stability of the Tet- virus RNA dimers was also unaffected and indistinguishable from wild type (figure 4.25). No increase in thermal stability was observed for any viral RNA dimers in the presence of potassium. These observations indicate that a guanine tetrad

involving G<sub>819</sub> or any other guanine quartet structure was not present in HIV-1 virion RNA, in agreement with previous results (Fu et al., 1994). Therefore the proposal of a bi-partite dimerisation domain claimed *in vitro*, where a guanine tetrad structure was proposed to have a role in dimer stability after dimer initiation by a kissing loop, is unlikely to apply *in vivo* (Paillart et al., 1994).

The tetrad mutagenesis described was appropriate given the *in vitro* results but is clearly not exhaustive. It remains possible that a tetrad may still form in the DLS requiring other purine residues, either at dimer formation or subsequently in maturation (Paillart et al., 1994). Furthermore the tetrad itself may be composed of several disparate elements, not involving any previously identified residues. These sequences could be a target for further mutagenesis of the virion RNA. Perhaps the tetrad, if it is present, is only a transient structure. Alternatively, the behaviour of a 20 kB RNA dimer is unlikely to be the same as a 300 nucleotide RNA dimer *in vitro* (Sundquist and Heaphy, 1993), so perhaps the thermal stability assay cannot be applied to virion RNA.

In summary these observations are consistent with the lack of guanine tetrad involvement in the dimerisation of HIV-1<sub>NL4-3</sub> RNA *in vivo* involving residues G817-821, in contrast to results obtained using sub-genomic RNAs *in vitro* (Sundquist and Heaphy, 1993; Awang and Sen, 1993; Weiss et al., 1993). This analysis cannot rule out the possibility of a tetrad in the virion, eg. if a tetrad is present in virion RNA it may not be functionally important. The only evidence for the guanine tetrad mechanism of RNA dimerisation is *in vitro*.

### Kissing Loop.

For the kissing loop model, the autocomplementary loop sequence and stem structure appear to be the primary determinants for RNA dimerisation *in vitro*. There is no evidence for the involvement of additional residues or redundancy in the sequence requirements for dimer initiation. Therefore if the kissing loop structure forms *in vivo*, it should be prevented from initiating dimerisation by the 711GCG to 711AAA mutation. However this does not exclude the possibility of other kissing loops in the HIV-1 leader which may be able form dimers.

*In vivo*, the 711AAA mutation has altered the properties of the HIV KL- and HIV Tet-KL- viral RNA dimers, compared to wild type (HIV SpeI-). All kissing loop mutated viruses (KL-) were infectious but had a reduced growth rate (figure 4.18). The intention of this experiment was to determine whether these sequences were required for growth of the viruses or not. Therefore the p24 titre in the culture supernatant and not specific growth rates were measured. An end point dilution assay would be a better test to quantify the reduction in growth rate of the KL- viruses. Nevertheless the KL- viruses grew more slowly. Analysis of the virion RNA showed that the KL- viruses contained a reduced amount of dimer (figure 4.23). Under native conditions the dimer band was not well defined, however it became more visible with increasing temperature and is best seen in figure 4.25.

For most retroviral RNA preparations, under native conditions, bands are present with a mobility greater than a dimer (figure 4.23, Darlix et al., 1990). The identity of these is uncertain, and are perhaps trimers or tetramers of virion RNA. Their thermal stabilities are similar to dimers, suggesting a similar linkage between the RNA strands (figure 4.25).

Alternatively these upper bands could be dimers with reduced electrophoretic mobility due to bound proteins or lipids. For the Tet-KL-virus, there was a larger proportion of the higher molecular weight RNA bands, the reason for this was unknown. Perhaps these upper bands are an artefact of RNA isolation.

The KL- and Tet-KL- viruses also appeared to contain a higher percentage of monomeric RNA under native conditions than wild type virus ie. the monomer : dimer ratio was increased (figure 4.23). This suggested that the dimer could be unstable, or dimer formation had been affected.

Thermal stability of the KL- dimers showed that there was no significant difference in the dissociation temperature when compared to wild type viruses (figure 4.25). This indicated that a low stability RNA dimer had probably not formed by any loop stacking interaction as described *in vitro* (Laughrea and Jette 1996; Muriaux et al., 1996). It also suggests that the dimers formed were not immature (Fu et al., 1994). The effect of increasing temperature during the thermal stability determinations for all of the virus RNAs lead to a better definition of the dimer band, especially the KL- dimers (figure 4.25). The KL- virion RNA was clearly different from the other viruses, this and the reduction in viral growth rate warrants further examination.

The behaviour of the KL- virus RNA dimers can be interpreted in two ways. Firstly the dimer band becomes better defined if the other (weaker) RNA contacts along the genome are dissociated first, removing the 'smearing' associated with the native samples. Alternatively, dimer formation may occur with increasing temperature. *In vitro*, loose and tight dimers have been proposed for HIV-1 RNA (Laughrea and Jette,

1996) which corresponded to the initial loop:loop interaction followed by strand unwinding and inter-strand hydrogen bond formation (Muriaux et al., 1996). This effect was achieved by incubation of HIV-1 RNA dimers at 55°C, presumably overcoming some thermodynamic block to stable dimer formation. Similar behaviour has been observed for HaSV RNAs *in vitro* (Feng et al., 1995). *In vivo* it has been proposed that the nucleocapsid protein would substitute for the incubation at elevated temperatures (Muriaux et al., 1996). Therefore the lack of a fully functional kissing loop structure for the KL- virus could be overcome by the nucleocapsid protein (and elevated temperature) to allow a dimer to form with similar thermal stability to wild type. So the kissing loop dimer initiated by sequences 711-716 was not essential for dimer stability.

Perhaps mutagenesis of the  $_{711}\text{GCGCGC}_{716}$  sequence lead to problems initiating dimer formation i.e. the DIS was disrupted. A reduction in the efficiency of dimer formation could account for the lower amount of RNA dimer in the KL- particles, along with the increase in the amount of monomer. However dimerisation was not completely prevented, and dimers which formed had similar thermal stabilities to wild type viruses. This suggested that events after dimer formation such as maturation proceeded normally and although important for genomic RNA dimerisation, the kissing loop structure was not absolutely essential.

One possible interpretation of the results obtained with the HIV KL- and HIV Tet-KL- virus cultures is that they contain two populations of virus. This situation could arise for several reasons. It was possible that the cultures had been contaminated with KL+ virus. This can be discounted because these observations have been reproduced on occasions when KL+ viruses were not being cultured in the laboratory. More compellingly

DNA amplified from HIV KL- and HIV Tet-KL- virion RNA by RT-PCR was resistant to BssHII cleavage (KL+ sequences are cleaved by BssHII, figure 4.21). Also the viruses isolated from cell culture had the expected DNA sequence (figure 4.22). It was estimated that any contaminant KL+ sequence would comprise less than 5% of the DNA and probably less than 1%. Another possibility was that a second site change has occurred in the KL- cultures. This cannot be excluded because it would entail sequencing the entire virus genome which was not performed. If it were true then it has happened quickly and on two separate occasions, but it has not restored dimerisation to wild type levels. A third and most likely possibility, was that although the KL- cultures were clonal they exhibited a complex phenotype, figure 5.4.

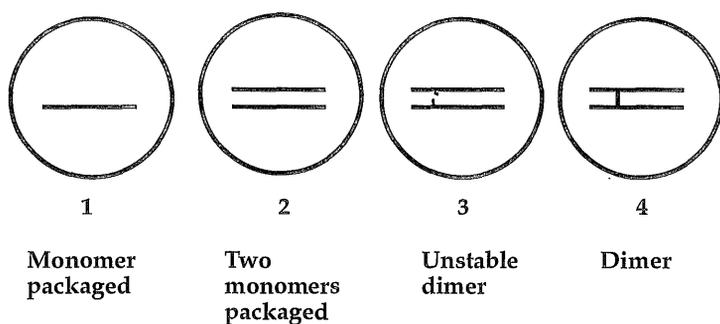


Figure 5.4. Schematic representation of possible HIV KL- virion RNA conformations.

A large proportion of the KL- particles may contain monomer, two unlinked RNAs, or an unstable dimer that may not be functional. These viruses would probably be replication defective, or at best replicate very slowly. A smaller proportion of particles may contain a functional genomic RNA dimer, this could also account for the reduced growth rate of the HIV KL- and Tet-KL- viruses.

Another explanation was that one population of virus existed, but the RNA dimer it contains had a reduced functionality. Perhaps the disruption of the kissing loop sequences prevented favourable RNA:RNA interactions occurring at the same rate as wild type dimer initiation. However dimer formation may still occur at a reduced rate, and is then stabilised by the other RNA contacts throughout the genome. It is possible that these contacts themselves may be the mechanism where intact dimers form in the KL- and other viruses, in an analogous situation to MoMuLV (Tchenio and Heidmann, 1995). In this study removal of the first two thirds of the MoMuLV genome containing the DLS region, still allowed the residual RNA to dimerise and establish a normal provirus. This presumably highlighted the importance of the other RNA strand contacts along the genome, or indicated that RNA dimerisation within the virion did not depend on specific cis-acting RNA sequences. Therefore a reduction in correct dimer initiation could be responsible for the lower amounts of dimeric RNA in the KL- viruses, however once a dimer was formed it could be stabilised by the other interstrand contacts in the genome.

This leads to the conclusion that the kissing loop structure was not absolutely essential for dimerisation *in vivo*, although its presence increases the probability of normal RNA dimer formation (Haddrick et al., 1996). Therefore the RNA sequences 711GCG are important for viral RNA dimerisation, perhaps as part of a kissing loop structure. This structure may act as a type of "catalyst" promoting an ordered annealing of packaged virion RNA.

#### Other Retroviral Autocomplementary Sequence Loop Structures.

There is evidence for the importance of autocomplementary sequences on stem loops for RNA dimerisation of other retroviruses, at least *in vitro*. For MoMLV a 5'-UAGCUA-3' sequence within the encapsidation domain may be important for RNA dimerisation (Girard et al., 1995). For HaSV a two stem loop structure was proposed in an RNA corresponding to sequences 205-271 (Feng et al., 1995). The self complementary sequence 5'-GGCC-3' is present in the loop and may be involved in RNA dimerisation. A similar loop sequence could be present in the structure of a dimerisation competent VL30 rat retrotransposon RNA (Torrent et al., 1994; S. Heaphy, personal communication). Similarly a structural analysis of BLV RNA showed an extensive stem loop structure with a 5'-UGAUCA-3' sequence exposed at the tip of a loop (Kato et al., 1993), this region was also important for RNA encapsidation (Kurg et al., 1995). However no mutagenesis of these sequences *in vivo* has been reported.

Further analysis is required to confirm the existence of a kissing loop complex for any retrovirus *in vivo*. It is formally possible that the 711GCG to 711AAA mutation described for HIV KL- and Tet-KL- may be unconnected with a kissing loop structure. The structure of virion RNA at the DLS is unknown *in vivo*, although the existence of the kissing loop stem has been supported from *in vitro* RNA analysis (Harrison and Lever, 1992; Skripkin et al., 1994; Paillart et al., 1995; Laughrea and Jette 1994 & 1996; Muriaux et al., 1994 & 1996; Clever et al., 1995; McBride and Panganiban 1996). For HIV-1 the kissing loop sequences are located upstream of the major 5' splice donor site, so they are present on all sub-genomic RNAs as well as unspliced genomic RNA. Only genomic RNA dimerises, this suggests that the kissing loop structure may not form on sub-genomic RNAs. Alternatively if the kissing loop structure does form

perhaps additional undefined interactions, or the action of the nucleocapsid protein, are necessary to dimerise virion RNA. Furthermore the overall parallel orientation of retroviral RNAs is inconsistent with an antisense kissing loop interaction, other structural features must be present within the DIS/DLS to re-align the two RNAs in parallel. It is also unknown to what extent the inter-strand base pairing actually is between the two RNAs following the initial kissing loop interaction. By electronmicroscopy MoMLV and RSV genomic RNAs were in contact within a 50 nucleotide region (Murti et al., 1981), presumably this is sufficient to account for their observed thermal stabilities.

Although the effects on RNA packaging of the G<sub>819</sub> and 7<sub>11</sub>AAA mutations were not specifically addressed, it was clear that neither mutation had drastically affected RNA packaging. Clearly complete RNA dimerisation was therefore not essential for RNA incorporation in the KL-viruses. It is difficult to make any further conclusions. The KL- virus may have packaged RNAs that failed to dimerise, or packaged a lower amount of stable dimer. It has recently been proposed that loop:loop RNA interactions are involved in the packaging of HIV-1 RNA, these may be related to RNA dimerisation as the kissing loop structure is one of the proposed packaging elements (Clever et al., 1995; McBride and Panganiban 1996).

### **5.5 Nature of the 5' DLS.**

This study supports the concept that the leader region of retroviral genomic RNA is involved in dimerisation. The 7<sub>11</sub>GCG to 7<sub>11</sub>AAA mutation significantly reduced the growth rate of and level of dimeric genomic RNA within KL- viruses. However the dimer which formed had similar thermal stability to wild type. Therefore the contribution of other

contacts throughout the genome may be an important component of RNA dimerisation.

The nature of the less stable contacts between the RNAs is unclear. Firstly many points of contact occur between retroviral RNA genomes, visualised by electronmicroscopy under semi-denaturing conditions (Mangel et al., 1974). Heat or chemical denaturation showed that the 35S monomeric RNA isolated from virions is substantially nicked. Yet the 70S dimer constituted from these monomers shows much lower signs of degradation (figure 4.23 to 4.25). This can be interpreted to mean that the two RNA strands are associated at many points along the genome and not just at the DLS. Therefore if mutagenesis of the kissing loop structure completely prevented RNA dimerisation then the other contacts may substitute, in agreement with a deletion analysis of MoMLV (Tchenio and Heidmann, 1995). It can be suggested then that the 5' DLS is probably the most stable point of contact along the RNAs, it has an important but not exclusive role in genomic retroviral RNA dimerisation. For HIV-1, electronmicroscopy of genomic RNA with different concentrations of formamide and urea identified a central dimer linkage structure, associated with a small loop also present in a fraction of the monomeric RNA (Hoglund et al., 1995).

Perhaps multiple dimerisation domains exist on the virion RNA, there is evidence for these *in vitro*. For HIV-1 a tetrad mediated DLS may form after the kissing loop DIS interaction (Paillart et al. 1995). Similarly RSV (Lear et al., 1995) and HaSV (Feng et al., 1995) RNAs also contain multiple dimerisation domains. These same interactions may occur throughout the virion RNA and be involved in RNA dimerisation. However by the very nature of isolating and focussing on dimerisation of small RNAs *in vitro* the analysis may be not representative. This was the case for the guanine

tetrad, evidence for its formation *in vitro* was good, but it appeared to be absent *in vivo*. In contrast it is likely that the kissing loop identified *in vitro*, does have a role in the dimerisation of HIV-1 RNA *in vivo*.

#### **5.6 Future Experiments.**

Future experiments will focus on the HIV KL- and HIV Tet-KL- viruses. Precise determinations of viral growth rates should be performed compared to wild type viruses. Further less dramatic mutagenesis of the kissing loop structure would be interesting. Perhaps partial disruption of the autocomplementary loop residues may identify a correlation between virion RNA dimer stability or formation, and the stability of the stem loop structure. The ability of RNA sequences outside of the 5' leader region may be considered for their contributory role in RNA dimerisation, as multiple points of contact exist between the two RNAs. It is difficult to design experiments to determine the contribution of these residues separate from those in the DLS, as the position and nature of the multiple contact residues is unknown. These could be identified by an *in vitro* analysis, however the correlation with virion RNA is not certain. Perhaps an analysis of non-retroviral RNAs is also required to determine whether oligomerisation can be a general feature of RNA molecules.

To determine if dimerisation precedes or occurs after packaging a dimerisation mutant must be obtained. However it is uncertain that this can ever become available if the multiple contacts throughout the genome are the mechanism of dimerisation. Similarly the extent of dimer maturation in the virion needs to be assessed, this could be examined by multiple virus harvests at short time intervals. Ideally an *in vivo* footprinting technique would be developed in which the reactivity of residues involved in RNA dimerisation could be followed during virion

assembly. In the long term, an understanding of RNA dimerisation may yield a target for antiretroviral therapy.

## **References**

Aboul-ela F., Murchie A. I. H. and Lilley D. M. (1992). NMR study of parallel-stranded tetraplex formation by the hexanucleotide d(TG<sub>4</sub>T). *Nature* 360: 280-282.

Adachi A., Gendelman H. E., Koenig S., Folks T., Willey R., Rabson A., and Martin M. A. (1986). Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* 59: 284-291.

Adam M. and Miller A. D. (1988). Identification of a signal in a murine retrovirus that is sufficient for packaging of non-retroviral RNA into virions. *J. Virol.* 62: 3802-3806.

Alberts et al., (1989). *Molecular biology of the cell.* Garland Publishing Inc., New York.

Aldovini A. and Young R. A. (1990). Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in the production of noninfectious virus. *J. Virol.* 64: 1920-1926.

Alford R. L., Honda S., Lawrence C. B., and Belmont J. W. (1991). RNA secondary structure analysis of the packaging signal for Moloney murine leukemia virus. *Virol.* 183: 611-619.

Armentano S., Yu S. F., Kantoff P. W., von Ruden T., Anderson W. F., and Gilboa E. (1987). Effect of internal viral sequences on the utility of retroviral vectors. *J. Virol.* 61: 1647-1650.

Aronoff R., Hajjar A. M., and Linial M. L. (1993). Avian retroviral RNA encapsidation: reexamination of functional 5' RNA sequences and the role of nucleocapsid cys-his motifs. *J. Virol.* 67: 178-188.

Awang G. and Sen D. (1993). Mode of dimerization of HIV-1 genomic RNA. *Biochem.* 32: 11453-11457.

Babe L. M., Rose J. and Craik C. S. (1995). Trans dominant inhibitory human immunodeficiency type 1 protease monomers prevent protease activation and virion maturation. *Proc. Natl. Acad. Sci. (USA)* 92: 10069-10073.

Bader J. P. and Ray D. A. (1976). Configurational variants of Oncovirus RNAs. *J. Virol.* 19: 810-819.

Bader J. P. and Steck T. L. (1969). Analysis of the ribonucleic acid of murine leukemia virus. *J. Virol.* 4: 454-459.

Baltimore D. (1970). Viral-RNA dependent DNA polymerase. *Nature* 226: 1209-1211.

Banner D. W., Kokkinidis M. and Tsernoglou D. (1987). Structure of the ColE1 ROP protein at 1.7Å resolution. *J. Mol. Biol.* 196: 657-675.

Barat C., Lullien V., Schatz O., Keith G., Nugeyre M., Gruninger-Leitch F., Barre-Sinoussi F., LeGrice S. F. J., and Darlix J-L. (1989). HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO J.* 8: 3279-3285.

Barre Sinoussi F., Chermann J-C., Rey F., Nugeyre M. T., Chamaret S., Gruest J., Dauguet C., Axler-Blin C., Vezinet-Brun F., Rouzioux C., Rozenbaum W., and Montagnier L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220: 868-871.

Baudin F., Marquet R., Isel C., Darlix J., Ehresmann B., and Ehresmann C. (1993). Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J. Mol. Biol.* 229: 382-397.

Bender M. A., Palmer T. O., Gelinas R. E., and Miller D. (1987). Evidence that the packaging signal of Moloney murine leukemia virus extends into the gag region. *J. Virol.* 61: 1639-1646.

Bender W., Chien Y-H., Chattopadhyay S., Vogt P. K., Gardiner M. B., and Davidson N. (1978). High-Molecular-Weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. *J. Virol.* 25: 888-896.

Berkhout B., Oude-Essink B. B., and Schoneveld I. (1993). *In vitro* dimerization of HIV-2 leader RNA in the absence of PuGGAPuA motifs. *FASEB J.* 7: 181-187.

Berkowitz R. D., Hammarskjold M.-L., Helga-Maria C., Rekosh D., and Goff S. P. (1995). 5' Regions of HIV-1 RNAs are not sufficient for encapsidation: implications for the HIV-1 packaging signal. *Virol.* 212: 718-723.

Bieth E., Gabus C., and Darlix J. (1990). A study of the dimer formation of Rous sarcoma virus RNA and of its effect on viral protein synthesis *in vitro*. *Nucleic Acids Research* 18: 119-127.

Bishop J. M., Levinson W. E., Quintrell N., Sullivan D., Fanshier L., and Jackson J. (1970a). The low molecular weight RNAs of Rous sarcoma virus. *Virol.* 42: 182-195

Bishop J. M., Levinson W. E., Sullivan D., Fanshier L., Quintrell N., and Jackson J. (1970b). The low molecular weight RNAs of Rous sarcoma Virus. *Virology* 42: 927-937.

Bonar R. A., Sverak L., Bolognesi D. P., Langlois A. J., Beard D., and Beard J. W. (1967). Ribonucleic acid components of BAI strain A (Myeloblastosis) avian tumour virus. *Cancer Res.* 27: 1138-1157.

Brahic M. and Vigne R. (1975). Properties of visna virus particles harvested at short time intervals: RNA content, infectivity, and ultrastructure. *J. Virol.* 15: 1222-1230.

Brown P. O., Bowerman B., Varmus H. E., and Bishop J. M. (1987). Correct integration of retroviral DNA *in vitro*. *Cell* 49: 347-356.

Campbell, S. and Vogt V. M. (1995). Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.* 69: 6487-6497.

Canaani E., Helm K. V. D., and Duesberg P. H. (1973). Evidence for 30-40S RNA as precursor of the 60-70S RNA of Rous sarcoma virus. *Proc. Natl. Acad. Sci. (USA)* 70: 401-405.

Cann A. J. (1990). Stable and safe HIV provirus clones. *Nucleic Acids Research.* 18: 6153-6154.

Cann A. J., Koyanagi Y., and Chen I. S. Y. (1988). High efficiency transfection of primary human lymphocytes and studies of gene expression. *Oncogene* 3: 123-128.

Cascone P. J., Haydar T. F., and Simon A. E. (1993). Sequences and structures required for recombination between virus-associated RNAs. *Science* 260: 801-805.

Cease K. B., and Berzofsky J. A. (1994). Toward a vaccine for AIDS: the emergence of immunobiology-based vaccine development. *Ann. Rev. Immunol.* 12: 923-89.

Chang K-Y. and Tinoco I. (1994). Characterisation of a "kissing" hairpin complex from the human immunodeficiency virus genome. *Proc. Natl. Acad. Sci. (USA)* 91: 8705-8709.

Chang S-Y. P., Bowman B. H., Weiss J. B., Garcia R. E., and White T. J. (1993). The origin of HIV-1 isolate HTLV-III<sub>B</sub>. *Nature* 363: 366-469.

Chazal N., Carriere C., Gay B., and Boulanger P. (1994). Phenotypic characterisation of insertion mutants of the human immunodeficiency virus type 1 gag precursor expressed in recombinant baculovirus-infected cells. *J. Virol.* 68: 111-122.

Chen H., Boyle T. J., Malim M. H., Cullen B. R., and Lyerly H. K. (1992). Derivation of a biologically contained replication system for human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. (USA)* 89: 7678-7682.

Cheung K-S., Smith R. E., Stone M. P., and Joklik W. K. (1972). Comparison of immature (rapid harvest) and mature Rous sarcoma virus particles. *Virology* 50: 851-864.

Clark S. J., Saag M. S., Decker W. D., Campbell-Hill S., Robertson J. L., Veldkamp P. J., Kappes J. C., Hahn B. H., and Shaw G. M. (1991). High titres of cytopathic virus in plasma of patients with symptomatic primary HIV-1 infection. *N. Engl. J. Med* 324: 954-960.

Clavel F. D., Guetard F., Brun-Vezinet S., Charmaret M-A., Rey M. O., Santos-Ferreira A. G., Laurent C., Dauguet C., Katlama C., Rouzioux C., Klatzmann D., Champalimaud J. L., and Montagnier L. (1986). Isolation of a new human retrovirus from West African patients with AIDS. *Science* 233: 343-346.

Clavel F. and Orenstein J. M. (1990). A mutant of human immunodeficiency virus with reduced RNA packaging and abnormal particle morphology. *J. Virol.* 64: 5230-5234.

Clever J., Sasseti C., and Parslow T. G. (1995). RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J. Virol.* 69: 2101-2109.

Coffin J. (1992). Genetic diversity and evolution of retroviruses. *Curr. Topics Microbiol. Immunol.* 176: 143-164.

Coffin J., Haase A., Levy J. A., Montagnier L., Oroszlan S., Teich N., Temin H., Toyoshima K., Varmus H., Vogt P., and Weiss R. (1986). Human immunodeficiency viruses. *Science* 232: 697.

Coffin J. M., Tschlis P. N., Barker C. S., and Voynow S. (1980). Variation in avian retrovirus genomes. *Ann. NY. Acad. Sci.* 354: 410-425.

Coffin J. M. (1990) in *Virology* (Fields B. N., Knope D. M., Chnock R. M. , Hirsch M. S., Meenick J. L., Monath T. P., and Roizman B., eds) 2nd Ed., pp 1437-1499. Raven Press Ltd., New York.

Conley S., Layne S., Raina J., Moore J., and Nara P. (1993). A perspective on ELISA based quantitative HIV-1 gp120 and p24 assays- current applications, limitations and potential solutions. *J. AIDS.* 6: 729.

Cullen B. R. (1994). The role of nef in the replication cycle of human and simian immunodeficiency viruses. *Virology* 205: 1-6.

Cullen B. R. (1986). Transactivation of HIV occurs via a bimodal mechanism. *Cell* 46: 973-982.

Cullen B. R. (1992). Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiological Reviews*. 56: 375-394.

Daar E. S., Moudgh T., Meyer R. D., and Ho D. D. (1991). Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection. *N. Engl. J. Med.* 324: 961-964.

Dagleish A. G., Beverly P. C. L., Clapham P. R., and Crawford D. H., Greaves M. F., Weiss R. A. (1984) The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* 312: 763-767.

Darlix J. (1986). Control of Rous sarcoma virus RNA translation and packaging by the 5' and 3' untranslated sequences. *J. Mol. Biol.* 189: 421-434.

Darlix J-L., Gabus C., and Allain B. (1992). Analytical study of avian reticuloendotheliosis virus dimeric RNA generated *in vivo* and *in vitro*. *J. Virol.* 66: 7245-7252.

Darlix J-L., Gabus C., Nugeyre M-T., Clavel F., and Barre-Sinoussi F. (1990). Cis elements and trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* 216: 689-699.

Deminie C. A. and Emerman M. (1994). Functional exchange of an oncoretrovirus and a lentivirus matrix protein. *J. Virol.* 68: 4442-4449.

DeRocquigny H., Ficheux D., Gabus C., Fournie-Zaluski M. C., Darlix J-L., and Roques B. P. (1991). First large scale chemical synthesis of the 72 amino acid HIV-1 nucleocapsid protein NCp7 in an active form. *Biochem. Biophys. Res. Commun.* 180: 1010-1018.

DeRocquigny H., Gabus C., Vincent A., Fournie-Zaluski M., Roques B., and Darlix J. (1992). Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers. *Proc. Natl. Acad. Sci. (USA)* 89: 6472-6476.

DeRocquigny H., Ficheux D., Gabus C., Allain B., Fournie-Zaluski M. C., Darlix J., and Roques B. P. (1993). Two short basic sequences surrounding the zinc finger of nucleocapsid protein NCp10 of Moloney murine leukemia virus are critical for RNA annealing activity. *Nucleic Acids Research.* 21: 823-829.

Detellier C. and Laszlo P. (1980). Role of alkali metal and ammonium cations in the self-assembly of the 5' guanosine monophosphate dianion. *J. Am. Chem. Soc.* 102: 1135-1141.

deWolf F., Roos M., Lange J. M. A., Houweling J. T. M., Coutinho R. A., Van Der Noordaa J., Schellekens P. T., and Goudsmit J. (1988). Decline in CD4+ cell numbers reflects increase in HIV-1 replication. *AIDS Res. Hum. Retrovir.* 4: 433-439.

Dube S., Kung H. J., Bender W., Davidson N., and Ostertag W. (1976). Size, subunit composition, and secondary structure of the Friend virus genome. *J. Virol.* 20: 264-272.

Duesberg P. H. (1968). Physical properties of Rous sarcoma virus. *Proc. Natl. Acad. Sci. (USA)* 60: 1511-1518.

Dupraz P., Oertle S., Meric C., Damay P., and Spahr P-F. (1990). Point mutations in the proximal cys-his box of Rous sarcoma virus nucleocapsid protein. *J. Virol.* 64: 4978-4987.

Earl P. L., Doms R. W., Moss B. (1990). Oligomeric structure of the human immunodeficiency virus type 1 envelope glycoprotein. *Proc. Natl. Acad. Sci. (USA)* 87: 648-652.

Ehrlich L. S., Fong S., Scarlata S., Zybarth G., and Carter C. (1996). Partitioning of HIV-1 gag and gag-related proteins to membranes. *Biochem.* 35: 3933-3943.

Embretson J. E. and Temin H. M. (1987). Lack of competition results in efficient packaging of heterologous murine retroviral RNAs and reticuloendotheliosis virus encapsidation-minus RNAs by the reticuloendotheliosis virus helper cell line. *J. Virol.* 61: 2675-2683.

Erickson R. (1969). Studies on the RNA of avian myeloblastosis virus. *J. Virol.* 37: 124-131.

Erickson-Viitanen S., Manfredi J. P., Viitanen P., Tribe D. E., Tritch R., Hutchinson III C. A., Loeb D. D., and Swanstrom R. (1989). Cleavage of HIV-1 gag polyprotein synthesised *in vitro*: sequential cleavage by the viral protease. *AIDS Res. Hum. Retrovir.* 5: 577-591.

Erikson E. and Erikson R. L. (1971). Association of 4S ribonucleic acid with oncornavirus ribonucleic acids. *J. Virol.* 8: 254-256.

Facke M., Janetzko A., Shoeman R. L., and Krausslich H-G. (1993). A large deletion in the matrix domain of the human immunodeficiency virus gag gene redirects virus particle assembly from the plasma membrane to the endoplasmic reticulum. *J. Virol.* 67: 4972-4980.

Fan H. and Baltimore D. (1973). RNA metabolism of Murine leukemia virus: detection of virus specific RNA sequences in infected and uninfected cells and identification of virus specific messenger RNA. *J. Mol. Biol.* 80: 93-117.

Fang G. and Cech T. R. (1993). Characterisation of a G-quartet formation reaction promoted by the B-Subunit of the *oxytricha* telomere-binding protein. *Biochem.* 32: 11646-11657.

Fang G. and Cech T. R. (1993). The B Subunit of *oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell.* 74: 875-885.

Feng Y., Broder C. C., Kennedy P. E., and Berger E. A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272: 872-877.

Feng Y-X., Fu W., Winter A. J., Levin J. G., and Rein A. (1995). Multiple regions of Harvey sarcoma virus RNA can dimerise *in vitro*. *J. Virol.* 69: 2486-2490.

Franke K. E., Yuan H. E. H., and Luban J. (1994). Specific incorporation of cyclophilin A into HIV-1 virions. *Nature.* 372: 359-362.

Freed E. O., Englung G., and Martin M. A. (1995). Role of the basic domain of human immunodeficiency virus type 1 matrix in macrophage infection. *J. Virol.* 69: 3949-3954.

Fu W. and Rein A. (1993). Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J. Virol.* 67: 5443-5449.

Fu W., Gorelick R. J. and Rein A. (1994). Characterisation of human immunodeficiency virus type 1 dimeric RNA from wild type and protease-defective virions. *J. Virol.* 68: 5013-5018.

Furth J., Seibold H. R., and Rathbone R. R. (1933). Experimental studies on lymphomatosis of mice. *Amer. J. Cancer* 19: 521-604.

Gallo R. C., Salahuddin S. Z., Popovic M., Shearer G. M., Kaplan M., Haynes B. F., Palker T. J., Redfield R., Oleske J., and Safai B. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science.* 224: 500-503.

Gehring K., Leroy J-L., and Gueron M. (1993). A tetrameric DNA structure with protonated cytosine:cytosine base pairs. *Nature.* 363: 561-565.

Geldenman H. E., Orenstein J. M., Baca L. M., Weiser B., Burger H., Katler D. C., and Meltzer M. S. (1989). The macrophage in the persistence and pathogenesis of HIV infection. *AIDS* 3: 475-495.

Gelderblom H. R. (1991). Assembly and morphology of HIV: potential effect of structure on viral function. *AIDS.* 5: 617-638.

Gellert M., Lipsett M. N., and Davies D. R. (1962). Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. (USA)* 48: 2013-2018.

Gheysen D. K., Jacobs E., deForesta F., Thiriart C., Francotte M., Thines D., and De Wilde M. (1989). Assembly and release of HIV-1 precursor Pr55<sup>gag</sup> virus-like particles from recombinant baculovirus-infected insect cells. *Cell* 59: 103-112.

Girard P-M., Bonnet-Mathoniere B., Muriaux D., and Paoletti J. (1995). A short autocomplementary sequence in the 5' leader region is responsible for dimerisation of MoMuLV genomic RNA. *Biochem.* 34: 9785-9794.

Gluzman Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* 23: 175-182.

Gonda M. A., Rice N. R., and Gilden R. V. (1980). Avian reticulo-endotheliosis virus: characterisation of the high molecular weight-viral RNA in transforming and helper virus populations. *J. Virol.* 34: 743-751.

Gorelick R. J., Henderson L. E., Hanser J. P., and Rein A. (1988). Point mutants of Moloney murine leukemia virus that fail to package viral RNA: Evidence for specific RNA recognition by a "zinc finger-like" protein sequence. *Proc. Natl. Acad. Sci. (USA)* 85: 8420-8424.

Gorelick R. J., Nigida S. M., Bess J. W., Arthur L. O., Henderson L. E. and Rein A. (1990). Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64: 3207-3211.

Gottlinger H. G., Dorfman T., Sodroski J. G. and Haseltine W. A. (1991). Effect of mutations affecting the p6 gag protein on HIV particle release. *Proc. Natl. Acad. Sci (USA)* 88: 3195-3199.

Gougeon M. L. and Montagnier L. (1993). Apoptosis in AIDS. *Science.* 260: 1269-1270.

Grew C., Beck A., and Gelderblom H. R. (1990). HIV: early virus-cell interactions. *J. AIDS* 3: 965-974.

Groenink M., Fouchier R. A. M., de Groede R. E. Y., de Wolf F., Gruters R. A., Cuypers H. T. M., Huisman H. G., and Tersmette M. (1991). Phenotypic heterogeneity in a panel of infectious molecular human immunodeficiency virus type 1 clones derived from a single individual. *J. Virol.* 65: 1968-1975.

Guo Q., Lu M., and Kallenbach N. R. (1992). Adenine affects the structure and stability of telomeric sequences. *J. Biol Chem.* 267: 15293-15300.

Guschlbauer W., Chantot J., and Thiele D. (1990). Four-Stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA. *J. Biomolecular Structure & Dynamics*. 8: 491-511.

Haddrick M., Lear A., Cann A. J., and Heaphy S. (1996a). Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J. Mol. Biol.* 259: 58-68.

Haddrick M., Liu Z-H., Lau A., Heaphy S., and Cann A. J. (1996b) Morphogenesis and RNA packaging in recombinant HIV VLPs. *Meth. Virol.*, in press.

Hahn B. H., Shaw G. M., Taylor M. E., Redfield R. R., Markham P. D., Salahuddin S. Z., Wong-Staal F., Gallo R. C., and Parks W. P. (1986). Genetic variation in HTLVIII/LAV over time in patients with AIDS or at risk for AIDS. *Science* 232: 1548-1553.

Harrison G. and Lever A. M. (1992). The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. *J. Virol.* 66: 4144-4153.

Harrison G. P., Hunter E., and Lever A. M. (1995). Secondary structure model of the Mason-Pfizer monkey virus 5' leader sequence: identification of a structural motif common to a variety of retroviruses. *J. Virol.* 69: 2175-2186.

Hayashi T., Ueno Y., and Okamoto T. (1993). Elucidation of a conserved RNA stem-loop structure in the packaging signal of HIV-1. *FEBS Lett.* 327: 213-218.

Hayashi T., Shioda T., Iwakura Y., and Shibuta H. (1992). RNA packaging signal of human immunodeficiency virus type 1. *Virol.* 188: 590-599.

Heaphy S., Finch J. T., Gait M. J., Karn J. and Singh M. (1991). Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich bubble located within the rev-responsive region of viral messenger RNAs. *Proc. Natl. Acad. Sci. (USA)* 88: 7366-7370.

Henderson L. E., Copeland T. D., Sowder R. C., Smythers G. W., and Oroszlan S. (1981). Primary structure of the low-molecular weight nucleic acid binding proteins of murine leukemia viruses. *J. Biol. Chem.* 256: 8400-8406.

Ho D. D., Neumann A. U., Perelson A. S., Chen W., Leonard J. M., and Markowitz M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373: 123-126.

Hockley D. J., Nermut M. V., Grief C., Jowett J. B. M., and Jones I. M. (1994). Comparative morphology of gag protein structures produced by mutants of the gag gene of human immunodeficiency virus type 1. *J. Gen. Virol.* 75: 2985-2997.

Hoglund S., Goncalves J., Ohagen A., and Gabuzda D. (1995). Ultrastructure of HIV-1 Genomic RNA. *AIDS Res. Hum. Retr.* 11 S1 S105 (abstract).

Hoglund, S., A. Ohagen, K. Lawrence and D. Gabuzda. (1994). Role of *vif* during Packing of the Core of HIV-1. *Virol.* 201: 349-355.

Hong S. S., and Boulanger P. (1993). Assembly-Defective point mutants of the human immunodeficiency virus type 1 gag precursor phenotypically expressed in recombinant baculovirus-infected cells. *J. Virol.* 67: 2787-2798.

Hu S. L., Travis B. M., Garrigues J., Zarling J. M., Sridhar P., Dykers T., Eichberg J. W., and Alpers C. (1990). Processing, assembly and immunogenicity of HIV core antigens expressed by recombinant vaccinia virus. *Virol.* 179: 321-329.

Hu W. and Temin H. M. (1990). Retroviral recombination and reverse transcription. *Science.* 250: 1227-1233.

Hunter E. (1994). Macromolecular interactions in the assembly of HIV and other retroviruses. *Sem. Virol.* 5: 71-83.

Jones J. S., Allan R. W., and Temin H. M. (1993). Alteration of location of dimer linkage sequence in retroviral RNA: little effect on replication or homologous recombination. *J. Virol.* 67: 3151-3158.

Jowett J. B. M., Hockley D. J., Nermut M. V., and Jones I. (1992). Distinct signals in human immunodeficiency virus type 1 Pr55 necessary for RNA binding and particle formation. *J. Gen. Virol.* 73: 3079-3086.

Kang C., Zhang X., Ratliff R., Moyzis R., and Rich A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature.* 356: 126-131.

Kaplan A. H., Zack J., Knigge A. M., Paul D. A., Kempf D. J., Norbeck D. W., and Swanstrom R. (1993). Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. *J. Virol.* 67: 4050-4055.

Karacostas V., Nagashima K., Gonda M. A., and Moss B. (1989). Human immunodeficiency virus-like particles produced by a vaccinia virus expression vector. *Proc. Natl. Acad. Sci. (USA)* 86: 8964-8967.

Karacostas V., Wolffe E. J., Nagashima K., Gonda M. A., and Moss B. (1993). Overexpression of the HIV-1 Gag-Pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology* 193: 661-679.

Karcewzski M.K. and Strebel K. (1996). Cytoskeletal association and virion incorporation of HIV-1 vif protein. *J. Virol.* 70: 494-507.

Kastelein R. A., Remaut E., Fiers W., and van Duin J. (1982). Lysis gene expression of RNA phage MS2 depends on a frameshift during translation of the overlapping coat protein gene. *Nature* 295: 35-41.

Katoh I., Kyushiki H., Sakamoto Y., Ikawa Y., and Yoshinaka Y. (1991). Bovine leukemia virus matrix associated protein MA(p15): Further processing and formation of a specific complex with the dimer of the 5'-terminal genomic RNA fragment. *J. Virol.* 65: 6845-6855.

Katoh I., Yasunaga T., and Yoshinaka Y. (1993). Bovine leukemia virus RNA sequences involved in dimerisation and specific gag protein binding: close relation to the packaging sites of avian, murine, and human retroviruses. *J. Virol.* 67: 1830-1839.

Katz R. A., Terry R. W., and Skalka A. M. (1986). A conserved cis-acting sequence in the 5' leader of Avian sarcoma virus RNA is required for packaging. *J. Virol.* 163-167.

Kaye J. F. and Lever A. M. (1996). trans-Acting proteins involved in RNA encapsidation and viral assembly in human immunodeficiency virus type 1. *J. Virol.* 70: 880-886.

Kellam P., Bouchers C. A. B., and Larder B. A. (1992). Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to the development of high-level resistance to zidovudine. *Proc. Natl. Acad. Sci. (USA)* 89: 1934-1938.

Kestler H. W., Ringler D. J., Mori K., Panicali D. L., Sehgal P. K., Daniel M. D., and Desrosiers R. C. (1991). Importance of the nef gene for maintenance of high virus load and for development of AIDS. *Cell* 65: 651-662.

Kewalaramani V. N., Park C. S., Gallombardo R. A., and Emerman M. (1996). Protein stability influences HIV-2 vpr virion incorporation and cell cycle effect. *Virology* 218: 326-334.

Khan R. and Giedroc D. P. (1991). Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. *J. Biol. Chem.* 267: 6689-6695.

Kim J., Cheong C., and Moore P. B. (1991). Tetramerisation of an RNA oligonucleotide containing a GGGG sequence. *Nature* 351: 331-332.

Kishi M., Zheng Y-H., Bahmani M., Tokunaga K. K., Takahashi H., Kakinuma M., Lai P., Nonoyama K. M., Luftig R. B., and Ikuta K. (1995). Naturally occurring accessory gene mutations lead to persistent human immunodeficiency virus type 1 infection of CD4-positive T cells. *J. Virol.* 69: 7507-7518.

Klatzmann D., Champagne E., Charmaret S., Gurest J., Guetard D., Hercend T., Gluckman J. C., and Montagnier L. The lymphocyte T4 molecule behaves as a receptor for human retrovirus LAV. (1984) *Nature* 312: 767-778.

Klikova M., Rhee S. S., Hunter E., and Ruml T. (1995). Efficient *in vivo* and *in vitro* assembly of retroviral capsids from gag precursor proteins expressed in bacteria. *J. Virol.* 69: 1093-1098.

Klimkait T., Strebel K., Hoggan M. D., Martin M. A., and Orenstein J. M. (1990). The HIV-1 specific protein vpu is required for efficient virus maturation and release. *J. Virol.* 64: 621-629.

Knight J. B., Si Z. H., and Stoltzfus C. M. (1994). A base paired structure in the avian sarcoma virus 5' leader is required for efficient encapsidation of RNA. *J. Virol.* 68: 4493-4502.

Kondo E., Mammano F., Cohen E. A., and Gottlinger H. G. (1995). The p6 gag domain of human immunodeficiency virus type 1 is sufficient for the incorporation of vpr into heterologous viral particles. *J. Virol.* 69: 2759-2764.

Koyanagi Y., Miles S., Mitsuyasin R. T., Merrill J. E., Vintners H. V., and Chen I. S. Y. (1987). Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms. *Science* 236: 819-822.

Kozak M. (1989). The scanning model for translation: an update. *J. Cell. Biol.* 108: 229-241.

Kozarsky K., Penman N., Basiripour L., Haseltine W., and Sodroski J. (1989). Glycosylation and processing of the HIV-1 envelope protein. *J. AIDS* 2: 163-169.

Krausslich H. G., Ochsenbauer C., Traenckner A. M., Mergener K., Facke M., Gelderblom H. R., and Bosch V. (1993). Analysis of protein expression and virus-like particle formation in mammalian-cell lines stably expressing HIV-1 gag and env gene-products with or without active HIV proteinase. *J. Virol.* 192: 605-617.

Kung H., Bailey J. M., Davidson N., Nicholson M. O., and McAllister R. M. (1975). Structure, subunit composition, and molecular weight of RD-114 RNA. *J. Virol.* 16: 397-411.

Kung H., Hu S., Bender W., Bailey J. M., and Davidson N. (1976). RD-114, baboon, and woolly monkey viral RNAs compared in size and structure. *Cell*. 7: 609-620.

Laughrea M. and Jette L. (1994). A 19-nucleotide sequence upstream of the 5' major splice donor is part of the dimerisation domain of human immunodeficiency virus type 1 genomic RNA. *Biochem*. 33: 13464-13474.

Laughrea M. and Jette L. (1996). Kissing loop model of HIV-1 genome dimerisation: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248-271 are dispensable for dimer formation. *Biochem*. 35: 1589-1598.

Lear A. L., Haddrick M., and Heaphy S. (1995). A Study of the dimerisation of Rous sarcoma virus RNA *in vitro* and *in vivo*. *Virology*. 212: 47-57.

Lee T. H., Coligan J. E., Sodroski J. G., Haseltine W. A., Salahuddin S. Z., Wongstaal F., Gallo R. C., and Essex M. (1984). Antigens encoded by the 3'-terminal region of Human T-cell leukemia virus - evidence for a functional gene. *Science* 226: 57-61.

Lee P. P. and Linial M. (1994). Efficient particle formation can occur if the matrix domain of human immunodeficiency virus type 1 gag is substituted by a myristylation signal. *J. Virol*. 68: 6644-6654.

Leffers H., Kjems J., Ostergaard L., Larsen N., and Garrett R. (1987). A comparative study of 23S ribosomal RNA of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. *J. Mol. Biol.* 195: 43-61.

LeGrice S. and Leitch G. (1990). Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. *Euro J. Biochem*. 187: 307-314.

Lever A., Gottlinger H., Haseltine W., and Sodroski J. (1989). Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. *J. Virol*. 63: 4085-4087.

Levin J. G., Grimley P. M., Ramseur J. M., and Berezsky I. K. (1974). Deficiency of 60 to 70S RNA in murine leukemia virus particles assembled in cells treated with actinomycin D. *J. Virol*. 14: 152-161.

Levy J. A. (1994). HIV and the pathogenesis of AIDS. ASM Press, ASM Washington.

Levy J. A., Hoffman A. D., Kramer S. M., Landis J. A., Shimabukuro J. M., and Oschiro L. S. (1984). Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* 225: 840-842.

Linial M. L. and Miller A. D. (1990). Retroviral RNA packaging: sequence requirements and implications. *Curr. Top. Microbiol. Immunol.* 157: 124-152.

Linial M., Mederios E., and Hayward W. (1978). An avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: biological and biochemical characterization. *Cell* 15: 1371-1381.

Lipanov A. A., Quintana J., and Dickerson R. E. (1990). Disordered single crystal evidence for a quadruple helix formed by guanosine 5'-monophosphate. *J. Biomolecular Structure & Dynamics* 8: 483-489.

Littman D. ed. (1995). The CD4 molecule. *Curr. Topics Microbiol Immunol.* 205.

Liu Z. P., Frantz J. D., Gilbert W., and Tye B. K. (1993). Identification and characterisation of a nuclease activity specific for G4 tetrastranded DNA. *Proc. Natl. Acad. Sci. (USA)* 90: 3157-3161.

Liu H., Wu X., Newman M., Shaw G. M., Hahn B. H., and Kappes J. C. (1995). The *vif* protein of human and simian immunodeficiency viruses is packaged into virions and associates with viral core structures. *J. Virol.* 69: 7630-7638.

Lori F., Vernose F., De Vico A. L., Lusso P., Reitz M. S., and Gallo R. C. (1992). Viral DNA carried by human immunodeficiency virus type 1 virions. *J. Virol.* 66: 5067-5074.

Luban J. and Goff S. P. (1994). Mutational analysis of cis-acting packaging signals in human immunodeficiency virus type 1 RNA. *J. Virol.* 68: 3784-3793.

Luborsky S. W. (1971). Sedimentation equilibrium analysis of the molecular weight of a tumour virus RNA. *Virol.* 45: 782-787.

Luo L., Li Y., and Kang Y. (1990). Expression of gag precursor protein and secretion of virus like particles of HIV-1 from recombinant baculovirus infected insect cells. *Virol.* 179: 874-880.

Luo G. and Taylor J. (1990). Template switching by reverse transcriptase during DNA synthesis. *J. Virol.* 64: 4321-4328.

Lustig A. J. (1992). Hoogsteen G-G base pairing is dispensable for telomere healing in yeast. *Nucleic Acids Research.* 20: 3021-3028.

Maddon P. J., Dalgleish A. G., McDougal J. S., Clapham P. R., Weiss R. A., and Axel R. (1986). The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 47: 333-348.

Maisel J., Bender W., Hu S., Duesberg P. H., and Davidson N. (1978). Structure of 50 to 70S RNA from Moloney sarcoma viruses. *J. Virol.* 25: 384-394.

Mangel W. F., Delius H., and Duesberg P. H. (1974). Structure and molecular weight of the 60-70S RNA and the 30-40S RNA of the Rous sarcoma virus. *Proc. Natl. Acad. Sci. (USA)* 71: 4541-4545.

Mansky L. M., Krueger A. E., and Temin H. M. (1995). The bovine leukemia virus encapsidation signal is discontinuous and extends into the 5' end of the *gag* gene. *J. Virol.* 69: 3282-3289.

Marino J. P., Gregorian R. S., Csankovszki G., and Crothers D. M. (1995). Bent helix formation between RNA hairpins with complementary loops. *Science* 268: 1448-1454.

Marquet R., Baudin F., Gabus C., Darlix J., Mougel M., Ehresmann C., and Ehresmann B. (1991). Dimerisation of human immunodeficiency virus type 1 RNA: stimulation by cations and possible mechanism. *Nucleic Acids Research* 19: 2349-2357.

Marquet R., Paillart J.-C., Skripkin E., Ehresmann C., and Ehresmann B. (1994). Dimerisation of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Research* 22: 145-151.

McBride M. S. and Panganiban A. T. (1996). The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. *J. Virol.* 70: 2963-2973.

McNicholl J. M. and McDougal J. S. Immunologic escape as a mechanism of viral persistence in HIV-1 infection. *Sem. Virol.* (1994). 5: 307-317.

Mergener K., Facke M., Welker R., Brinkman V., Gelderblom H., and Krausslich H. G. (1992). Analysis of HIV particle formation using transient expression of sub-viral constructs in mammalian cells. *Virol.* 186: 25-39.

Meric C. and Spahr P. (1986). Rous sarcoma virus nucleic acid-binding p12 is necessary for viral 70S RNA dimer formation and packaging. *J. Virol.* 60: 450-459.

Meric C. and Goff S. P. (1989). Characterisation of Moloney murine leukemia virus mutants with single amino-acid substitutions in the cys-his box of the nucleocapsid protein. *J. Virol.* 63: 1558-1568.

Messer L. I., Levin J. G., and Chattopadhyay S. (1981). Metabolism of viral RNA in murine leukemia virus-infected cells: evidence for differential stability of viral message and virion precursor RNA. *J. Virol.* 40: 683-690.

Messing J. (1983). New M13 vectors for cloning. *Meth. Enzymol.* 101: 20-78.

Miele G., Mouland A., Harrison G. P., Cohen E., and Lever A. M. (1996). The human immunodeficiency virus type 1 5' packaging signal structure affects translation but does not function as an internal ribosome entry site structure. *J. Virol.* 70: 944-951.

Montagnier L., Golde A., and Vigier P. (1969) A possible subunit structure of RSV RNA. *J. Gen Virol.* 4: 449-452.

Montagnier L., Chermann J., Barre-Sinoussi F., Chamaret S., Gruest J., Nugeyre M. T., Rey F., Dautet C., Axler-Blin C., Vezinet-Brun F., Rouzioux C., Saimot A. G., Rozenbaum W., Gluckman J. C., Klatzmann D., Vilmer E., Griselli C., Gazengel C., and Brunet J. B. (1984). A new human T-lymphotropic retrovirus: Characterization and possible role in lymphadenopathy and acquired immune deficiency syndromes. p. 363-379 *In.* Gallo R. C., Essex M. E., Gross L. (ed.), *Human T-cell leukemia/lymphoma virus.* Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.

Mougel M., Tounekti N., Darlix J-L., Paoletti J., Ehresmann B., and Ehresmann C. (1993). Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. *Nucleic Acids Research.* 21: 4677-4684.

Muriaux D., Fosse P., and Paoletti J. (1996). A kissing complex together with a stable dimer is involved in the HIV-1<sub>LAI</sub> RNA dimerisation process *in vitro*. *Biochem.* 35: 5075-5082.

Muriaux D., Girard P-M., Bonnet-Mathoniere B., and Paoletti J. (1995). Dimerisation of HIV-1<sub>LAI</sub> RNA at low ionic strength. *J. Biol. Chem.* 270: 8209-8216.

Murti K. G., Bondurant M., and Tereba A. (1981). Secondary structural features in the 70S RNAs of Moloney murine leukemia and Rous sarcoma viruses as observed by electron microscopy. *J. Virol.* 37: 411-419.

Nabel G. and Baltimore D. (1987). An inducible transcription factor activates expression of HIV in T-cells. *Nature* 326: 711-713.

Nermut M. V., Hockley D. J., Jowett J. B. M., Jones I. M., Garreau M., and Thomas D. (1994). Fullerene-like organization of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. *Virol.* 198: 288-296.

Niedrig M., Gelderblom H. R., Pauli G., Marz J., Bickhard H., Wolf H., and Modrow S. (1994). Inhibition of infectious human immunodeficiency

virus type 1 particle formation by gag protein-derived peptides. J. Gen. Virol. 75: 1469-1474.

Oertle S. and Spahr P. (1990). Role of the *gag* polyprotein precursor in packaging and maturation of Rous sarcoma virus genomic RNA. J. Virol. 64: 5757-5763.

Ogawa K., Shibata R., Kiyomasu T., Higuchi I., Kishida Y., Ishimoto A., and Adachi A. (1989) Mutational analysis of the human immunodeficiency virus *vpr* open reading frame. J. Virol. 63: 4110-4114.

Oka Y. and Thomas C. A. J. (1987). The cohering telomeres of *Oxytricha*. Nucleic Acids Research. 15: 8877-8898.

Overton H. A. , Fujii Y., Price I. R., and Jones I. M. (1989) The protease and gag gene products of human immunodeficiency virus: authentic cleavage and post translational modification in an insect expression system. Virol. 170: 107-116.

Pager J. (1993). A liquid chromatographic preparation of retroviral RNA. Anal. Biochem. 215: 231-235.

Paillart J.-C., Marquet R., Skripkin E., Ehresmann B., and Ehresmann C. (1994). Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J. Biol Chem. 269: 27486-27493.

Panganiban A. and Fiore D. (1988). Ordered interstrand and intrastrand DNA transfer during reverse transcription. Science. 241: 1064-1069

Pantaleo G., Graziosi C., Demarest J. F., Butini L., Montroni M., Fox C. H., Orenstein J. M., Kotler D. P., and Fauci A. S. (1993). HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. Nature 362: 355-358.

Pantaleo G., Graziosi C., Butini L., Pizzo P. A., Schnittman S. M., Kotler D. P., and Fauci A. S. (1991). Lymphoid organs function as major reservoirs for human immunodeficiency virus. Proc. Natl. Acad. Sci. (USA) 88: 9838-9842.

Park J. and Morrow C. D. (1991). Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. J. Virol. 65: 5111-5117.

Patterson S. and Knight S. C. (1987). Susceptibility of human peripheral-blood dendritic cells to infection by HIV. J. Gen. Virol. 68: 1177-1181.

Peattie D. A. and Gilbert W. (1980). Chemical probes for higher-order structure in RNA. *Proc.Natl.Acad.Sci., (USA)* 77: 4679-4682.

Peliska J. A. and Benkovic S. J. (1992). Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science* 258: 1112-1118.

Persson C., Wagner E. G. H., and Nordstrom K. (1990). Control of replication of plasmid R1: structures and sequences of the antisense RNA, CopA, required for its binding to the target RNA, CopT. *EMBO J.* 9: 3767-3775.

Pinnavaia T. J., Marshall C. L., Mettler C. M., Fisk C. L., Miles H. T., and Becker E. D. (1978). Alkali metal ion specificity in the solution ordering of a nucleotide, 5'-guanosine monophosphate. *J. Am. Chem. Soc.* 100: 3625-3627.

Platt E. J. and Haffar O. K. (1994). Characterisation of human immunodeficiency virus type 1 Pr55<sup>gag</sup> membrane association in a cell-free system: Requirement for a C-terminal domain. *Proc.Natl.Acad.Sci. (U.S.A)* 91: 4594-4598.

Poblotzki A., Wagner R., Niedrig M., Wanner G., Wolf H., and Modrow S. (1993). Identification of a region in the Pr55<sup>gag</sup>-polyprotein essential for HIV-1 particle formation. *Virology* 193: 981-985.

Poiesz B. J., Ruscetti F. W., Reitz M. S., Kalyanaraman V. S., and Gallo R. C. Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sezary T-cell leukemia. *Nature* (1981) 294: 268-271.

Popovic M., Sarngadharan M. G., Read E., and Gallo R. C. (1984). Detection, isolation and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*: 224: 497-500.

Prats A. C., Sarik L., Gabus C., Litvak S., Keith G., and Darlix J-L. (1988). Small finger protein of avian and murine retroviruses has nucleic acid annealing activity and postions the replication primer tRNA onto genomic RNA. *EMBO J.* 7: 1777-1783.

Prats A. C., Roy P., Wang M., Erard V., Housset C., Gabus C., Paoletti C., and Darlix J. (1990). cis Elements and trans -acting factors involved in dimer formation of murine leukemia virus RNA. *J. Virol.* 64: 774-783.

Raghuraman M. K. and Cech T. R. (1989). Assembly and self-association of *Oxytricha* telomeric nucleoprotein complexes. *Cell* 59: 719-728.

Ratner L. and Niederman T. M. J. (1995). Nef. *Curr. Topics in Microbiol. and Immunol.* 193: 169-208.

Ray F. A. and Nickloff J. A. (1992). Site-specific mutagenesis of almost any plasmid using a PCR-based version of unique site elimination. *BioTechniques*. 13: 342-348.

Reicin A. S., Paik S., Berkowitz R. D., Luban J., Lowy I., and Goff S. P. (1995). Linker insertion mutations in the human immunodeficiency virus type 1 gag gene: Effects on virion particle assembly, release, and infectivity. *J. Virol.* 69: 642-650.

Rhee S. S. and Hunter E. (1990). A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a C-type retrovirus. *Cell* 63: 77-86.

Rhodes D. and Klug A. (1993). Zinc fingers. *Sci. Am.* Feb. 32-39.

Rizvi T. A. and Panganiban A. T. (1993). Simian immunodeficiency virus RNA is efficiently encapsidated by human immunodeficiency virus type 1 particles. *J. Virol.* 67: 2681-2688.

Robinson W. S., Pitkanen A., and Rubin H. (1965). The nucleic acid of the Bryan strain of Rous sarcoma virus: purification of the virus and isolation of the nucleic acid. *Proc. Natl. Acad. Sci. (USA)* 54: 137-144.

Rohdewohld H., Weiher H., Reik W., Jaenisch R., and Breindl M. (1987). Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase I hypersensitive sites. *J. Virol.* 61: 336-343.

Rous P. (1911) A sarcoma of the fowl transmissible by an agent separable from the tumour cells. *J. Exp. Med.* 13: 397.

Roy C., Tounekti N., Mougél M., Darlix J., Paoletti C., Ehresmann C., Ehresmann B., and Paoletti J. (1990). An analytical study of the dimerization of *in vitro* generated RNA of Moloney murine leukemia virus MoMuLV. *Nucleic Acids Research*. 18: 7287-7292.

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

Sakaguchi K., Zambrano N., Baldwin E. T., Shapiro B. A., Erickson J. W., Omichinski J. G., Clore G. M., Gronenborn A. M., and Appella E. (1993). Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. *Proc. Natl. Acad. Sci. (USA)* 90: 5219-5223.

Sakalian M., Wills J. W., and Vogt V. M. (1994). Efficiency and selectivity of RNA packaging by Rous sarcoma virus gag deletion mutants. *J. Virol.* 68: 5969-5981.

Salahuddin S. Z., Markham P. D., Wongstaal F., Franchini G., Kalyanaraman V. S., and Gallo R. C. (1983). Restricted expression of Human T-cell leukemia lymphoma virus (HTLV) in transformed human umbilical cord blood lymphocytes. *Virology* 129: 51-64.

Sambrook J., Fritsch E. F., and Maniatis T. (1989). *Molecular cloning - a laboratory manual*. Cold Spring Harbour Laboratory Press. Cold Spring Harbour, N.Y.

Sanger F., Nicklen S., and Coulson A. R. (1977) DNA-sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. (USA)* 74: 5463-5467.

Scaria P. V., Shire S. J., and Shafer R. H. (1992). Quadruplex structure of d(G<sub>3</sub>T<sub>4</sub>G<sub>3</sub>) stabilized by K<sup>+</sup> or Na<sup>+</sup> is an asymmetric hairpin dimer. *Proc. Natl. Acad. Sci. (USA)* 89: 10336-10340.

Schlesinger S., Makino S., and Linial M. L. (1994). Cis-acting genomic elements and trans-acting proteins involved in the assembly of RNA viruses. *Semin. Virol.* 5: 39-49.

Sen D. and Gilbert W. (1988). Formation of parallel 4-stranded complexes by guanine rich motifs in DNA and its implications for meiosis. *Nature* 334: 364-366.

Sen D. and Gilbert W. (1990). A sodium-potassium switch in the formation of 4-stranded G4-DNA. *Nature* 344: 410-414.

Sen D. and Gilbert W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334: 364-366.

Sen D. and Gilbert W. (1991). Novel DNA superstructures formed by telomere-like oligomers. *Biochem.* 31: 65-70.

Shank P. R. and Linial M. (1980). Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: characterisation of a deletion in the provirus. *Virology* 36: 450-456.

Sheng N. and Erickson-Viitanen S. (1994). Cleavage of p15 protein *in vitro* by human immunodeficiency virus type 1 protease is RNA dependent. *J. Virol.* 68: 6207-6214.

Shioda T. and Shibuta H. (1990). Production of human immunodeficiency virus (HIV)-like particles from cells infected with recombinant vaccinia viruses carrying the gag gene of HIV. *Virology* 175: 139-148.

Simons R.W. and Kleckner N. (1988). Biological regulation by antisense RNA in prokaryotes. *Ann. Rev. Genet.* 22: 567-600.

Skalka A. M., Boone L., Junghans R., and Luk D. (1982). Genetic recombination in avian retroviruses. *J Cell Biochem.* 19: 293-304.

Skripkin E., Paillart J.-C., Marquet R., Ehresmann E., and Ehresmann B. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerisation *in vitro*. *Proc. Natl. Acad. Sci. (USA)* 91: 4945-4949.

Smith A. J., Cho M.-I., Hammarskjöld M.-L., and Rekosh D. (1990). Human immunodeficiency virus type 1 Pr55gag and Pr160gagpol expressed from a simian virus 40 late replacement vector are efficiently processed and assembled into virus like particles. *J. Virol.* 64: 2743-2750.

Smith F. W. and Feigon J. (1992). Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature* 356: 164-168.

Spearman P., Wang J. J., Vander Heyden N., and Ratner L. (1994). Identification of human immunodeficiency virus type 1 gag protein domains essential to membrane binding and particle assembly. *J. Virol.* 68: 3232-3242.

Stanley S. K. and Fauci A. S. (1993). T-cell homeostasis in HIV infection: Part of the solution, or part of the problem? *J. AIDS.* 6: 142-143.

Stewart L., Schatz G., and Vogt V. M. (1990). Properties of avian retrovirus particles defective in viral protease. *J. Virol.* 64: 5076-5092.

Stoltzfus C. M. and Snyder P. N. (1975). Structure of B77 sarcoma virus RNA: Stabilisation of RNA after packaging. *J. Virol.* 16: 1161-1170.

Sundquist W. I. (1991). The Structures of Telomeric DNA, In Eckstein F., (ed). *Nucleic Acids and Molecular Biology Vol 5.* Springer-Verlag Press, Berlin, Germany.

Sundquist W. I. and Klug A. (1989). Telomeric DNA dimerises by formation of guanine tetrads between hairpin loops. *Nature* 342: 825-829.

Sundquist W. I. and Heaphy S. (1993). Evidence for interstrand guanine base tetrad formation in the dimerisation of HIV-1 genomic RNA. *Proc. Natl. Acad. Sci (USA)* 90: 3393-3397.

Tchenio T. and Heidmann T. (1995). The dimerization/packaging sequence is dispensable for both the formation of high-molecular-weight RNA complexes within retroviral particles and the synthesis of proviruses of normal structure. *J. Virol.* 69: 1079-1084.

Temin H. (1991). Sex and recombination in retroviruses. *Trends in Genetics.* 7: 71-74.

Temin H. M. and Mizutani S. (1970). Viral-RNA dependent DNA Polymerase. *Nature* 226: 1211-1213.

Thali M., Bukovsky A., Kondo E., Rosenwirth B., Walsh C. T., Sodroski J., and Gottlinger H. G. (1994). Functional association of cyclophilin A with HIV-1 virions. *Nature* 372: 363-365.

Torrent C., Gabus C., and Darlix J-L. (1994). A small and efficient dimerization/packaging signal of rat VL30 RNA and its use in murine leukemia virus-VL30 derived vectors for gene transfer. *J. Virol.* 68: 661-667.

Torrent C., Bordet T., and Darlix J-L. (1994). Analytical study of rat retrotransposon VL30 RNA dimerization *in vitro* and packaging in murine leukemia virus. *J. Mol. Biol.* 240: 434-444.

Tounekti N., Mougél M., Roy C., Marquet R., Darlix J., Paoletti J., Ehresmann B., and Ehresmann C. (1992). Effect of dimerization on the conformation of the encapsidation psi domain of moloney murine leukemia virus RNA. *J. Mol. Biol.* 223: 205-220.

Wain-Hobson S. (1992). Human immunodeficiency virus type 1 quaspecies *in vivo* and *ex vivo*. *Curr. Topics in Micro. Immunol.* 176: 181-193.

Wain-Hobson S. (1995). Virological Mayhem. *Nature* 373: 102.

Walsh K. and Gualberto A. (1992). MyoD binds to the guanine tetrad nucleic acid structure. *J. Biol Chem.* 267: 13714-13718.

Wei X., Ghosh S. K., Taylor M., Johnson V. A., Emini E. A., Deutsch P., Lifson J. D., Bonhoeffer S., Nowak M. A., Hahn B. H., Saag M. S., and Shaw G. M. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373: 117-122.

Weiss R. A. (1996). Foamy viruses bubble on. *Nature* 380: 201.

Weiss R. A., Mason W. S., and Vogt P. K. (1973). Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumour viruses. *Virology* 52: 535-552.

Weiss R. (1993). How does HIV cause AIDS. *Science* 260: 1273-1279.

Weiss S., König B., Morikawa Y., and Jones I. (1992). Recombinant HIV-1 nucleocapsid protein p15 produced as a fusion protein with glutathione S-transferase in *Escherichia coli* mediates dimerization and enhances reverse transcription of retroviral RNA. *Gene* 121: 203-212.

Weiss S., Hausl G., Famulok M., and Konig B. (1993). The multimerization state of retroviral RNA is modulated by ammonium ions and affects HIV-1 full-length cDNA synthesis *in vitro*. *Nucleic Acids Research*. 21: 4879-4885.

Whittle H., Egboga A., Todd J., Corrah T., Wilkins A., Demba E., Morgan G., Rolfe M., Berry N., and Tedder R. (1992). Clinical and laboratory predictors of survival in Gambian patients with symptomatic HIV-1 and HIV-2 infection. *AIDS* 6: 685-689.

Wiley R. L., Bonifacino J. S., Potts B. J., Martin M. A., and Klausner R. D. (1988). Biosynthesis, cleavage and degradation of the HIV-1 envelope glycoprotein gp 160. *Proc. Natl. Acad. Sci. (USA)* 85: 9580-9584.

Williams K. J. and Loeb L.A. (1992). Retroviral reverse transcriptases: Error frequencies and mutagenesis. *Curr. Topics in Micr & Immunol*. 176: 165-180.

Williamson J. R., Raghuraman M. K., and Cech T. R. (1989). Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell*. 59: 871-880.

Wills J. W. and Craven R. C. (1991). Form, function and use of retroviral Gag proteins. *AIDS*. 5: 639-654.

Wilson W., Braddock M., Adams S. E., Rathjen P. D., Kingsman S. M., and Kingsman A. J. (1988). HIV expression strategies - ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. *Cell* 55: 1159-1169.

Wu W., Blumberg B. M., Fay P. J., and Bambara R. A. (1995). Strand transfer mediated by human immunodeficiency virus reverse transcriptase *in vitro* is promoted by pausing and results in misincorporation. *J. Biol. Chem*. 270: 325-332.

Yamada K., Morozumi H., and Okamoto T. (1995). LTR-directed homologous recombination of full-length HIV-1 provirus clone in *recA(-)* bacteria. *Arch. Virol*. 140: 1007-1014.

Yu X., Yuan X., Matsuda Z., Lee T., and Essex M. (1992). The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J. Virol*. 66: 4966-4971.

Zahler A. M., Williamson J. R., Cech T. R., and Prescott D. M. (1991). Inhibition of telomerase by G-quartet DNA structures. *Nature* 350: 718-720.

Zhang H., Zhang Y., Spicer T., Henrard D., and Poesz B. J. (1995). Nascent human immunodeficiency virus type 1 reverse transcription occurs within an enveloped particle. *J. Virol*. 69: 3675-3682.

Zhang Y. and Barklis E. (1995). Nucleocapsid protein effects on the specificity of retrovirus RNA encapsidation. *J. Virol.* 69: 5716-5722.