Localisation of the RNA-binding domain on the Respiratory Syncytial Virus Nucleocapsid protein

Lindsay Bryony Murphy

MRC Virology Unit

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

Institute of Virology
University of Glasgow
Church Street
Glasgow
G11 5JR

December, 2001

ProQuest Number: 13818489

![ProQuest logo]

ProQuest 13818489

Summary

The purpose of the study presented herein was to investigate the RNA-binding properties of the nucleocapsid (N) protein of respiratory syncytial virus (RSV). The position of the RNA binding-site within the protein was investigated. In addition, the ability of N protein to assemble into nucleocapsid-like structures was studied, thus determining whether these facets were distinguishable.

The sequence specificity of RNA-binding by the N protein was investigated using several RNA probes, one of which represented the RSV leader sequence. Various *in vitro* assays, such as the yeast three-hybrid system, gel mobility shift assay and northwestern blots, were employed to this end. In the absence of other viral proteins, N possessed no sequence preference in the manner of its association with RNA.

The locale of the RNA-binding site within the N protein was addressed by employing two methodologies. The first approach employed a UV crosslinking assay with subsequent analysis by mass spectrometry. A single N-derived peptide was identified within the carboxy-domain of N (amino acids 352 to 358) as being potentially modified by RNA-binding. Secondly, the use of carboxy-terminal deletion mutants of the N protein demonstrated the presence of an RNA-binding domain within the first 92 amino acids of the N protein. The two individual binding regions identified demonstrate the conformational nature of the RNA-binding domain. However, the role of the carboxy terminus was shown to be minimal, as deletion of this region did not prevent RNA-binding.

Concurrent with the binding of RNA is the formation of nucleocapsids, helical structures formed from N proteins encapsulating the viral genomic RNA. The ability of recombinant N to assemble into such structures was investigated by EM. Histidine-tag N demonstrated the capacity to form nucleocapsid-like structures and N:RNA rings, the latter representing a single turn of the nucleocapsid helix. The rings were shown to consist of 10 N protein monomers, a situation different from Sendai virus (13 monomers) and more similar to that of rabies virus (10 monomers). A mutant representing amino acids 1 to 92 of the RSV N protein formed RNase-resistant structures that were nucleocapsid-like in appearance. This suggests that characteristics of RNA-binding and nucleocapsid assembly are situated within the same domain. A potential map of the domain structure of RSV N protein, derived from this work and from studies on-going in the laboratory, is presented.

Declaration

I declare that the following thesis embodies the results of my own special work, that it has been composed by myself and that it does not include work forming part of a thesis presented successfully for a degree in this or another University.

Lindsay Bryony Murphy
December, 2001

Acknowledgements

Table of Contents

List of Figures

List of Tables

Abbreviations

aa	Amino acid
ADP	Adenosine diphosphate
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CAT	Chloramphenicol acetyl transferase
CI	Inclusion body
CIP	Calf intestinal phosphotase
CO_2	Carbon dioxide
cpm	Counts per minute
CTL	Cytotoxic T lymphocytes
CTP	Cytosine triphosphate
dH_2O	De-ionised water
DI	Defective Interfering genomes
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
E.coli	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetra-acetic acid disodium salt
EM	Electron microscopy
FCS	Foetal calf serum
FI-RSV	Formalin inactivated respiratory syncytial virus
GMSA	Gel mobility shift assay
HRP	Horseradish peroxidase
IF	Immunofluorescence
IFN	Interferon
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	Kilo bases
kDa	Kilo Daltons
L	Litres
M	Molar

mA	Milli Amperes
MAb	Monoclonal antibody
µCi	Micro Curie
µl	Microlitres
µm	Micrometers
µM	Micromolar
m.o.i.	Multiplicity of infection
ml	Millilitres
mm	Millimetres
mM	Millimolar
mwco	Molecular weight cut-off
nm	Nanometers
nM	Nanomolar
N-P40	Nonidet-P40
nt	Nucleotide
O.D.	Optical density
°C	Degrees Celsius
ORF	Open reading frame
%	Percentage
PBS	Phosphate buffered saline
pM	Picomolar
pfu	Plaque-forming units
PNK	Polynucleotide kinase
PTA	Phosphotungstic acid
RER	Rough endoplasmic reticulum
RIPA	Radiolabelled immuno-precipitation assay
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
SDS	Sodium dodecyl sulphate
TCA	Trichloro-acetic acid
TFA	Trifluoro-acetic acid
UV	Ultra violet
V	Volts
v/v	Volume by volume
w/v	Weight by volume

w/w   Weight by weight

WHO   World Health Organisation

xg   centrifugal force (times gravity)

Amino Acid Abbreviations

A   Ala   alanine

B   Asx   asparagine or aspartic acid

C   Cys   cysteine

D   Asp   aspartic acid

E   Glu   glutamic acid

F   Phe   phenylalanine

G   Gly   glycine

H   His   histidine

I   Ile   isoleucine

K   Lys   lysine

L   Leu   leucine

M   Met   methionine

N   Asn   asparagine

P   Pro   proline

Q   Gln   glutamine

R   Arg   arginine

S   Ser   serine

T   Thr   threonine

V   Val   valine

W   Trp   tryptophan

Y   Tyr   tyrosine

Z   Glx   glutamine or glutamic acid

Virus Abbreviations

APV   Avian Pneumovirus

bRSV   Bovine respiratory syncytial virus

hRSV   Human respiratory syncytial virus

MV   Measles virus

NDV   Newcastle disease virus

PIV   Parainfluenza virus

PVM   Pneumovirus of mice

RV   Rabies virus

SeV   Sendai virus

SV5   Simian virus 5

VSV   Vesicular stomatitis virus

1 Introduction

1.1 History

Respiratory syncytial virus (RSV) was first reported in 1956 upon its identification as the causative agent of coryza, a respiratory illness in chimpanzees (Morris *et al.*, 1956). Shortly after its discovery, the virus was isolated from children in the Baltimore area of New York suffering from either pneumonia or from croup. The formation of lung epithelial cell syncytia, a characteristic of RSV infection from which it derives its name, was first observed in these patients (Chanock *et al.*, 1957). An extensive serological investigation of children within the Baltimore area revealed that most individuals had been exposed to the virus before the age of four (Chanock & Finberg, 1957). Current serological data indicates that by the age of two years, up to 90% of children have encountered RSV. Peak rates of RSV infection occur in infants aged between six weeks and six months and, in addition to the very young, RSV may also cause serious infection in the immunosuppressed and the elderly (Hall, 2000; Simoes, 1999). Reinfection is a frequent event in both children and adults due to the highly variable nature of circulating strains, similar to the situation with influenza virus.

1.1.1 Epidemiology

In 1992, the World Health Organisation (WHO) estimated that, of the 12.2 million annual deaths of children under the age of five, one third were due to infections of the lower respiratory tract (Garenne *et al.*, 1992). The predominant respiratory pathogens identified in this study were *Streptococcus pneumonia*, *Haemophilus influenza* and RSV.

RSV has a worldwide distribution. In temperate regions, a clear seasonality may be observed with the occurrence of epidemics in the winter months. Epidemics are an annual event and have a mean duration of five months with peaks in infection during the months of February and March in the Northern hemisphere. In tropical and sub-tropical regions of the world, epidemics often occur during the cooler, rainy season. There is correlation with the spread of infection and large gatherings of people, for example, at religious festivals (Simoes, 1999).

1.1.2 Antigenic Subgroups

Investigations of viral isolates using monoclonal antibodies (MAbs) demonstrated the existence of two antigenically distinct subgroups, termed A and B (Anderson *et al.*, 1985; Gimenez *et al.*, 1986; Mufson *et al.*, 1985). Within the laboratory these are represented by the strains A2 and 18537, respectively. Within each subgroup, viruses appear antigenically relatively stable. Comparison between subgroups indicates an antigenic relatedness of only 25%. Such antigenic dimorphism is predominantly the result of variation in the virus attachment protein, the G protein, which possesses 1 to 7% antigenic relatedness between subgroups A and B in comparison with 50% antigenic relatedness observed between the different subgroup fusion (F) proteins (Hendry *et al.*, 1988); (Johnson *et al.*, 1987a). Examination of the protective antigenic regions on the F protein demonstrates that there has been neither extensive nor rapid subgroup divergence. In addition, the accumulation of immune-evasive traits with time, within a subgroup, is not apparent. This suggests that, unlike influenza A, RSV does not undergo antigenic drift (reviewed by Collins *et al.*, 1996b).

Sequences from eight of the genes from RSV strains A2 and 18537 were compared by sequencing (Johnson *et al.*, 1987b). Considering only the envelope proteins it was shown that the small hydrophobic (SH) and the F protein genes show moderate sequence divergence. The G protein gene, however, had extensive divergence between the two subgroups. Upon translation of the nucleotide sequence of the open reading frames, this represented a 47% difference in amino acid sequence. The sequence variation was concentrated in the extracellular domain of the protein which suggests selective pressure for the divergence of G protein, perhaps because of immune pressure. It is hypothesised that the F protein is less tolerant of sequence alteration due to more stringent structural and functional restrictions.

The antigenic divergence between subgroups may contribute to the high incidence of reinfection during early childhood. However, it is probable that this has only a modest effect, as reinfections may involve viruses from the same subgroup due to rapid RSV evolution (Mufson *et al.*, 1987). The G protein has been found not only to be highly variable between subgroups but that some subgroup A isolates, from the same locale, show up to 20% variability (Cane *et al.*, 1991).

1.2 Disease in Humans

In temperate climes, RSV is the principal cause of hospitalisation of children, due to respiratory illness, in their first year of life (van den Hoogen *et al.*, 2001). However, mortality due to RSV is uncommon in the developed world. It's occurrence is predominantly in infants with underlying illness, such as congenital heart disease and pulmonary hypertension (Simoes, 1999). RSV is now considered a major cause of death in the elderly. Within this age group, at least 10% of hospital admissions during the winter months are caused by RSV, a further 10% of which are fatal, presenting a similar fatality rate to influenza (Falsey *et al.*, 1995). Acute RSV infections are also a common occurrence in adults, though with no requirement for hospitalisation.

RSV is spread in infected respiratory secretions, carried in large droplet form or fomite contamination rather than by small-particle aerosols (Hall & Douglas, 1981). Therefore, virus spread requires close personal contact or object contamination with fomite, the latter of which is more common in hospital-acquired infection (Hall, 2000; Isaacs, 1991).

RSV has an incubation period of approximately five days (Johnson *et al.*, 1961). Initially, virus replicates in the nasopharynx before spreading from the upper to the lower respiratory tract. Signs of lower respiratory tract infection appear between one and three days after the onset of rhinorrhea, possibly reflecting the timing of viral spread to the bronchi and the bronchioles. Infection is limited to the superficial layers of the respiratory epithelium. However, virus spread to other tissues and establishment of persistent infection may occur in severely immuno-compromised patients (Simoes, 1999). In severe cases of RSV in infants, the pathology of resulting bronchiolitis or pneumonia is similar to that obtained with other respiratory viruses. It is believed that the predominant site for RSV infection is the small bronchioles, accompanied by inflammation and leukocyte infiltration of surrounding tissue (Hussell *et al.*, 1998). Such inflammation is the predominant cause of RSV pathology, with minor contribution from direct virus cytopathology. Therefore, treatment with both anti-inflammatory agents and anti-virals may be administered to alleviate the disease state (Simoes, 1999).

1.2.1 Clinical Features

With RSV, the more common infection is of the upper respiratory tract causing symptoms of rhinitis and cough, with low-grade fever in some patients. Upon development of lower respiratory infection, bronchiolitis and pneumonia are the most common manifestations in children. Croup is a less frequent development of RSV infection (Simoes, 1999). In healthy infants that first encounter RSV at six weeks to nine months, RSV usually causes upper respiratory symptoms. 25 to 40% of these cases progress to lower respiratory infection (Glezen et al., 1986). Most frequently, uneventful recovery occurs after illness of between seven and 12 days. However, following RSV-induced bronchiolitis or pneumonia during infancy, abnormalities in pulmonary function may be observed (Sigurs, 2001). Factors associated with development of serious disease include prior underlying anatomic or functional abnormality, environmental factors such as family smoking, lower socio-economic circumstances and a tendency of the individual to secrete IgE in response to RSV (Brandenburg et al., 2001; Simoes, 1999).

RSV has been detected in the lungs of a percentage of children dying from sudden infant death syndrome (SIDS) (Ogra et al., 1975). In such instances, the role of RSV is unknown. However, it is probable that a subset of sudden deaths is caused by respiratory illness and therefore possible that a proportion of these is due to RSV.

1.2.2 Diagnosis

A variety of assays exist for the laboratory diagnosis of RSV infection in older children and in adults. An increase in the levels of serum antibody, complement-fixing and neutralising antibody may all be used as indicators of infection (Johnson et al., 1961). Diagnostic assays for the detection of virus antigens include immunofluorescent microscopy and enzyme-linked immunosorbant assays (ELISA) performed on nasopharyngeal secretions. Reverse transcription of patient RNA followed by polymerase chain reaction with primers specific for virus sequence is an alternative method for the diagnosis of RSV in patients with limited virus shed from the nasopharynx, a situation frequent in elderly patients (Simoes, 1999).

1.2.3 Therapeutic Treatment of RSV

Considerable supportive care is required in the treatment of hospitalised patients with RSV including the mechanical removal of secretions, positioning of the infant, administration of humidified oxygen and possible provision of respiratory assistance (Behrendt *et al.*, 1998). In addition to reducing the discomfort of the patient, treatment of RSV infection is beneficial by shortening the duration of hospital stay and the intensity of nursing required.

A variety of treatments are employed for RSV infection including both drug- and immunotherapy. The compound ribavirin, a guanosine analogue, has been used in the treatment of RSV since the mid-1980s. Its antiviral properties were first observed in cultured mammalian cells and in laboratory cotton rats (Hruska *et al.*, 1982). In RSV infection, the drug is administered by aerosol over 18 to 20 hours a day to reduce, or eliminate, the acute systemic toxicity of the drug. Studies into the efficacy of the use of ribavirin have had conflicting results. The drug has been reported to provide a modest but consistent beneficial effect on virus shedding and clinical illness (Hall *et al.*, 1983). However, there is no evidence that treatment of the virus infection reduces the frequency of mortality, the duration of hospitalisation or the need for supportive therapies (Law *et al.*, 1997). In most centres, its use is now restricted to immunocompromised patients and the severely ill; this limitation arises from the high cost of the drug, its application and lack of consensus on its efficacy (Simoes, 2001).

Inhaled salbutamol, a bronchodilator, is used in the therapy of RSV. A meta-analysis demonstrated that the drug induced moderate short-term improvement in mild or moderate bronchiolitis. However, treatment had no effect on the risk of hospitalisation (Kellner *et al.*, 1996). An alternative drug used in the treatment of RSV is racemic epinephrine. It is superior to salbutamol in improving airway resistance and clinical scores and decreasing the need for hospitalisation of babies with bronchiolitis (Menon *et al.*, 1995).

Immunotherapy is used in the treatment of RSV. The use of purified human intravenous immune globulin (IVIG) was first tested in laboratory cotton rats. Upon inoculation at the height of RSV infection, a reduction in the level of pulmonary virus was effected in the rats and there was no evidence of pathology (Prince *et al.*, 1987). However, in human trials performed with both intravenous and inhaled antibody

application, no substantial beneficial effects on lower respiratory tract infection were observed (Rodriguez *et al.*, 1997).

Immunoglobulin therapy has been used more successfully in the prevention rather than the therapy of RSV infection. Children identified as being at high-risk of serious infection can be protected against severe disease by the monthly administration of IVIG, or humanised MAb, during the RSV season. In trials, the use of Synagis (or palivizumab, as it was previously called), a humanised MAb with specificity for the F protein, resulted in fewer hospitalisations within the patient group. Though a cumbersome procedure of application, Synagis was found to be well tolerated and effective (Johnson *et al.*, 1997). This treatment was licensed in the USA in 1998. Synagis is administered to premature babies, born with an undeveloped immune system and limited maternal antibody, to protect them from RSV infection.

1.2.4 Immunology

The effectors of both protection and recovery from RSV infection are secretory antibodies, serum antibodies, MHC class I-restricted cytotoxic T lymphocytes (CTL) and maternally derived antibodies in the very young. Following RSV infection, CTL-mediated protection appears to be short-lived. Both serum and secretory antibodies contribute but provide incomplete protection against repeat infection (Brandenburg *et al.*, 2001). Cumulative immunity after repeated reinfections provides protection against more serious lower respiratory tract infection in older children and adults (Kimman & Westenbrink, 1990).

Work performed in mice suggests CD8$^+$ lymphocytes play a role in the pathogenesis of RSV infection. Passively transferred RSV-specific CTLs, obtained from mice infected with RSV, were observed to clear virus from the lungs of persistently infected, gamma-irradiated mice. However, acute pulmonary damage resulted and the mice frequently died (Cannon *et al.*, 1988). Experiments involving spleen cells, enriched in CD4$^+$ or CD8$^+$ lymphocytes, indicated that both were able to decrease pulmonary virus titre in mice, with CD4$^+$ cells being more efficient. Concurrently, CD4$^+$ cells induced pulmonary eosinophilia and secreted a T$_{H2}$ cytokine profile (Alwan *et al.*, 1992). These observations, confirmed by related studies, suggest that there is a fine balance between the protective and the disease-producing effects of T cells. Despite the

successful removal of virus by CTLs, it is the mediators of the inflammatory response that cause the disease pathology associated with RSV infection. In a variety of experimental animals, recombinant M2 and, to a lesser extent, recombinant N protein induced a protective effect, mediated by CTLs, and of short duration (Connors et al., 1991). In adult humans, the CTL response repertoire included the N, SH, F, M, M2 and NS2 proteins (Cherrie et al., 1992). The F protein is the major CD4[+] cell target antigen in both humans and BALB/c mice. Recombinant F protein primarily induced a T_{H1} response in CD4[+] cells compared with the T_{H2} response elicited by the G protein. This corresponded with respectively minimal or enhanced lung histopathology. Therefore, the antigen composition of a vaccine can influence the quality of the immune response as well as the immunopathology generated upon virus challenge (Openshaw et al., 2001).

Specific antibodies are generated upon RSV infection. Antiviral IgE may be involved in pathogenesis. Upon association with antigen, IgE binds both basophils and mast cells stimulating a local allergic response (Campbell, 1993). Association of IgE production with pathogenesis again confirms that inflammation is the major cause of RSV pathology. Protection against reinfection by respiratory viruses is traditionally attributed to secretory antibodies. However, serum antibody appears to have a protective role in RSV infection, demonstrated by epidemiological studies and antibody trials. The incidence and severity of lower respiratory tract infection in infants with RSV has been observed to correlate inversely with serum antibody levels (Brandt et al., 1997; Glezen et al., 1981; Kimman & Westenbrink, 1990). The source of protection within the respiratory tract of cotton rats appears to be divided. The upper respiratory tract is protected primarily by local immunity, including secretary antibody, whereas the lower respiratory tract is protected by serum-neutralising antibody (reviewed by Collins et al., 1996b). The major neutralising antibody response is targeted against the RSV F protein whereas the G protein induces only a weak neutralising antibody response (Johnson & Collins, 1988).

1.2.5 Interferon

RSV infection frequently fails to induce detectable levels of interferon (IFN) in the nasal secretions, which is in contrast to infection by parainfluenza and influenza viruses where IFN levels are high at the peak infection and correlate with clearance of the virus

(Hall *et al.*, 1978). It was suggested that IFN, together with cell-mediated immunity, modulates the severity of RSV-induced disease. Severe RSV disease was correlated with decreased IFN-γ production by peripheral blood monocytes (Bont *et al.*, 1999). Atreya & Kulkarni (1999) showed that RSV, in tissue culture, was resistant to the antiviral effects of type I IFNs and to MxA, unlike some other paramyxoviruses. The mechanism of the resistance is unknown but Schlender *et al.* (2000) demonstrated that the bovine RSV (bRSV) non-structural proteins NS1 and NS2 are both required for anti-IFN effect. Young *et al.* (2000) showed that simian virus 5 (SV5), Sendai virus (SeV) and parainfluenza types 2 and 3 exert anti-IFN effects by interfering with the STAT1/STAT2 transcription pathway, effectively blocking the signalling pathway. RSV did not block the IFN signalling response therefore it must use some other mechanism to combat IFN effects. Thus, it appears that while RSV replication is unaffected by IFN, the effect of decreasing IFN is to enhance the inflammatory response that results in disease. There is evidence that use of IFN therapy, given with an appropriate antiviral may help abate the severity of the disease (van Schaik *et al.*, 2000).

1.2.6   Vaccine

A clinically approved vaccine is not yet available for RSV. Problems such as the insufficient attenuation of vaccine virus, its thermolability and the need to include both A and B strains have hindered vaccine development (Simoes, 1999). In addition, to be effective against infant bronchiolitis and pneumonia, the vaccine must be capable of stimulating effective resistance before the second month of life. However, the immunosuppressive effects of maternally derived serum RSV-neutralising antibody, and immunological immaturity of the infant, present additional problems in vaccine design and delivery (Brandenburg *et al.*, 2001).

In the mid-1960s, a formalin-inactivated, parenterally administered RSV vaccine (FI-RSV) was developed. Vaccine recipients developed a high titre of serum complement-fixing antibody and a moderate titre of serum-neutralising antibody. Despite this, an unusually large proportion developed severe lower respiratory tract illness upon subsequent RSV challenge, similar to the severe bronchiolitis observed during infancy (Chin *et al.*, 1969; Fulginiti *et al.*, 1969; Kim *et al.*, 1969). Therefore, the vaccine appeared to induce sensitisation leading to severe disease in subsequent infection. The vaccine may have stimulated an imbalance in the systemic immunoglobulin response. It is possible that vaccine failed to induce appreciable local respiratory tract

RSV secretory IgA antibody because of its parenteral administration. In addition, a high titre of serum antibody directed towards the F protein was generated but exhibited only a low level of neutralising activity. Formalin treatment of the virus may affect the antigenicity of viral glycoproteins F and G and expose less favourable epitopes. *In vitro* analysis of lymphocytes obtained from vaccinees, demonstrated that RSV antigens stimulated an exaggerated proliferative CTL response. It was proposed that formalin-inactivated vaccine induces a high level of RSV-specific $CD4^+$ but not $CD8^+$ memory cells (Kim *et al.*, 1976). In mice immunised with FI-RSV and subsequently challenged with RSV, $CD4^+$ lymphocytes were observed to predominate in the pulmonary infiltrate. In addition, contrary to the usual observation in mice during unmodified RSV infection, the $CD4^+$ lymphocyte cytokine profile T_{H2} predominated (Graham *et al.*, 1993). These observations suggested that vaccinees might also experience an imbalance in their cell-mediated immune response.

To over-come the failure of FI-RSV, the development of attenuated viruses for use as a live vaccine was attempted. Live vaccines have been developed for other members of the *Paramyxoviridae*, the viral super-family to which RSV belongs, such as measles and mumps viruses. To this effect, attenuated, immunogenic cold-passaged (cp) or temperature sensitive (ts) viruses were generated (Crowe *et al.*, 1994a; Crowe *et al.*, 1994b; Friedewald *et al.*, 1968; Gharpure *et al.*, 1969; McKay *et al.*, 1988). Early trials in susceptible infants and young children yielded discouraging results as immunisation by parenteral inoculation of live virus proved ineffective (Belshe *et al.*, 1982). Attenuated strains possessed only a low level of residual virulence and were genetically unstable. To produce mutants free of these deficiencies, cp-RSV was further mutated. Two distinct lineages of mutagenised virus each yielded promising candidate vaccine strains (Crowe *et al.*, 1994b). Despite reduced replication in chimpanzees, these mutants induced almost complete protection to subsequent RSV infection (Crowe *et al.*, 1994a). Their usefulness will not be known until they have been evaluated in susceptible sero-negative infants.

The application of viral proteins to induce protection has been investigated. F protein, purified from RSV culture (Walsh *et al.*, 1987), and a chimeric F-G protein, expressed from recombinant baculovirus, have been studied (Kakuk *et al.*, 1993; Wathen *et al.*, 1991). In cotton rats, RSV protection was provided when challenged several months post-immunisation. However, though a high level of F-specific antibody was stimulated,

only a low level of neutralising activity was exhibited (Connors *et al.*, 1992). Vaccinia virus recombinants expressing the F and the G proteins generated encouraging results in both rodents and monkeys (Olmsted *et al.*, 1986). Unfortunately, they were only weakly immunogenic in chimpanzees and failed to induce resistance to RSV challenge.

Bovine RSV has been considered as a live vaccine for use in humans. Initial studies in rodents yielded encouraging results (Piazza *et al.*, 1993). However, the virus does not replicate efficiently in chimpanzees and therefore fails to induce resistance to hRSV challenge. Bovine RSV is unable to infect human cells, however, it may yet prove a viable vaccine candidate upon generation of recombinant bRSV expressing hRSV F and G proteins.

1.3 Animal RSV

Bovine RSV is an economically important disease in cattle. The naturally occurring virus may be divided into at least two antigenic subgroups. The genome organisation of human and bovine RSV is identical. Differences in nucleotide and amino acid sequence occur in similar regions to those between subgroups of human RSV, though differences are approximately two-fold greater. In addition, antibody cross-reactivity exists between several human and bovine RSV proteins (N, P, M and F). Despite such similarities, bovine RSV is unable to replicate efficiently in chimpanzees and to infect human cells.

Other ruminant forms of the virus exist. Caprine RSV is very similar to the bovine virus. It is possible that ovine and bovine RSV are two antigenic subgroups rather than distinct viruses. Therefore, it is possible that ruminant RSVs represent antigenic subgroups rather than the distinct viruses of different hosts (Collins *et al.*, 1996b).

1.4 Virology

1.4.1 Classification

The genome of RSV is composed of nonsegmented, single-strand, negative-sense RNA, placing the virus within the taxonomic Order, *Mononegavirales* (Pringle, 1991). Figure 1.1 shows the taxonomic organisation of the *Mononegavirales*, with particular regard to the *Paramyxoviridae*. The Order *Mononegavirales* contains four virus
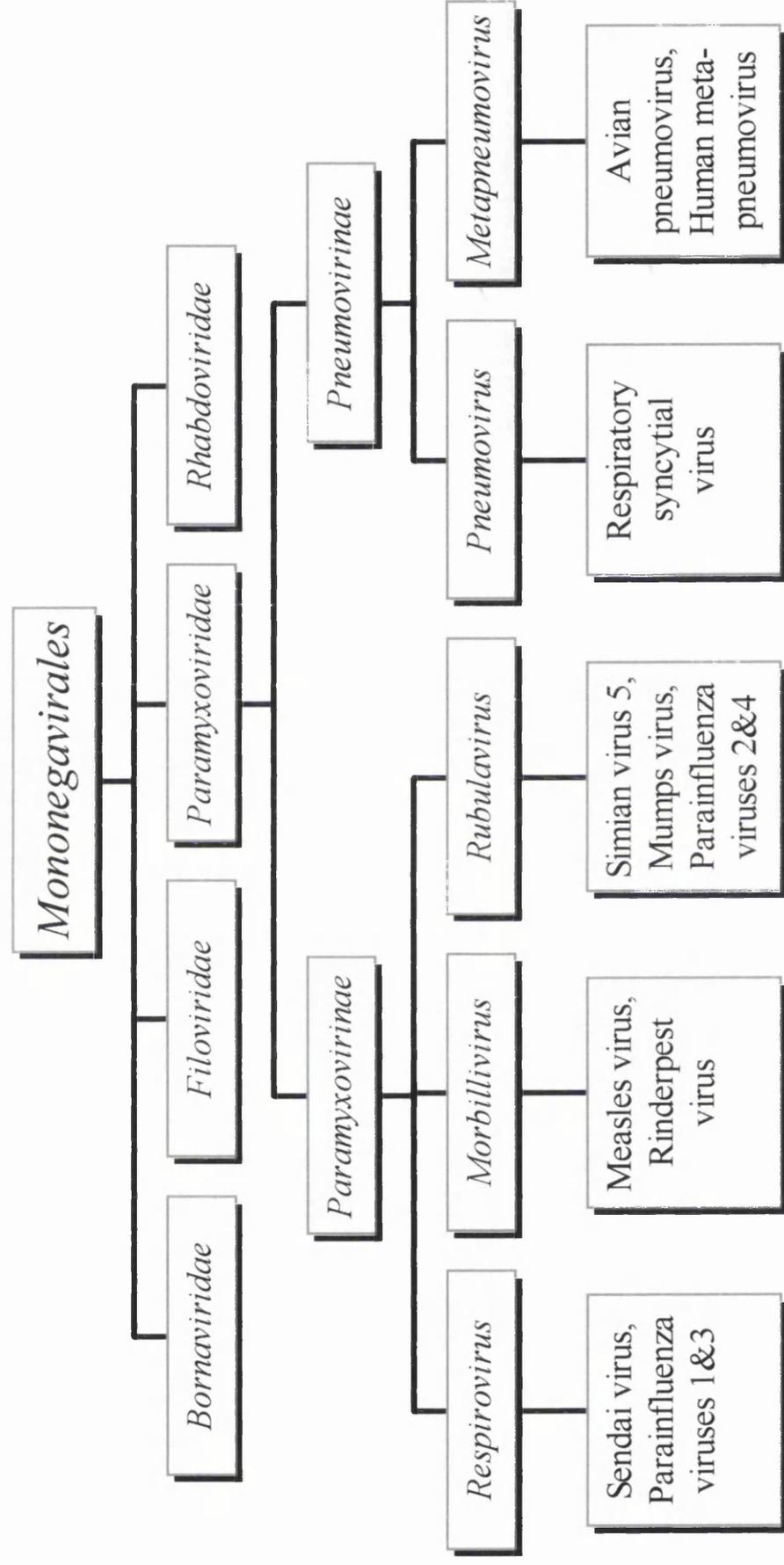
Figure 1.1: The taxonomy of the Order *Mononegavirales* with particular emphasis on the *Paramyxoviridae*.

Mononegavirales

- *Bornaviridae*
- *Filoviridae*
- *Paramyxoviridae*
  - *Paramyxovirinae*
    - *Respirovirus* — Sendai virus, Parainfluenza viruses 1&3
    - *Morbillivirus* — Measles virus, Rinderpest virus
    - *Rubulavirus* — Simian virus 5, Mumps virus, Parainfluenza viruses 2&4
  - *Pneumovirinae*
    - *Pneumovirus* — Respiratory syncytial virus
    - *Metapneumovirus* — Avian pneumovirus, Human meta-pneumovirus
- *Rhabdoviridae*

families, the *Filoviridae*, *Bornaviridae*, *Rhabdoviridae* and the *Paramyxoviridae*, example members of which are listed in Table 1.1. The viruses of this order most extensively studied include SeV (Murine Parainfluenza Subtype 1, family *Paramyxoviridae*) and VSV (Vesicular Stomatitis Virus, family *Rhabdoviridae*). From investigations into these prototype viruses, much information on the biology of the *Mononegavirales* has been elucidated. Parallels are drawn and investigated in related viruses. There are a number of similarities, which result in the grouping of viruses into the *Mononegavirales*, some of which are listed in Table 1.2. However, there are marked differences between the virus families (Table 1.2) such as the method of cell entry (pH dependent or independent), presence or absence of heamagglutinin activity or site of replication. Nevertheless, it is the common replication strategy, and the use of a virally encoded RNA-dependent RNA polymerase, that unites the group.

The RS viruses are members of the genus *Pneumovirus*, of the family *Paramyxoviridae*. In 1993, the *Paramyxoviridae* family was further partitioned forming two subfamilies, the *Paramyxovirinae* and the *Pneumovirinae*. Division of the family was based on morphology, genome organisation and the biological activity and sequence relatedness of the proteins. Subfamily *Paramyxovirinae* contains three genera, *Respirovirus*, *Morbillivirus* and *Rubulavirus* (Table 1.1). The genera *Pneumovirus* and *Metapneumovirus* are present in the subfamily *Pneumovirinae* (Lamb & Kolakofsky, 1996; Pringle, 1997). RSV is a member of the genus *Pneumovirus*. The genus *Metapneumovirus* was generated to accommodate the avian pneumoviruses. Subsequently, a novel virus has been isolated from children with respiratory infection in the Netherlands; virus sequence homology and genome organisation suggest it too may be classified as a *Metapneumovirus*. The provisionally named "human metapneumovirus" causes symptoms similar to those of RSV (van den Hoogen *et al.*, 2001).

1.4.2 Virus Growth

RSV grows relatively poorly in tissue culture and in many experimental animals. However, virus growth is permissible in a wide variety of human and animal cells. Upon growth, RSV does not shut down host macromolecular synthesis. Approximately 90% progeny virus remains cell associated, in the form of viral filaments (Bachi, 1988). RSV rapidly looses infectivity upon storage, handling and purification and is consequently

Family	Subfamily	Genus	Example Species
Filoviridae		*Filovirus*	Ebola Virus Marburg Virus
Bornaviridae		*Bornavirus*	Borna Virus Disease
Rhabdoviridae		*Vesiculovirus* *Lyssavirus* *Ephemerovirus* *Cytorhabdovirus* *Nucleorhabdovirus*	Vesicular Stomatitis Virus Rabies Virus Bovine Ephemeral Fever Virus Lettuce Necrotic Yellows Virus Potato Yellow Dwarf Virus
Paramyxoviridae	*Paramyxovirinae*	*Respirovirus* *Morbillivirus* *Rubulavirus*	Sendai Virus Human Parainfluenza Virus Types 1&3 Measles Virus Mumps Simian Virus 5 Newcastle Disease Virus
	Pneumovirinae	*Pneumovirus* *Metapneumovirus*	Human RSV Bovine RSV Pneumonia Virus of Mice Avian Pneumovirus (Turkey Rhinotracheitis Virus)

Table 1.1: Members of the Order *Mononegavirales*.

Similarities	Differences
1 Possess a similar gene order.	1 There is variation in genome size. e. g. 11 161 kilobases (kb) of vesicular stomatitis virus (*Rhabdovirus*) compared with 19 108 kb of Marburg virus (*Filovirus*).
2 Genome is always tightly associated with viral protein along its entire length, forming the nucleocapsid.	
3 Transcription of viral genes is by sequential, interrupted synthesis of mostly non-overlapping mRNAs.	2 Differences in virion morphology. 3 Viruses differ in the cellular site of replication.
4 Replication is achieved by production of a complete positive-sense transcript.	4 Variation in pathogenic potential. 5 Differences in virus host range. e.g. the *Rhabdoviridae* host range includes plants, invertebrates and vertebrates; the filoviruses are restricted to primates.
5 Virus maturation occurs by budding through the host cell plasma membrane that the virion then retains.	
6 Virus encodes a virion-associated RNA-dependent RNA polymerase.	6 Method of cell entry varies. e.g. rhabdoviruses use a ph-dependent lysosomal pathway; paramyxoviruses use a ph-independent pathway.

Table 1.2: The similarities and differences between members of the Order *Mononegavirales*.

(Reviewed by Lamb & Kolakofsky, 1996).

difficult to concentrate (Fernie & Gerin, 1980). Within tissue culture, RSV has been shown to bud from the apical membrane of the host cell (Roberts *et al.*, 1995). This might have implications *in vivo* and could explain the restriction of RSV to the lungs of infected individuals.

Several small experimental animals, for example the cotton rat, have been used as models for RSV infection. However, in most cases replication is only semi-permissive and disease does not occur. Of non-human primates, only the chimpanzee appears to be highly permissive to RSV replication (Collins *et al.*, 1996b). However, chimpanzee use in vaccine studies is naturally restricted by availability, expense and ethical considerations. As a model for human disease, the study of bRSV in cattle may be useful.

1.4.3 Virus Morphology

The virions of RSV are of a similar size and shape to those of the *Paramyxovirinae*. They are generally spherical with a diameter of between 150 nm and 300 nm but are pleomorphic in appearance. Figure 1.2 shows a stylised image of an RSV virion. In addition, filamentous forms of virus have been observed in tissue culture but their significance is unknown (Pringle *et al.*, 1981; Roberts *et al.*, 1995). The RSV envelope is derived from the host cell as it buds from it (Lamb and Kolakofsky, 1996). Inserted into the membrane are viral glycoproteins which can be visualised by negative-stain electron microscopy as spikes extending 8 to 12 nm from the virion surface (Bachi & Howe, 1973; Lehmkuhl *et al.*, 1980). In the case of RSV, there are three glycoprotein species, the SH (small hydrophobic), F (fusion) and G (viral attachment) proteins (Collins *et al.*, 1996b)

Within the virion is the nucleocapsid; its core structure is composed of viral genomic RNA and the N (nucleocapsid) protein. The viral P (phospho-) protein and L (large) protein are also attached to this helical structure, which is the template for RNA synthesis. The nucleocapsid is very stable, able to withstand the high salt and gravity forces of caesium chloride density gradient centrifugation. Nucleocapsids band at a density of between 1.29 and 1.31 g / cm^3 within caesium chloride gradients, typical of a ribonucleoprotein (RNP) complex. The nucleocapsids of the *Paramyxoviridae* may be viewed by negative-stain EM; those of the *Paramyxovirinae* are approximately 18 nm in
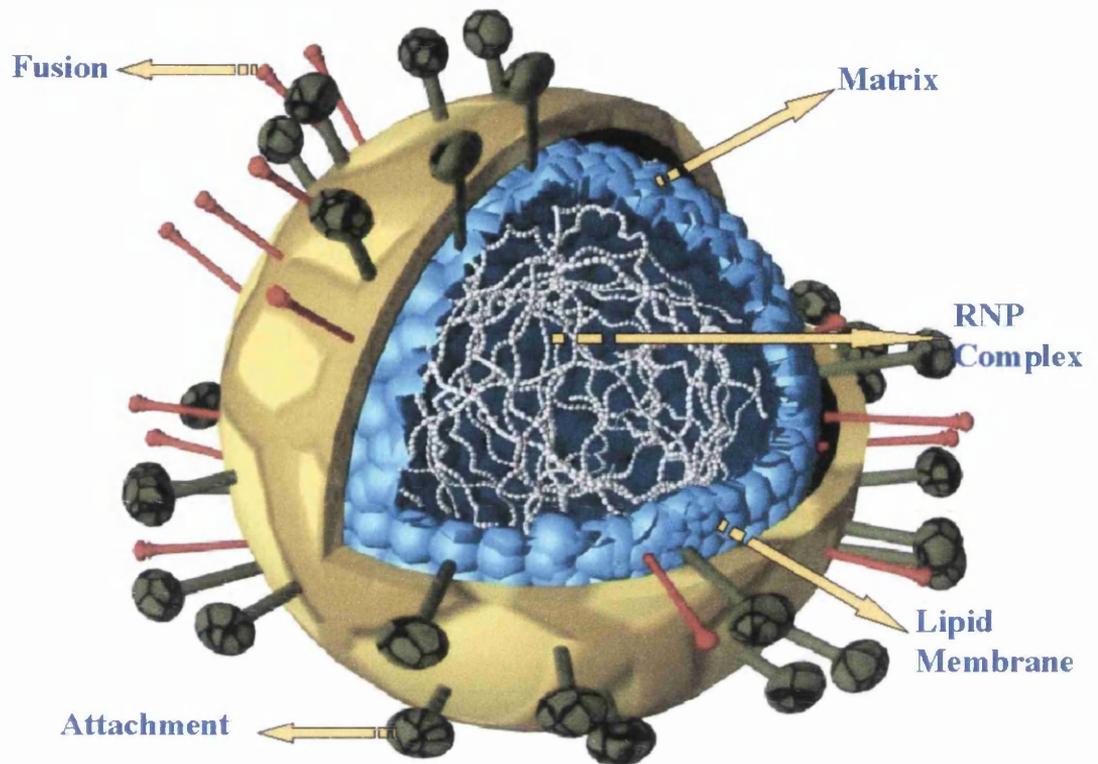
Figure 1.2: Stylised image of a Pneumovirus virion.

The Pneumovirus virion consists of a ribonucleoprotein (RNP) core containing genomic RNA and viral proteins N, P, L and M2. The RNP complex is surrounded by a protein matrix (M), itself surrounded by cell-derived lipid membrane containing the F and the G proteins, responsible for fusion and attachment respectively. The viral membrane protein SH is absent from this image.

Figure courtesy of A. Easton, University of Warwick, UK.

diameter, 1 µm in length and 5.5 nm in pitch length, although variation does occur (Egelman *et al.*, 1989). The nucleocapsids of the *Pneumovirinae* are of similar dimensions but are narrower, with a diameter of between 12 nm and 15 nm (Bachi & Howe, 1973; Lehmkuhl *et al.*, 1980).

1.5 The Genome

The genomes of the *Paramyxoviridae* are of similar size, ranging between 15.2 and 15.9 kilo bases (kb). A cartoon of the genome organisation of family members is present in Figure 1.3. The genome sequence of human RSV strain A2 has been determined; the RNA is 15 222 nucleotides in length and encodes ten genes (see Genbank, accession number M75468) (Collins *et al.*, 1987; Mink *et al.*, 1991; Stec *et al.*, 1991). This gene number is shared by all members of the *Pneumovirinae* subfamily, with the exception of the avian pneumoviruses, which lack the genes encoding non-structural proteins, NS1 and NS2 (Randhawa *et al.*, 1997). In contrast, the genomes of the *Paramyxovirinae* subfamily contain only six (*Respirovirus* and *Morbillivirus*) or seven (some members of genus *Rubulavirus*) genes (Figure 1.3). The genes for NS1 and NS2 are the first on the RSV genome map, which is a significant difference to the *Paramyxovirinae* subfamily where N is the first gene encountered on the genome (Lamb & Kolakofsky, 1996). As well as lacking NS1 and NS2, APV and the human metapneumovirus differ from other *Pneumovirinae* in their gene order, F-M2-SH-G compared with SH-G-F-M2.

A feature that appears to be unique to RSV is that the start of the L gene overlaps with the end of the M2 gene. As discussed later, this overlap may down-regulate expression of the L gene.

1.5.1 Extending the coding potential of the genome

Other members of the *Paramyxoviridae* have extended their genome coding potential by using alternative expression methods. Using the P gene of SeV as an example, the coding potential of this gene has been expanded by the use of alternative translational open reading frames (ORFs) by a mechanism of "leaky" scanning by the ribosome, the presence of AUG codons in poor translational contexts and the use of non-AUG codons to initiate translation (Curran & Kolakofsky, 1988a; Curran & Kolakofsky, 1988b; Curran *et al.*, 1986). SeV also extends the coding potential of the P gene by
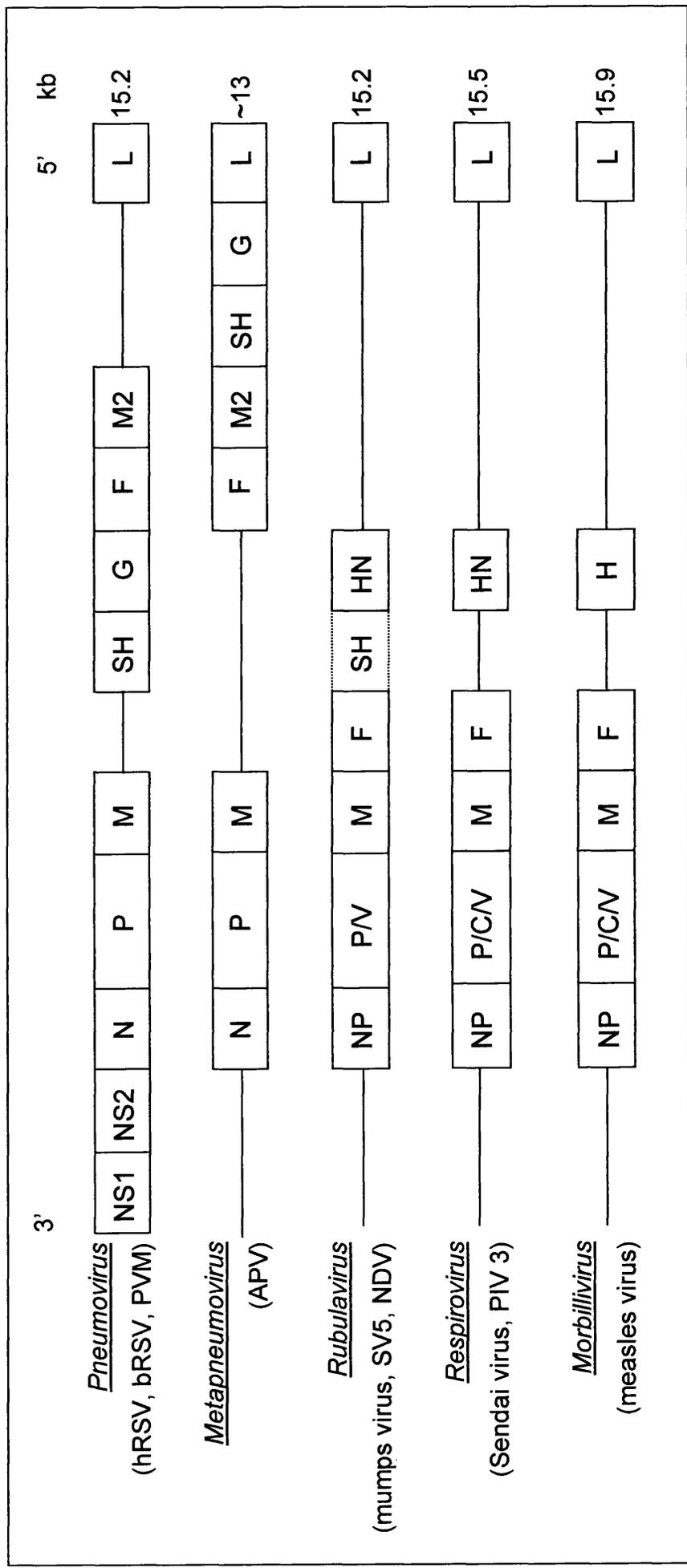
Figure 1.3: A cartoon presenting the gene order in genera of the family *Paramyxoviridae*.

Virus name abbreviations are as follows: Human/Bovine Respiratory Syncytial Virus (h/b RSV), Pneumonia Virus of Mice (PVM), Avian Pneumovirus (APV), Simian Virus 5 (SV5), Newcastle Disease Virus (NDV) and Parainfluenza Virus 3 (PIV 3). Gene name abbreviations are as follows: nucleocapsid protein gene (N/NP), phosphoprotein gene (P) and internal derivatives C & V, matrix protein gene (M), fusion protein gene (F), small hydrophobic protein gene (SH), the attachment protein gene (HN/H/G) and the large protein gene (L). Individual genes are not drawn to scale. (Collins *et al.*, 1986, reviewed by Kingsbury, 1991).

RNA editing, which involves the insertion of non-templated nucleotide residues into the mRNA (Vidal *et al.*, 1990). Unmodified mRNAs from the SeV P gene encode the P protein. Upon insertion of extra nucleotides within the mRNA, a frameshift occurs downstream of the insertion point allowing the ribosome to translate the V carboxy region, characterised by a zinc-finger motif. There is evidence that RSV utilises an RNA editing mechanism in the G protein to generate antibody escape mutants (Garcia-Barreno *et al.*, 1990). RSV extends the coding potential of the M2 gene. The gene contains two ORFs, producing the M2-1 and M2-2 proteins which use a unique translational event. Within the mRNA, the start codon of M2-2 is 5' proximal to the M2-1 protein stop codon. Ahmadian *et al.* (2000) demonstrated the ability of ribosomes to terminate on the M2-1 ORF, and effectively backtrack to re-initialise translation on the M2-2 ORF.

1.5.2 Defective Interfering genomes

Defective interfering (DI) particles possess deleted genomes resulting from errors in replication. Such errors are caused by the polymerase complex creating either internal deletions or inverted terminal repeats, also referred to as copy-back DIs. Only copy backs duplicating the 3' end of the antigenome have been identified naturally. The necessary proteins, supplied by co-infecting helper genomes, replicate DI genomes. The DI particles interfere with helper virus growth as, being smaller, they replicate faster and remove proteins that would normally be used in the formation of infectious genomes (Re, 1991). DI particles were found in RSV preparations after four passages in Hep2 cells (Treuhaft & Beem, 1982).

1.5.3 Rescue of Non-Segmented Negative Strand RNA Virus

The rescue of negative strand RNA virus, referred to as reverse genetics, is a useful tool for the investigation of both genome and protein function. Collins *et al.* (1991) developed the RSV rescue system. Genome leader and trailer sequences were retained but viral genes were replaced by marker gene chloramphenicol transferase (CAT), effectively constructing an element that could be described as a DI genome. Investigation with the mini-genome confirmed that the genome-end sequences were sufficient to allow replication and transcription of the model genome (Collins *et al.*, 1991; Collins *et al.*, 1993). The ability to rescue virus entirely from cDNA was later achieved (Collins *et al.*, 1995). The technique produces infectious virus from cDNA-

encoded RNA and from cDNA-encoded viral genes the expression of which is driven by T7 polymerase supplied, normally, by a vaccinia virus recombinant. By this method, mutations can be introduced into cDNA encoding either RNA or proteins making it possible to create defined mutants for study and vaccine development.

1.6 Virus Replication and assembly

Following G protein-mediated adsorption to the target host cell and fusion of viral and cell membranes, the nucleocapsid is released into the cytoplasm. Primary transcription, which does not require concurrent protein synthesis, is performed by virion-associated polymerase complexes, carried into the infected cell by the nucleocapsid (Portner & Murti, 1986; Portner *et al.*, 1988).

1.6.1 Transcription

The ten genes of RSV are transcribed in their 3' to 5' genomic order. This suggests entry of the polymerase complex into the genome at a single upstream position. Using the reverse genetic system, Collins and co-workers have demonstrated that the 3' leader sequence is the promotor for RSV (Collins *et al.*, 1991; Collins *et al.*, 1995). Thus, the model of sequential transcription proposed in VSV and SeV may be applicable to RSV (Lamb & Kolakofsky, 1996). This model proposes that the viral polymerase associates with the 3' leader sequence of the genome. Transcription initiates within the leader region and the polymerase proceeds until it reaches the junction between the leader and first gene (NS1 in the case of RSV). Presumably, the leader is released and the polymerase transcribes the first gene until it reaches the gene-end sequence. The complex then initiates transcription of the next gene, thus the polymerase progresses along the length of the genome transcribing each of the genes in turn. Signals directing the capping and polyadenylation of the nascent mRNA, again achieved by the polymerase, are present within the gene-start and stop sequences (see below).

The RSV polymerase complex recognises gene-stop signals with a frequency of between 90 and 95%. In the remaining incidents, readthrough, or polycistronic, mRNAs, are generated (Fearns & Collins, 1999b). Only the first ORF present in a polycistronic transcript is translated (Wong & Hirano, 1987). This may be a means by which virus down-regulates (attenuates) expression of 5' proximal genes. Messenger

RNA abundance decreases with gene distance from the 3' leader sequence. This phenomenon is unrelated to mRNA half-lives and is due mainly to polymerase fall-off during sequential transcription, affected by increasing distance from the 3' promoter as has been shown for the *rhabdovirus* VSV. The overlap of 68 nucleotides existing between RSV genes M2 and L reduces transcription of this last gene. Fearns and Collins (1999a) propose that a proportion of polymerases that terminate at the M2 gene-end sequence, scan back along the genome to initiate transcription at the L gene-start signal.

1.6.2 Replication

RSV genome replication requires the concurrent synthesis of viral proteins to enable concomitant encapsidation of the nascent RNA. The same viral polymerase complex is responsible for both transcription and RNA replication. A model describing the change from transcription to RNA replication has been proposed following the study of VSV and SeV. Investigations suggest that there is a switch from the stop-start mechanism of transcription to an anti-termination read-through mode, producing full-length positive-sense antigenome. This is complementary to the genome, encapsidated with N, and behaves as a template for genome synthesis. The model proposed that N protein availability is a limiting factor, determining whether or not genome replication may occur (Lamb & Kolakofsky, 1996). It suggests that N is able to bind an encapsidation signal, present in the leader sequence of the negative-strand genome template. When N concentration is not limiting, the protein binds and therefore suppresses the termination signal, allowing the transcription of full-length antigenome and initiating nucleocapsid assembly. A similar termination signal is present within the promoter of the antigenome to ensure genome production does not occur in the absence of N protein. Upon its utilisation, N concentration decreases. When N is limiting, the protein does not bind the encapsidation sequence, therefore replicative RNA synthesis terminates at the leader junction and the polymerase reinitiates at the first gene and synthesises mRNAs (Blumberg *et al.*, 1981). The model of an N-controlled switch is very much debated. Upon its investigation in RSV, no evidence was obtained of a dynamic switch between transcription and RNA replication. Results indicated that the ratio between these processes remained relatively constant even when an excess of N was supplied which should, according to the model, favour replication (Fearns *et al.*,

1997). However, at present, this model remains a popular explanation of the mechanism of control between transcription and RNA replication.

1.6.3 Rule of Six

The study of the replication of some members of the *Paramyxoviridae*, using various rescue systems, indicated the importance of genome length. Work performed with defective interfering virus particles established that efficient replication was achieved only when genome length was a multiple of six. Calain & Roux (1993) found that the addition or deletion of a single nucleotide destroyed viral replication in the SeV system; replication was restored by the addition of nucleotides returning the genome to a multiple of six. The phenomenon was referred to as the "Rule of Six", and can be applied to the *Morbillivirus* genus e.g. MV (Radecke *et al.*, 1995) and to a lesser extent the *Rubulavirus* genus e.g. SV5 (Murphy & Parks, 1997). The situation for RSV appears to be different, as it shows no requirement for its genome to be a multiple of six or any other integer (Samal & Collins, 1996). Vulliemoz & Roux (2001) postulate that where the Rule of Six applies, the interaction of six nucleotides, in context with the N protein, is critical for the presentation of the viral promoter to the polymerase. This may be a reflection of differences in the organisation between the 3' viral promotor of the *Pneumovirinae* and the *Paramyxovirinae*. The *Paramyxovirinae* promotor is effectively bi-partite, with one element in the leader region and the other within the N gene (i.e. downstream of the leader/N gene junction). It is suggested that the inter-relationship between the correct number, and therefore correct spacing of nucleotides, and the N protein bring the two elements into some form of functional alignment. Incorrect spacing, by the addition or deletion of nucleotides, disrupts the alignment and has a deleterious effect on functionality. However, the *Pneumovirinae* promoter is mono-partite, as shown by the rescue system (Collins *et al.*, 1991). It is found solely within the leader region and can accommodate changes in length, as it does not have to align with downstream sequences.

1.7 Regulatory RNA Sequences

The genome contains all the necessary signals to regulate the synthesis of RNA, either as mRNA or as genome/antigenome. Production of mRNA requires transcriptional modification (for example capping and polyadenylation) such that the viral messages are functionally identical to those of the cell. Since RSV replicates in the cytoplasm and

utilises it's own polymerase, the virus requires specific control signals to regulate RNA production and the switch from mRNA production to replication of the genome.

1.7.1 Extragenic RNA – Sequences at Genome Termini

Short, non-coding sequences are present at both the 3'- (leader) and 5'- (trailer) termini of the paramyxovirus family genome; these contain the viral promotors and the genome encapsidation signal. In RSV, the leader sequence is 44 nucleotides long and the trailer sequence, 155 nucleotides in length. The extreme 3'-terminal and 5'-terminal sequences of the leader and trailer share 81% complementarity (21/26 bases may be aligned) (Mink *et al.*, 1991) suggesting that they base pair or share a common function. The first 12 nucleotides of the leader region are well conserved between the paramyxoviruses and the pneumoviruses. This sequence conservation suggests the region may be of functional importance. The ability of the trailer region to act as a promoter to produce genome RNA was demonstrated by the deleterious effects of mutating this region (Fearns *et al.*, 2000; Peeples & Collins, 2000).

1.7.2 Intragenic RNA – Gene-Start and Gene-Stop Sequences

A characteristic common to all members of the *Paramyxoviridae* family is the presence of non-coding sequence surrounding the open reading frame (ORF) within each gene. That occurring at the 3'-end of the gene is described as the gene-start sequence and the region present at the 5'-end, as the gene-end sequence. Both regions are transcribed into the mRNA.

The RSV gene-start sequences are nearly identical. The nine nucleotide sequence 3'-CCCCGUUUA-5' is present at the beginning of each gene, with the exception of L, where it differs by two nucleotides (Collins *et al.*, 1986). Transcription by the polymerase complex begins at the first nucleotide of this sequence and continues to the gene-end sequence, a semi-conserved signal that contains a short stretch of U residues. The gene-end sequence directs polyadenylation and transcription termination. It is suggested, by analogy to SeV and VSV, that the polymerase complex stutters upon transcription of the U region producing the polyadenylated tail of the mRNA. This proposed reiterative copying produces a tail of approximately 200 adenosine (A) residues (Lamb *et al.*, 1996). Analysis of the gene-start and gene-end

signals has shown that mutations can adversely affect the replication of recombinant mini-genomes (Kuo *et al.*, 1996b).

1.7.3 Intergenic RNA – Non-Coding Sequence between Genes

There are intergenic regions of non-coding RNA present in the *Paramyxoviridae* family genome (Amesse *et al.,* 1982). Among members of the *Morbillivirus* and the *Respirovirus* genera, these regions consist of a conserved three nucleotide sequence (3'-GAA-5'). Such conservation suggests they may be regulatory signals. The rubulaviruses do not possess intergenic sequence conservation. Within the *Pneumovirus* genus, the intergenic sequences are highly variable, ranging from one base (between the N and P genes) to 52 nucleotides (between the G and F genes) in length (Collins *et al.*, 1986). The intergenic regions are not transcribed into mRNA, though they are passed over by the polymerase complex without its becoming detached from the genome. Excepting the requirement for at least one base, the variation in length of RSV intergenic regions does not appear to affect the transcription of the proceeding gene (Kuo *et al.*, 1996a).

1.8 Virion Assembly

The assembly of the nucleocapsid occurs in the cell cytoplasm. The nucleocapsids are transported to the plasma membrane where they associate with the envelope proteins. The mechanism of virus particle assembly at the cell surface is thought to be mediated by the M protein though this remains to be determined. Virion maturation appears to occur in circumscribed regions of the cell surface. This may reflect dependence of virus budding from the cell membrane on localised sub-cellular structures (for example, lipid rafts) or the clustering of virus components (P. Yeo, personal communication). Teng & Collins (1998) demonstrated that N, P, L, M and F constituted the minimal requirements for packaging of the genome into an infectious particle.

1.9 The Protein components of RSV

There are ten major mRNA products of the RSV genome. Of the 11 resulting proteins (the M2 gene gives rise to two proteins), eight are known to be virion-associated, structural proteins. These include the transmembrane surface proteins G, F and SH, responsible for host cell adsorption and virus entry. The non-glycosylated M protein

forms the virion matrix beneath the viral membrane. The N, P, L and M2-1 proteins are associated with the genomic RNA to form the viral nucleocapsid and are the functional components of the polymerase. Protein distribution within the virus particle of RSV is demonstrated in Figure 1.2. The two remaining mRNA products are the non-structural proteins, NS1 and NS2, present only in trace amounts in the virion (if at all).

1.10 The Viral Envelope Proteins

1.10.1 The Fusion (F) Protein

The fusion protein of RSV is related in both structure and function to its counterparts in the *Paramyxovirinae* subfamily and provides perhaps the best reason for grouping pneumoviruses into the same taxonomic family (Collins *et al.*, 1984; Spriggs *et al.*, 1986). RSV F is similar in size (540-580 amino acids, aa) to those of the paramyxoviruses though it possesses low sequence relatedness with these proteins. The F proteins of this family resemble one another in the location of hydrophobic domains and the possession of an internal cleavage peptide. As suggested by the name, the F protein mediates the fusion between the viral and host cell membranes, enabling cell entry. It is also responsible for syncytium formation during host infection.

The F protein is a Type I glycoprotein with a cleaved amino-terminal signal sequence and a membrane anchor near the carboxy-terminus. The protein spans the membrane and possesses a short cytoplasmic tail. It is synthesised in the rough endoplasmic reticulum (RER) as the precursor polypeptide F0. Within the RER, F0 is co-synthetically modified by the addition of N-linked sugars and is assembled into a trimer. This precursor is cleaved, possibly in the *trans*-Golgi network, by the cellular protease furin, a subtilisin-like endoproteinase, creating the biologically active protein form (Sugrue *et al.*, 2001). Cleavage generates a short amino-terminal F2 peptide and the longer carboxy-terminal F1 peptide, linked by disulphide bridges. The newly exposed amino-terminus of F1 is termed the fusion peptide and is well conserved among the *Pneumovirinae* but not between sub-families, although all are highly hydrophobic. The hydrophobic residues are involved in the membrane insertion function of the protein and mediates the fusion event (Gonzalez-Reyes *et al.*, 2001; Zimmer *et al.*, 2001).

1.10.2 The Attachment Protein (G)

The viral attachment protein G binds the receptor on the target cell membrane, facilitating adsorption of the virus. This was confirmed by the prevention of virus adsorption by a G protein-specific Mab (Levine *et al.*, 1987). The cellular receptor for RSV *in vivo* is, as yet, unknown but within tissue culture, heparin has been implicated as a ligand for G protein dependent cell-surface attachment (Krusat & Streckert, 1997).

The G protein (298 aa) possesses no significant sequence relatedness or structural similarity with those of the paramyxovirus subfamily (Wertz *et al.*, 1985) and has neither neuraminidase nor haemagglutinating activity. The G protein is a Class II glycoprotein with an extracellular carboxy-terminal domain and an uncleaved signal-anchor sequence at its amino-terminus. The RSV G protein is palmitylated and is heavily glycosylated with both N- and O- linked sugars, possessing a higher carbohydrate content than its counterparts in the paramyxoviruses. Both N-linked glycosylation and the assembly of polypeptide into homotrimers occur within the RER. Addition of O-linked sugars takes place in the *trans*-Golgi network. The carbohydrate may serve to shield the polypeptide chain and prevent its recognition by the host (reviewed by Melero *et al.*, 1997).

Approximately 15% of the RSV G protein is expressed in a soluble form. It is created when translation initiates from an internal AUG start codon and subsequent post-translational removal of amino-terminus amino acids (Hendricks *et al.*, 1988). Secretion of G protein may influence the host immune response in a manner favourable for the virus although there is no evidence as to what.

1.10.3 The Small, Hydrophobic (SH) Protein

SH is a small glycosylated envelope protein, with a carboxy-terminal extracellular domain and an amino-terminal membrane anchor. This Type II envelope protein contains a single internal hydrophobic domain and possesses potential sites at both the amino- and carboxy-termini for the addition of carbohydrate. A variety of SH protein species accumulate within infected cells, differing in their extent of glycosylation and their use of internal start codons (Collins & Mottet, 1993). The significance of these different forms is unknown though it appears that only certain glycosylated derivatives are packaged within the virion. The function of the RSV SH protein is unknown though

it has been shown to greatly enhance fusion *in vitro* when co-expressed with G and F proteins (Pastey & Samal, 1997). However, within the context of the virus, when recombinant RSV with the SH gene deleted was propagated in tissue culture, no effect on the fusion potential of the virus was observed. However, in an animal model, such viruses are attenuated (Whitehead *et al.*, 1999).

1.10.4 The Matrix (M) Protein

M is the most abundant virion protein. RSV M is a basic, non-glycosylated protein of 256 aa in length. It lacks sequence relatedness with the matrix proteins of the paramyxoviruses. However, they share the common feature of a hydrophobic domain in the carboxy-terminal half and presumed similarity of function. The hydrophobic domain is insufficient in length to span the lipid membrane. Fractionation studies suggest that M protein is peripherally associated with the viral envelope, possibly located in an electron dense layer beneath this membrane (Lamb & Kolakofsky, 1996). M protein is believed to mediate interaction between the nucleocapsid, envelope proteins and the lipid bilayer, behaving as the central organiser of virion assembly and morphology. In Sendai Virus, M protein interacts specifically with membrane expressing F and HN, indicating an interaction with the glycoprotein cytoplasmic tails (Sanderson *et al.*, 1993). In VSV, interaction between M and the nucleocapsid appears to arrest RNA synthesis, forcing the nucleocapsid into the dormant state observed in the virion (Carroll & Wagner, 1979); similar behaviour may occur with the *Paramyxoviridae* M proteins.

1.11 The Nucleocapsid-Associated Proteins

1.11.1 The L Protein

The RSV large protein L is the least abundant of the structural proteins. It is similar in size to the L proteins of paramyxoviruses and rhabdoviruses (approximately 2200 aa) but shares little overall sequence homology. However, there are five regions of high homology near the protein centre. These conserved domains are also present within the RNA-dependent RNA polymerases of other virus families. It is not within the scope of this thesis to give a full discussion on the complexity of L, suffice to say that it contains the necessary enzymatic functions for the production of RNA, both genome and message. The protein has been assigned functions such as methylation, capping and polyadenylation (Collins *et al.*, 1996b). As with the L proteins of all

Mononegavirales, there is a requirement for an accessory factor, the Phospoprotein (see Section 1.11.4). The template for activity is the nucleocapsid, naked RNA will not serve.

1.11.2 The Non-Structural Proteins

The non-structural proteins NS1 and NS2 are present in certain members of the *Pneumovirinae*, including RSV, but are absent from the *Paramyxovirinae* subfamily. NS1 and NS2 genes are located adjacent to the leader sequence and encode polypeptides of predicted length 139 aa and 124 aa, respectively (Johnson & Collins, 1989). The proteins accumulate within the cytoplasm of infected cells and have not been detected within the virion (Evans *et al.*, 1996). At present, the functions of NS1 and NS2 are not known. NS1 has been shown to be a negative regulator of genome replication and transcription (Atreya *et al.*, 1998). However, neither NS1 or NS2 are essential for virus replication (Teng *et al.*, 2000). Recent work on the NS1 and NS2 of bRSV has implicated them in playing a role in combating the IFN response (see Section 1.2.5).

1.11.3 The M2-1 and M2-2 Proteins

As mentioned previously, the M2 gene gives rise to two protein species M2-1 and M2-2, both of which are translated from the same mRNA (Ahmadian *et al.*, 2000). The M2-1 protein was originally identified as a fourth membrane protein, with supposedly, a matrix-type function (Huang *et al.*, 1985). However, this is no longer the case. M2-1 has been ascribed a function as an elongation factor necessary for the efficient production of long RNAs during replication and transcription (Fearns & Collins, 1999b). M2-2 was shown to be a down-regulator of RNA synthesis (Collins *et al.*, 1996a). Subsequently, Bermingham & Collins (1999) implicated M2-2 in the switch between transcription and replication. This has obvious implications for the model, discussed in Section 1.6.2 in which N protein concentration regulates the switch. However, RSV N is significantly smaller than those in the *Paramyxovirinae*, thus M2-2 may supply a function normally associated with the N protein.

1.11.4 The P Protein

The P protein of RSV is highly phosphorylated and is smaller than its counterparts within the *Paramyxovirinae* subfamily (241 aa compared with 392 aa for SV5 and 568

aa for SeV) (Morgan, 1991). By analogy to the SeV P, the RSV P protein may perform two functions, one as a transcription and replication factor, in complex with L, and the other as a chaperone to unassembled nucleocapsid protein N (P-N^0), maintaining the latter in a soluble, active form preventing illegitimate nucleocapsid assembly of non-viral RNA (Curran *et al.*, 1993; Curran *et al.*, 1995).

The RSV P protein is modular in organisation with the separation of amino- and carboxy-terminal domains by a hypervariable region. Domains responsible for P protein oligomerisation, association with the L protein and for binding N protein in the assembled (i.e. the nucleocapsid), and possibly unassembled, form appear to be within the proteins carboxy-terminus (Garcia-Barreno *et al.*, 1996; Khattar *et al.*, 2001; Slack & Easton, 1998). All studies agree that the carboxy-terminus is essential to enable P protein to function in the interaction with N, notably when P behaves as a transcription co-factor. The role of the amino-terminus is less clear but it may participate in the chaperone function, as proposed for SeV (Curran *et al.*, 1995).

1.11.5 The N Protein

N is the most abundant protein in the nucleocapsid core and is associated with the P, L and M2-1 proteins (Garcia-Barreno *et al.*, 1996; Khattar *et al.*, 2001; Samal *et al.*, 1993). The N protein performs a number of functions. Its major role is the encapsidation of the viral genome generating the RNase-resistant nucleocapsid. N tightly binds the RNA along its entire length, providing support, and protection from RNAses, of the phosphodiester backbone of the genome, yet enabling the continuation of transcription and RNA replication (Lamb & Kolakofsky, 1996).

The human RSV N protein was sequenced by Collins *et al.* (1985). It is smaller than the *Paramyxovirinae* counterparts, being only 391 aa in length and has a mass of approximately 43 kDa. Upon sequence comparison of human and bovine strains of RSV, N is observed to be a highly conserved protein with 93 to 94% sequence identity at the amino acid level (Samal *et al.*, 1991). There is relatively low N sequence homology with other non-segmented negative strand viruses. However, alignment of the RSV N protein with the N-terminal 400 aa of *Paramyxovirinae* N proteins shows the presence of three segments of limited sequence similarity. These sequences are semi-conserved and appear to share a predicted secondary structure. This sequence similarity suggests the region is of fundamental importance to protein function (Morgan, 1991).

1.12 N proteins and nucleocapsid assembly

Many nucleocapsid protein sequences of the negative-strand RNA viruses have now been determined. Within the *Paramyxoviridae*, little more than 50% identity exists between N proteins, though conservation is greater within genera. In the *Paramyxovirinae*, N proteins range in length from 489 to 553 aa and are between 53 and 58 kDa in size (Morgan, 1991). With the exception of Mumps Virus N protein, their net charges fall within the range of -7 to -12. This is an unexpected observation of proteins known to associate with negatively-charged RNA. Several RNA-binding protein families have been identified, each possessing a common sequence motif believed to interact with RNA. A family may have representatives in a diverse range of organisms ranging across viruses, prokaryotes and eukaryotes (Mattaj, 1993). The N proteins of the *Mononegavirales* are not classical RNA-binding proteins and lack recognised RNA-binding motifs. Leader regions of the *Mononegavirales,* which presumably contain the putative encapsidation signal, have no secondary structure. It is the primary RNA sequence that is responsible for binding N (Grinnell & Wagner, 1984; Hoffman & Banerjee, 2000). When expressed alone, N proteins do not appear to have sequence specificity and may bind cellular RNAs forming structures that are morphologically identical to nucleocapsids produced during normal viral infection (Meric *et al.,* 1994; Myers *et al.,* 1997; Spehner *et al.,* 1997. Warnes *et al.,* 1995) The specificity of RNA binding would appear, at least for SeV and MV, to be a function of the N-P complex (Curran *et al.,* 1995).

1.13 The nucleocapsids of the *Paramyxoviridae*

Despite their importance, very little ultra-structural information is available for *Paramyxoviridae* nucleocapsids, visualised by electron microscopy as single left-handed helices (Compans *et al.,* 1972). Studies on SeV show that its nucleocapsid varies extensively in the coiling and flexibility of the helical structure (Egelman *et al.,* 1989), which may be affected by salt concentration (Heggeness *et al.,* 1980). Upon salt-induced conformational change, accessibility of the N protein to trypsin cleavage alters. At low salt concentrations, the nucleocapsid has an extended conformation, leaving the N protein accessible to protease digestion. Upon the partial degradation of N, the SeV nucleocapsid structure remains intact. This is mediated by regions within the amino-terminal domain of the protein. Through-out investigations, the viral genome remained protected from RNase digestion. This demonstrated that neither the tightness

of the helix nor the proteolytic cleavage in N protein affected the binding of RNA by N or its protection from ribonuclease digestion (Heggeness *et al.*, 1980; Heggeness *et al.*, 1981). The ability to vary the nucleocapsid structure may be important in the mechanism of RNA transcription and replication. It was determined that proteolytic removal of the carboxy-terminus prevented RNA synthesis from the nucleocapsid.

1.14 Aims of this study

Within the Institute of Virology, RSV research was a new initiative thus there was no historical background upon which to build. Although C. Pringle worked on RSV in the early eighties, most of his reagents were not available. Paul Yeo instigated several areas of research developing his interests in the RSV proteins N and M and in virus assembly. The long-term objectives of the group were to achieve structural determination of the proteins above and to map biological interactions. Therefore, a feature of the work presented herein was to develop reagents and protocols that may be of use in future structural and biological investigations of the RSV N protein. Several areas were targeted as being of interest and have been addressed within this thesis.

- The purification of recombinant N protein, suitable for *in vitro* analysis
- Determination of the N protein's sequence specificity for binding RNA
- Identification of the RNA binding domain on the N protein
- Effect of mutations on assembly of nucleocapsid-like structures in bacteria

2 Materials and Methods

2.1 Statement of Safety

Tissue culture and work with both Respiratory Syncytial Virus and with Baculovirus were restricted to a sterile Class II flow cabinet. The cabinet was wiped with 70% (v/v) ethanol before and after use and was sterilised by UV light. Bacterial work was performed in a Class II laboratory.

Used growth medium and buffers were neutralised with 1% (w/v) virkon and incubated overnight prior to being discarded. Used glassware was steeped in 1% (v/v) chloros overnight prior to being washed and autoclaved.

All cell and chemical reagent work was performed wearing protective gloves and laboratory coat to minimise personal contact. When working with radioactivity, a double layer of gloves and a laboratory coat were worn. Radioactive work was performed behind Perspex shielding.

2.2 Source of Chemical Reagents

Tissue culture reagents were obtained from GibcoBRL (Scotland).

BDH (Poole, England) provided the majority of laboratory chemicals. Remaining chemicals were obtained from Sigma, R.P. Normapur A.R. Prolabo (France) and from Boehringer Mannheim, GmbH (Germany).

2.3 Sources of Enzymes

Restriction endonucleases were obtained from New England Biolabs. Additional enzymes were provided by Promega and by Sigma.

2.4 Sources of Antibodies

Antibodies directed against various RSV proteins were generated in the laboratory. Those frequently used in the project are listed in Table 2.1. Purified antibodies directed against Histidine-tag were obtained from Sigma.

Antibody Designation	Antibody Type	Species	RSV Protein Specificity
N003	Monoclonal	Mouse	N
N009	Monoclonal	Mouse	N
N015	Monoclonal	Mouse	N
P001	Monoclonal	Mouse	P
M024	Monoclonal	Mouse	M
22k/302	Monoclonal	Mouse	M2-1
P658	Polyclonal	Rabbit	P
N14	Polyclonal	Rabbit	N
N15	Polyclonal	Rabbit	N

Table 2.1: RSV-specific antibodies generated in the laboratory and used during the project.

2.5 Tissue Culture

2.5.1 Maintenance of Cultured Cells

The laboratory cell line CV-1, derived from African green monkey kidney, was used to culture respiratory syncytial virus. Cells were cultivated as a monolayer in T180 flasks and maintained in Dulbecco's Modified Eagle Medium (DMEM) with 2 mM L-glutamine, 1% (w/v) non-essential amino acids, 1% (w/v) penicillin and streptomycin antibiotics and 10% (v/v) heat-inactivated foetal calf serum (FCS). The cells were incubated at 37°C, 5% (v/v) carbon dioxide (CO_2). Cells were passaged at 80-100% monolayer confluency. The cell layer was washed with 0.25% (w/v) trypsin (Sigma) in versene. The supernatant was discarded and the flask incubated for three to five minutes. The loosened cells were resuspended in 10 ml fresh 10% FCS growth medium. The desired volume of cell suspension was placed in a new T180 flask and combined with fresh growth medium. Cells were returned to incubation at 37°C, 5% CO_2.

2.6 Virus Growth and Maintenance

2.6.1 Preparation of Respiratory Syncytial Virus

RSV strain A2 was provided by C. Pringle, Warwick University, and was used throughout the project. A CV-1 cell monolayer was grown to 70% confluency in a T180 flask. The monolayer was washed twice with phosphate-buffered saline (PBS). RSV was added to the monolayer in growth medium with 2% FCS at a multiplicity of infection (m.o.i.) of 0.1. The flask was incubated for one to two hours at $33^{\circ}C$ to enable virus adsorption to the cells. The inoculum was removed and the cells washed twice with PBS. 20 ml fresh 2% FCS medium was added to the flask and the infected cells incubated for three to five days at $33^{\circ}C$, 5% CO_2 until cytopathology, in the form of fused cells, was evident. To harvest virus from the infected culture, the medium was first removed. 10 ml fresh medium, followed by half a bijoux of sterile glass beads (2.5 mm diameter) were poured into the flask. The flask was shaken vigorously for approximately one minute to cause cell lysis and disrupt cell-associated virus. The resulting supernatant was decanted and cell debris and glass beads removed by centrifugation for 10 minutes at 3000 xg. The clarified supernatant was divided into 1 ml aliquots and stored at $-70^{\circ}C$.

2.6.2 Measurement of Virus Titre by Plaque Assay

CV-1 cells cultured in growth medium in a six-well plate (35 mm diameter) were grown to 80-90% confluency. The medium was removed and the cells washed twice with sterile PBS. Ten-fold serial dilutions of the virus sample were added to each well in 1 ml 2% FCS growth medium. The cells were incubated for one to two hours at $33^{\circ}C$, 5% CO_2. The inoculum was removed and the cells washed with PBS. An overlay mixture of 2% FCS, 0.8% Noble agar and DEAE-dextran (100 μg ml^{-1}) in DMEM was prepared and maintained at $42^{\circ}C$. Three millilitres of overlay mixture was added to each well and allowed to set at room temperature. The plate was then inverted and incubated for five to seven days. Three millilitres glutaldehyde fixative (50% Cidex in PBS) was added to each well in a fume hood. The plate was incubated for three to four hours at room temperature. The fixative and agar were removed by washing with water. The cells were stained with giemsa stain (BDH) for one hour at room temperature. The stain was removed by washing with water. The wells were examined under the light microscope and the number of virus plaques, visible by the absence of stain, was recorded.

2.6.3 Immunofluorescence (IF) of Virus-Infected Cells

Sterile glass cover-slips (13 mm diameter) were placed in individual wells of a 24-well plate. $4x10^4$ CV-1 cells were seeded in each well in a volume of 1 ml 10% FCS growth medium. The plate was incubated overnight at $37^\circ C$, 5% CO_2. Upon 80% confluency of the cell layer, the medium was removed and the cells washed twice with PBS. RSV was added to each well at m.o.i. 1.0 in 200 μl 2% FCS growth medium. The cells were incubated for one to two hours at $33^\circ C$, 5% CO_2. The virus medium was removed and the cells washed with PBS. One millilitre 2% FCS growth medium was added to each well and the cells, incubated for 16 to 20 hours at $33^\circ C$, 5% CO_2. The medium was removed and the cells washed with PBS. The cells were fixed with cold methanol (-$20^\circ C$) and incubated at -$20^\circ C$ for up to 20 minutes or overnight. The methanol was removed and the cells washed three times with PBS. Antibodies generated in the laboratory and directed against various RSV proteins are listed in Table 2.1. Purified RSV monoclonal antibodies were diluted 1/1000 in PBS with 1% FCS. RSV polyclonal antibodies were diluted 1/500 in PBS with 1% FCS. Antibodies were added individually or in combination to the wells. The plate was incubated for one hour at room temperature. Primary antibody was removed and the cells washed with PBS with 1% FCS. The appropriate anti-species antibody conjugated to FITC or TRITC, diluted 1/200 in PBS with 1% FCS, was added to each well. The plate was incubated for one hour at room temperature in the dark. Antibody was removed and the cells washed three times with PBS. The cover-slips were removed from each well, allowed to air-dry and mounted on glass slides with Citifluor. Immuno-staining was visualised under a fluorescence microscope (Nikon) or by confocal microscopy (courtesy of J. Murray).

2.6.4 Labelling of RSV-infected CV-1 Cells with [^{35}S]-Methionine

CV-1 cells, grown as a monolayer in a T180 flask, were infected with RSV at m.o.i. 0.1 as described in Section 2.6.1. At 48 hours post-infection, the cell monolayer was starved of methionine for one hour by incubating in methionine-free growth medium (Gibco), followed by the addition of [^{35}S]-methionine at 100 μCi ml^{-1}. Cells were incubated for three hours then harvested into RIPA buffer and clarified by centrifugation at 10 000 xg. The supernatant was pre-cleared using Pansorbin cells (Calbiochem) and a non-specific antibody.

2.6.5 Radiolabelled Immunoprecipitation Assay

[^{35}S]-methionine-labelled RSV-infected CV-1 cells were used in a radiolabelled immunoprecipitation assay (RIPA) with antibodies directed against RSV proteins. One microlitre purified mono- or polyclonal antibody was added to a 50 μl aliquot of radiolabelled, infected cell lysate (Section 2.6.4) and incubated for one hour at 4°C with shaking. Sepharose-bound protein A was used to precipitate antibody-antigen complexes from solution. Protein A associates more strongly with rabbit- rather than with mouse-derived antibodies. An extra labelling step was therefore included with the addition of rabbit anti-mouse (Sigma) to each sample and further incubation for one hour, 4°C with shaking. 20 μl protein A (50% (w/v) in lysis buffer) was added and the samples incubated for one hour, 4°C with shaking. The protein A complexes were pelleted by centrifugation for one minute at 6000 xg. The supernatant was discarded and the pellet resuspended in 200 μl lysis buffer. This process was repeated five times. The pellet was combined with sample loading buffer, heated for five minutes at 95°C and separated by SDS-PAGE (Section 2.10.7).

2.6.6 Extraction of Cellular and Viral RNA

The Total RNA Isolation Kit (Ambion) was used to extract cellular and viral RNA from RSV-infected CV-1 cells. The growth medium from a T180 flask of infected CV-1 cells was discarded and the cells washed three times with PBS. Ten millilitres denaturing solution was added and the flask, agitated for approximately three minutes. The resulting supernatant was combined with 0.1 volume of 3 M sodium acetate and one volume of phenol:chloroform (1:1) solution, shaken vigorously for one minute and placed on ice for 15 minutes. The mixture was subjected to a centrifugal force of 12 000 xg for 15 minutes. The aqueous layer (top) was retained and combined with one volume of acid phenol:chloroform (1:1) solution. The sample was again shaken and centrifuged as above. The resulting aqueous layer was combined with one volume iso-propanol and incubated overnight at -20°C. The sample was subjected to a centrifugal force of 7500 xg for 10 minutes to pellet the RNA. The supernatant was discarded and the pellet washed with 70% (v/v) ethanol and allowed to air-dry. The RNA pellet was resuspended in 1 ml de-ionised water (dH$_2$O) and its optical density (O.D.$_{260}$) determined (Section 2.9.7).

2.6.7   Recombinant Baculovirus Infection in Sf21 Insect Cells

Sf21 insect cells were grown in a monolayer in T180 flasks and maintained in TC100 growth medium with 10% FCS at 27°C, 5% CO_2. Upon confluency, cells were passaged by knocking the adherent cells into 10 ml growth medium and reseeding 1 ml into a fresh T180 flask. Baculovirus Bac-N_{wt} was obtained from P. Yeo. The recombinant virus expresses the N protein gene of RSV strain A2 under the control of the baculovirus polyhedrin promoter. Sf21 cells were infected with Bac-N_{wt} at an m.o.i. of 5-10 and maintained at 27°C, 5% CO_2 for 72 hours before harvesting the protein (Section 2.12.1.2).

2.7   **Bacterial Work**

2.7.1   Strains of Bacteria

Various laboratory strains of *Escherichia coli* (*E.coli*) were employed during the course of the project (Table 2.2). Strains DH5 and JM109 were used for the purpose of plasmid DNA amplification and maintenance. BL21(DE3)pLysS and NovaBlue (DE3) strains were used for the expression of recombinant protein. Strain BMH 71–18 mutS is tolerant of initial plasmid strand-mismatch. It was used in site-directed mutagenesis to obtain mutant DNA.

Bacterial Strain (source)	Genotype
DH5 (Stratagene)	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*
BL21(DE3)pLysS (Novagen)	F⁻ *ompT hsdS$_B$($r_B^-$$m_B^-$) gal dcm (DE3)* pLysS (Cm^R)
NovaBlue (DE3) (Novagen)	*endA1 hsdR17($r_{K12}^-$$m_{K12}^+$) supE44 thi-1 recA1 gyrA96 relA1 lac [F' proA⁺B⁺ lacI^{q}ZΔM15 ∷Tn10(Tc^R)] (DE3)*
BMH 71-18 mutS (Promega)	*thi supE Δ(lac-proAB) [mutS ∷Tn10] [F' proAB laqI^{9}ZΔM15*
JM109 (Promega)	*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi Δ(lac-proAB) F'[traD36 proAB⁺ lacI^q lacZΔM15]*

Table 2.2 : Laboratory strains of *E.coli* employed during the course of the project.

2.7.2   Growth and Maintenance of Bacteria

Routinely, bacteria were grown on L-broth agar or in 2YT liquid medium. Bacterial samples were either streaked using a sterile loop or evenly spread over the surface of an L-broth agar plate. Antibiotics were frequently included to select for the growth of bacteria possessing plasmid containing a particular resistance gene. Bacterial plates were inverted and incubated for 16 to 20 hours at 37°C. To isolate a single bacterial clone, an individual colony was picked from the plate using a sterile tooth-pick and was placed in 10 ml selective 2YT medium. The culture was incubated for 16 to 20 hours at 37°C with shaking at 200 rpm.

2.7.3   Storage of Bacterial Stocks

Bacterial stocks were prepared by combining an aliquot of the 10 ml culture grown from a single colony with an equal volume of sterile glycerol. Samples were transferred to a glass bijoux and placed at -70°C for long term storage.

2.7.4   Antibiotic Stocks

Stock ampicillin was prepared at 100 mg ml^{-1} in dH$_2$0 at -20°C; its working concentration was 100 µg ml^{-1}. Stock chloramphenicol was prepared at 34 mg ml^{-1} in ethanol at -20°C; its working concentration was 34 µg ml^{-1}.

2.7.5   Production of Competent Bacteria

2.7.5.1   Electro-competent Bacteria

A single bacterial colony was picked from an L-broth agar plate and used to inoculate 10 ml 2YT medium. The culture was incubated overnight at 37°C with shaking at 200 rpm. To generate bacteria competent for transformation by electroporation, 500 ml 2YT medium was inoculated with a dilution of 1/500 overnight culture and incubated at 37°C, 200 rpm until an O.D.$_{600}$ 0.6-0.8 was obtained. The culture was placed on ice for up to one hour and then subjected to a centrifugal force of 1700 xg for 10 minutes, 4°C. The supernatant was removed and the cell pellet resuspended in 300 ml ice-cold dH$_2$O. This procedure was repeated with 150 ml, 75 ml, 30 ml and 15 ml volumes of ice-cold 10% (v/v) glycerol. The volume of the final pellet was estimated and resuspended in an

equal volume of 10% glycerol. The resulting competent bacteria were stored at -70°C in 200 μl aliquots.

2.7.5.2 Calcium Chloride-Competent Bacteria

To generate bacteria competent for transformation by calcium chloride, 100 ml 2YT medium was inoculated with 1 ml overnight culture and incubated at 37°C, 200 rpm until O.D.$_{600}$ 0.5. The culture was placed on ice for 15 minutes and then subjected to a centrifugal force of 1700 xg for 15 minutes, 4°C. The supernatant was discarded and the cell pellet, resuspended in 50 ml ice-cold 100 mM calcium chloride. The bacteria were incubated on ice for 30 minutes to one hour followed by centrifugation. Steps were repeated with 10 ml and 1 ml 100 mM calcium chloride. The resulting 1 ml of competent cells was incubated on ice for 30 minutes prior to their immediate use.

2.7.6 Transformation of Competent Bacteria

2.7.6.1 Electroporation

40 μl electro-competent bacteria were combined with either ~0.1 μg plasmid DNA or 3 μl ligation reaction (Section 2.9.9) in an ice-cold cuvette with gap size 0.1 cm. The cuvette was tapped to dislodge air bubbles and placed in the ice-cold cuvette holder. The electroporater (BioRad) was set to the following parameters: Capacitance 25 μFD; Resistance 400 Ohms and Volts 1.6 kV. After application of the electric current, 1 ml 2YT broth was added to the cuvette and the bacteria transferred to a universal tube. The bacteria were incubated for one hour at 37°C, with shaking. One microlitre, 10 μl and 100 μl aliquots were spread on L-broth agar plates, with appropriate selective antibiotics. The plates were inverted and incubated overnight at 37°C.

2.7.6.2 Calcium Chloride-Heat Shock

0.5 μg plasmid DNA and 300 μl competent cells were combined and incubated for 30 minutes on ice. The mixture was transferred to 42°C for two minutes and then returned to ice for a further two minutes. 700 μl 2YT broth was added and the bacteria, incubated for one hour at 37°C, 200 rpm. Aliquots were spread on selective L-broth agar and the plates incubated, as previously described (Section 2.7.6.1).

2.7.7 Plasmid DNA Extraction

QIAprep Spin Miniprep Kit (Qiagen) was used to extract plasmid DNA from small volume cultures of transformed bacteria. The kit employs an alkaline method of bacterial lysis and captures plasmid DNA on a resin from where it is washed and subsequently eluted. A single bacterial colony was used to inoculate 10 ml selective 2YT medium. The culture was incubated for 16 to 20 hours at 37°C, 200 rpm. Between 1.5 and 6 ml culture was then subjected to a centrifugal force of 14 000 xg for five minutes. The supernatant was discarded and the pellet, resuspended in 250 μl Buffer P1 (with 100 μg ml⁻¹ RNase A). 250 μl lysis buffer (P2) was added and the reagents mixed by inverting the tube. Following a five minute incubation at room temperature, 350 μl neutralising buffer (N3) was added and the reagents again mixed by inverting. The mixture was subjected to a centrifugal force of 14 000 xg for 10 minutes to pellet cell debris. The resulting supernatant was retained and applied to a QIAprep spin column. The column was centrifuged for one minute and the flow-through discarded. 750 μl wash buffer (PE) was applied and the column centrifuged for one minute. The flow-through was discarded and the column was centrifuged for a further one minute to remove residual buffer. DNA was eluted by the addition of 50 μl dH₂0 to the column and centrifugation for one minute. The DNA was either used immediately or stored at -20°C.

2.7.8 Expression of His-Tagged Nucleocapsid Protein

E.coli strain BL21(DE3)pLysS (Novagen) was transformed with pET16b-Nucleocapsid gene plasmid DNA by calcium chloride-heat shock. An individual colony from the resulting transformation was picked and placed in 10 ml 2YT medium, with 10 μg ml⁻¹ ampicillin and 34 μg ml⁻¹ chloramphenicol. The culture was incubated for 16 to 20 hours at 37°C, 200 rpm. 100 ml selective 2YT medium was inoculated with overnight culture creating a 1/100 dilution. The culture was incubated at 37°C, 200 rpm until O.D.$_{600}$ 0.8 was obtained. Recombinant protein expression was induced upon addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and the culture incubated for a further three hours at 37°C, 200 rpm. The culture was subjected to a centrifugal force of 1700 xg for 15 minutes. The supernatant was discarded and the pellet processed for protein purification (Sections 2.10.1 and 2.12.1).

2.8 Yeast Three Hybrid System

2.8.1 Yeast Growth and Transformation

Recombinant yeast strain L40-coat was used in the yeast-three-hybrid assay. The strain possesses a gene encoding the MS2 bacteriophage coat protein fused to the Lex A DNA binding domain integrated into the chromosome. Daughter strains L3-2 RSV and L40-HCV were generated upon transformation of L40-coat with plasmids expressing RSV-MS2-2 and HCV 3'NCR-MS2-2 transcripts, respectively. The latter plasmid was kindly provided by J. Wood (MRC Virology Unit). Plasmids encoding RNA transcripts possess a gene enabling the yeast to survive in uracil-depleted medium (Ura⁻). Yeast strains were later transformed with plasmid expressing N-fusion protein. These plasmids encode a gene enabling yeast survival in the absence of leucine. Individual colonies of the transformed yeast strains L40-coat, L3-2 RSV and L40-HCV were picked, resuspended in 1 ml dH$_2$O and added to 50 ml aliquots YPD medium (L40-coat) or Ura⁻ medium (L3-2 RSV and L40-HCV). The cultures were incubated overnight at 30°C, 200 rpm. 30-40 ml overnight cultures were added to 300 ml YPD medium (O.D.$_{600}$ 0.2). The cultures were incubated at 30°C, 200 rpm until O.D.$_{600}$ 0.4-0.6 was obtained. The cultures were subjected to a centrifugal force of 1200 xg for five minutes at room temperature. The supernatant was discarded and the pellets, each resuspended in 200 ml dH$_2$O. The spin was repeated, the supernatants discarded and the pellets, each resuspended in 1.5 ml Lithium Acetate/TE buffer. 100 µl aliquots yeast suspension were combined with 1 µg plasmid DNA and 100 µg herring sperm DNA and vortexed. 600 µl 40% (w/v) PEG3000 Lithium Acetate/TE buffer was added and the mixture again vortexed. The yeast aliquots were incubated for 30 minutes at 30°C, 200 rpm. 70 µl dimethyl sulphoxide (DMSO) was added and the yeast transformed by heat-shock – transferral from 42°C for 15 minutes to ice for two to five minutes. The yeast was pelleted by centrifugation at 14 000 xg for five minutes. The supernatants were discarded and the pellets, each resuspended in 500 µl dH$_2$O. 100 µl aliquots of transformed yeast were spread on agar plates, with relevant supplements excluded from the medium. The yeast was incubated for two to four days at 30°C.

2.8.2 Colony Lift Filter Assay

The colony lift filter assay was performed to investigate the expression of the reporter gene β-galactosidase by transformed yeast. It's expression is indicative of an

association between the recombinant RNA and protein resulting in stimulation of the GAL4 activation domain. Colonies were transferred to Whatman #5 filter papers by placing papers on the yeast agar plates and smoothing over the surface. The filters were removed and submerged in liquid nitrogen for 10 to 30 seconds. The filters were thawed and then transferred colony-side up to blotting paper soaked in Z Buffer/X-Gal solution. The filters were incubated at room temperature for eight hours. The degree of colour development, resulting from reporter gene expression, was estimated by eye on a scale from zero to three.

2.9 DNA Manipulation

All nucleic acid samples were dissolved in either dH_2O or TE buffer. Samples were maintained at -20°C.

2.9.1 Reverse-Transcription (RT)

The First Strand cDNA Synthesis Kit for RT-PCR (Boehringer Mannheim) was used to generate cDNA from the total RNA extracted from RSV-infected CV-1 cells (Section 2.6.6). Random hexamers were used to prime cDNA synthesis.

10 x RT-PCR Reaction Buffer	2 µl
25 mM Magnesium Chloride	4 µl
Deoxyribonucleotide (dNTPs) Mix	2 µl
Random Primer p(dN)$_6$	2 µl
RNase Inhibitor	1 µl
(AMV) Reverse Transcriptase	0.8 µl
Gelatine	0.4 µl
dH_2O	6.8 µl
RNA Sample	1 µl / 20 µl

The reaction was incubated for 10 minutes at 25°C, followed by 60 minutes at 42°C. The sample was heated to 99°C for five minutes to denature the enzyme then placed on ice for a further five minutes.

2.9.2 Polymerase Chain Reaction (PCR)

The Expand High Fidelity PCR System (Roche) was used to amplify the viral nucleocapsid gene using specific primers (Appendix). The system employs the thermo-stable DNA polymerases Taq and Pwo. The latter enzyme has proof-reading abilities.

Mix 1		Mix 2	
200 µM	Each dNTP	10 x	Expand HF Buffer with 15 mM Magnesium Chloride
300 nM	Upstream Primer		
300 nM	Downstream Primer	2.6 units	Enzyme Mix
0.01-0.1 µg	Template DNA		Up to 25 µl dH₂O
or			
1µl RT products			
Up to 25 µl dH₂O			

Mixes 1 and 2 were assembled individually and combined immediately before use. The sample was incubated for three minutes at 94°C in a Hybaid Touchdown thermo-cycler. The sample was then incubated for 15 seconds at 94°C, 30 seconds at 50°C and 90 seconds at 72°C for 30 cycles. After five cycles, the annealing temperature was reduced by 0.5° per cycle. The reaction was finally incubated for 10 minutes at 72°C. PCR products were analysed by agarose gel electrophoresis.

2.9.3 Agarose Gel Electrophoresis

Agarose and TAE buffer were combined creating a gel of the required percentage (frequently 1 to 2%). The mixture was heated until the agarose dissolved. Ethidium bromide (10 µg ml⁻¹) was then added. The mixture was poured into a gated mini-gel tank (BioRad), the comb put in place and the gel allowed to set. The comb and gates were removed and the gel, immersed in TAE buffer. DNA sample was combined with the appropriate volume of loading buffer and loaded into a well. The gel was subjected to a potential difference of 150 V for 10 to 15 minutes. The gel was removed from the apparatus and viewed under ultra violet (UV) light (280 or 302nm).

2.9.4 DNA Isolation and Extraction from Agarose

The Gene Clean Kit (BIO 101) was used to extract DNA bands from agarose gel. This was performed to isolate DNA of a specific size and to remove nucleotides and primers residual from synthesis steps. The DNA sample was separated by agarose gel electrophoresis and visualized under long wave UV light (302 nm). Keeping UV exposure time to a minimum, the desired DNA band was excised from the gel with a clean scalpel. The gel slice was combined with 400 µl GC Spin Glassmilk® in a Spin™ filter and heated to 55°C for five minutes, inverting the tube every minute to prevent settling of the matrix. The tube was subjected to a centrifugal force of 14 000 xg for one minute and the resulting liquid in the catch tube was discarded. 500 µl GC Spin Wash was added and the tube centrifuged for 30 seconds. The liquid was discarded and the wash, repeated. The tube was centrifuged for one minute to remove remaining liquid from the filter. The Spin™ filter was transferred to a fresh catch tube. 25 µl GC Spin Elution Solution was added and the Glassmilk® resuspended by flicking the tube. The tube was centrifuged for 30 seconds and eluted DNA collected.

2.9.5 Phenol:Chloroform Extraction of Nucleic Acid

Phenol:chloroform extraction was performed to remove contaminating protein from nucleic acid. Addition of iso-amyl-alcohol (IAA) further removes lipids.

The DNA sample was increased to a volume of 100 µl with dH$_2$O. An equal volume of phenol (pH >6) (Sigma) was added and the sample vortexed for 30 seconds then placed on ice for one minute. The sample was subjected to a centrifugal force of 14 000 xg for one minute. The aqueous layer (top) was carefully removed and transferred to a fresh tube. 100 µl dH$_2$O was added to the phenol. The mixture was vortexed, placed on ice and centrifuged to back extract remaining DNA. The second aqueous layer was combined with the first extraction. This procedure was repeated. A single extraction with chloroform:IAA (24:1) was then performed. The extraction of RNA follows the same procedure excepting the use of acidic, in place of alkaline, phenol (i.e. pH <6).

2.9.6 Ethanol Precipitation of Nucleic Acid

Ethanol precipitation removes salt from a nucleic acid sample and enables it's resuspension in the desired buffer and volume.

The sample was combined with 0.1 volume 3 M sodium acetate (pH 4.8) and 2.5 volume ethanol and incubated at -20°C for one hour. The sample was exposed to a centrifugal force of 14 000 xg for 15 minutes, the supernatant discarded and the pellet, washed with 70% (v/v) ethanol. The sample was again centrifuged, the supernatant discarded and the pellet allowed to air-dry. The pellet was resuspended in dH_2O.

2.9.7 Determination of Nucleic Acid Concentration by Spectrophotometry

One microlitre nucleic acid sample was combined with 1 ml dH_2O in a quartz cuvette. Using dH_2O alone as reference, the absorbence at 260 nm ($O.D._{260}$) was measured. Nucleic acid mass was calculated by approximating as follows:

1 $O.D._{260}$ = 50 µg DNA

              = 40 µg RNA

              = 30-35 µg Oligonucleotide (depending on sequence and length)

2.9.8 Restriction Endonuclease Digestion of DNA

DNA digestion by restriction enzymes was performed to create the correct ends in plasmid and PCR-generated DNA enabling later ligation. It was also used to confirm the presence and orientation of a gene insert in plasmids. The restriction enzymes *Bam HI*, *Xba I* and *Hind III* were frequently used throughout the course of the project. Digestions were performed with phenol:chloroform-extracted, ethanol precipitated DNA. Approximately 1 µg DNA was combined with 10 units of restriction enzyme and the appropriate restriction buffer. To reduce the possibility of vector DNA self-ligation, calf intestinal phosphotase (CIP) was occasionally added to the reaction mix. The reaction was incubated for three hours or overnight at 37°C. The reaction mixture was ethanol precipitated (Section 2.9.6) and the resulting pellet resuspended in 10 µl dH_2O.

2.9.9 DNA Ligation

The RSV nucleocapsid gene was introduced into plasmid vectors pGEM3Z (Promega) and pET16b (Novagen). Digested PCR product and plasmid were combined in a molar ratio of 3:1, with a total DNA mass of approximately 100 ng. The DNA, ligation buffer and 4 units T4 DNA ligase were combined in a reaction volume of 15 µl and incubated

overnight at 16°C. Again, the ligation mixture was ethanol precipitated and the resulting pellet resuspended in 10 µl dH$_2$O.

2.9.10 Sequencing Analysis

The majority of DNA sequencing was performed by the Institute's Sequencing Unit using an automated system. Upon occasion, the Thermo Sequenase Cycle Sequencing Kit (Amersham) was used to manually determine cloned DNA sequence. T7 promoter and terminator primers and internal primers specific to the RSV nucleocapsid gene were used. Primer details are listed in the Appendix.

<u>Primer-Labelling Reaction</u>

0.5 pmole Primer

~0.1 µg Template DNA

5µCi [α-^{32}P]-dNTP 1

dNTP 2

dNTP 3

Reaction Buffer

Thermo sequenase

The primer was first extended in the presence of a ^{32}phosphorous (^{32}P)-labelled deoxynucleotide (dNTP). Primer and template DNA were combined with two unlabelled and one radiolabelled nucleotides. Upon addition of enzyme, the reaction was cycled between 95°C for 15 seconds and 60°C for 30 seconds 50 times to achieve primer-extension to a length limited by absence of the fourth nucleotide (approximately a further 10 nucleotides). The exact reagent conditions varied between vectors.

Four reaction mixtures were provided with the kit, each containing one di-deoxynucleotide species and three deoxyribonucleotides. Four microlitres of each was placed in individual 0.5 ml tubes. 3.5 µl radiolabelled reaction sample was added to each tube and the tube contents were mixed by pipetting. The reactions were cycled between 95°C for 30 seconds and 72°C for 90 seconds 50 times. Four microlitres stop solution was added to each reaction. The samples were heated to 70°C for five minutes, loaded and separated on a large 5% acrylamide TBE sequencing gel for approximately 1.5 hours at 80 Watts. The gel was fixed in 10% (v/v) methanol, 10%

(v/v) acetic acid for 30 minutes at room temperature, dried for one hour at 80°C and exposed to a phosphorimager screen. The image was analysed on a BioRad phosphorimager using Quantity One software.

2.9.11 Site-Directed Mutagenesis

The GeneEditor Mutagenesis Kit (Promega) was used to generate deletion mutations within N gene in pET16b. Mutation primers were designed to create terminal deletions of 100 and 200 amino acids in the resulting protein. In addition, primers designed to generate internal deletions of 40 amino acids through-out the length of the protein were synthesised. Mutation primer details are listed in the Appendix.

2.9.11.1 Preparation of Mutagenic Primers

The primers employed in this mutagenesis system become incorporated into the mutant plasmid strand. Therefore, they require phosphorylation to enable subsequent ligation of the newly-synthesised mutant strand to their 5' terminus. 12.5 nmol each mutagenic primer was combined with polynucleotide kinase (PNK), PNK buffer and 25 nmol ATP in a volume of 25 µl. The phosphorylation reactions were incubated for 30 minutes to one hour at 37°C. The enzyme was denatured by incubation for 50 minutes at 65°C. Primers were combined with ethanol and 7.5 M ammonium acetate and precipitated. The resulting DNA pellets were dissolved in dH$_2$O to create primer solutions of 1.25 µM.

2.9.11.2 Preparation of Template DNA

0.5-5 pmol DNA was denatured in 0.2 M sodium hydroxide, 1 mM EDTA in a volume of 20 µl for five minutes at room temperature. DNA was then ethanol precipitated from solution. The resulting DNA pellet was dissolved in 100 µl TE buffer pH 8. Denatured DNA recovery was confirmed by agarose gel electrophoresis.

2.9.11.3 Site-Directed Mutagenesis

Two GeneEditor primers, specific to the top and the bottom DNA strands of the same sequence, are provided with this system. Both primers introduce a mutation into the ampicillin-resistance gene carried in the plasmid. This alters the translation product enabling host resistance to both ampicillin and an additional, related antibiotic. Within this system, the GeneEditor primer is incorporated into the mutant strand. Therefore,

the choice between GeneEditor primers is determined by the orientation of the mutagenic primer. It is required that both primers adhere to the same strand to generate DNA with mutations in both N and the ampicillin gene. Approximately 0.1 pmol denatured template DNA was combined with 1.25 pmol mutation primer, 0.25 pmol GeneEditor primer and annealing buffer in a volume of 25 µl. Primer and template DNA were annealed by incubation for five minutes at 75°C followed by slowly cooling to 37°C. 5-10 units T4 DNA polymerase, 1-3 units T4 DNA ligase and synthesis buffer were added to the hybridisation reaction, increasing the volume to 35 µl. The reaction was incubated for 1.5 hours at 37°C. DNA was ethanol precipitated and the resulting pellet, resuspended in 10 µl dH$_2$0.

2.9.11.4 Bacterial Transformation and Selection

40 µl electro-competent BMH 71-18 were transformed with 5 µl mutated DNA sample by electroporation. Transformed bacteria were incubated in 800 µl 2YT broth for one hour at 37°C, 200 rpm. Four millilitres 2YT broth and 100 µl GeneEditor selection antibiotic were added and the cultures were incubated overnight at 37°C, 200 rpm. Plasmid DNA was extracted from the cultures using QIAprep Spin Miniprep Kit. DNA recovery was confirmed by agarose gel electrophoresis. 40 µl competent JM109 were transformed with 5 µl Miniprep DNA by electroporation. Transformed bacteria were combined with 800 µl SOC and incubated for one hour at 37°C, 200 rpm. 100 µl culture was spread on an antibiotic selection mix L-broth agar plate and incubated for 16 to 20 hours at 37°C. Individual colonies were picked, placed in 10 ml 2YT broth with antibiotic selection mix and incubated overnight at 37°C, 200 rpm. Plasmid DNA was again extracted using QIAprep Spin Miniprep Kit. Restriction endonuclease digestions were performed to determine resulting N gene fragment size. DNA sequence was determined manually using Thermo Sequenase Cycle Sequencing Kit (Amersham). *E.coli* strains BL21(DE3)pLysS and NovaBlue were transformed with mutant pET16b-N constructs by electroporation and mutant His-tag protein expression induced (Section 2.7.8).

2.9.12 PCR Mutagenesis

A series of primers were designed for use in PCR to create carboxy-terminal deletions within the N gene in pET16b. Primers introduced a premature stop codon and a *Bam* HI restriction site. Expand High Fidelity PCR System (Roche) was used to amplify N

gene mutants. Reagents previously listed were used with 50 pmol each primer and 12.5 ng template DNA in each reaction. Samples were thermo-cycled using the programme previously described (Section 2.9.2). DNA amplification was confirmed by agarose gel electrophoresis. The resulting PCR products were purified by Gene Clean. The products and fresh pET16b vector were then digested with the restriction enzyme *Bam* HI, the latter DNA in the presence of CIP to prevent re-ligation. Digested DNA was precipitated in ethanol and salt. Digested PCR product and pET16b were combined and ligated. Resulting plasmid DNA was used to transform *E.coli* strain DH5 by electroporation. Transformed bacteria were cultured on selective L-broth agar for 16 to 20 hours at 37°C. 10 ml aliquots of selective 2YT medium were inoculated with individual colonies and incubated overnight at 37°C, 200 rpm. Plasmid DNA was extracted by QIAprep Spin Miniprep. Resulting DNA was individually digested with restriction endonucleases *Bam* HI, *Xba* I and *Hind* III and the products visualised on an agarose gel to confirm insert size and orientation within the plasmid. BL21 (DE3) pLysS and NovaBlue *E.coli* strains were transformed with plasmids containing insert of the desired size and orientation by calcium chloride-heat shock. His-tag protein expression was induced as previously described.

2.10 The Purification of Histidine-tag Nucleocapsid Protein

2.10.1 Metal Chelate Chromatography

Histidine (His)-tag N protein was purified by HiTrap affinity chromatography (Pharmacia Biotech). Nickel ions, immobilised on an agarose matrix, are able to bind histidine residues under initial buffer conditions. Captured His-tag protein may be washed and subsequently eluted using phosphate buffer with increasing imidazole concentration. Protein purification was initially achieved manually using a syringe to apply sample and pass buffer through the column in steps of increasing imidazole molarity. This was later replaced by use of the automated Äkta Purifier Chromatography System and generation of an elution buffer gradient. His-N protein was expressed in BL21 (DE3) pLysS (Section 2.7.8). The resulting bacterial pellet, resuspended in 5 ml 10 mM imidazole start buffer, was lysed by repeated freeze/thaw in a dry ice-methanol bath and 37°C water bath. Two units per millilitre DNase (Sigma) and 160 μg ml^{-1} RNase (Sigma) were added and the lysate, incubated at 37°C for a further 30 minutes, until viscosity had reduced. The lysate was ultra-centrifuged for 20 minutes at 70 000 xg at 4°C. The resulting supernatant was passed through a 0.45 μm filter to remove any

remaining solid debris. A 1 ml HiTrap chelating column, charged with nickel ions, was connected to an Äkta Purifier Chromatography System, controlled by Unicorn software. The column was equilibrated with 10 ml 10 mM imidazole phosphate buffer at a flow rate of 2 ml min^{-1}. The clarified lysate was then applied to the column. The column was washed with 10 ml 10 mM imidazole phosphate buffer. A gradient of 10 mM to 500 mM imidazole phosphate buffer was applied to the column over a 25 ml volume. Elutant was collected in 1 ml fractions. Throughout elution, protein and nucleic acid distribution was determined by optical density (wavelengths of 280 nm and 260 nm, respectively). The presence and purity of His-tag protein in the different fractions was assessed by SDS-PAGE and western blot analysis (Sections 2.10.7 and 2.10.11). Purified His-N protein was stored at 4°C.

2.10.2 Iso-electric Focusing (IEF) Gels

The iso-electric point (pI) of His-N protein was determined using the Phast System and PhastGel IEF media (Pharmacia Biotech). Approximately 100 ng His-N protein was placed in a gel applicator at each end of a PhastGel IEF3-9 (broad pH range). Four microlitres pI calibration marker (Pharmacia Biotech) was placed in an applicator at only one gel end. The pre-formed gel was placed in the Phast system and the electrodes put in position. Electrophoresis was performed at 3.5 W, 15°C and the voltage was varied. Initially, the voltage was maintained at 2000 V for 75 Vh. It was then reduced to 200 V for 15 Vh during sample application before returning to 2000 V for 410 Vh. The gel was removed from the apparatus and proteins visualised by silver or coomassie blue stain (Sections 2.10.9 and 2.10.10). The pI of His-N protein was estimated by comparison with marker proteins of known pI.

2.10.3 Ion-Exchange Liquid Chromatography

Ion-exchange liquid chromatography was employed to improve His-N protein purity. Initial experiments were performed with sepharose-bound ions SP and DEAE to establish which was better able to bind His-N. These suggested the weak cation DEAE as the preferred choice for His-tag N protein purification by ion-exchange liquid chromatography (Chapter 7). A 1 ml DEAE-sepharose column was fitted to the Äkta Purifier Chromatography System. The column was equilibrated with 5 ml 0.2M Tris (pH 6.5) at a flow rate of 1 ml min^{-1}. His-N protein in 0.2 M Tris (pH 6.5), purified by HiTrap affinity chromatography, was applied to the column. The column was washed with 10

ml 0.2 M Tris (pH 6.5). A gradient from 0 M to 1 M sodium chloride in 0.2 M Tris (pH 6.5) was applied to the column over a volume of 15 ml. Fraction volumes of 1 ml were collected and investigated for presence and purity of His-N protein by SDS-PAGE and western blot (Sections 2.10.7 and 2.10.11).

2.10.4 Size-Exclusion Liquid Chromatography

To further purify His-tag N protein, size-exclusion liquid chromatography was employed. A Superdex 6 column (Pharmacia Biotech) was used for this purpose. Superdex is a matrix of beads with an internal pore of pre-determined size. Upon sample application, large proteins pass around the beads and are swiftly washed through the column, appearing in early elution fractions. Smaller proteins pass through the bead pores. This hinders their progress through the column resulting in their appearance in later fractions. A 30 ml Superdex 6 column was attached to the Äkta Purifier Chromatography System. The column was equilibrated with 30 ml binding buffer (Appendix). 0.5 ml DEAE-purified His-N protein was applied to the column. 30 ml binding buffer was passed through the column and the elutant retained in 1 ml fractions. Fractions were investigated by SDS-PAGE and western blot (Sections 2.10.7 and 2.10.11).

2.10.5 Protein Dialysis

2.10.5.1 Slide-A-Lyzer Dialysis

Protein samples were dialysed using Slide-A-Lyzer dialysis cassettes (Pierce) with a 10 000 molecular weight cut-off (mwco). The cassette was pre-wetted with the buffer into which protein was to be transferred. Using a needle and syringe, 0.5-1.5 ml protein sample was introduced into the cassette at one corner. The cassette was floated in 500 ml buffer for one hour at room temperature, with stirring. The buffer was then refreshed and the sample incubated overnight at 4°C. The membrane was gently rubbed with gloved fingers to dislodge any adhered protein. The sample was then removed from the cassette using needle and syringe.

2.10.5.2 Centricon 10 Buffer Exchange

Centricon 10 (Millipore), containing a 10 000 mwco membrane, offered an alternative method of buffer transfer. 1 ml protein sample was placed in a Centricon 10 tube and centrifuged for one hour at 3000 xg, 4°C. Under centrifugal force, the buffer passes

through the membrane and structures with a molecular weight greater than 10 000 are retained on the membrane. 0.5 ml buffer was used to wash the membrane by pipetting. The tube was centrifuged for a further 30 minutes. The retained protein sample was suspended in the desired volume of buffer and removed from the tube. In addition to buffer transfer, this technique enabled protein samples to be concentrated.

2.10.6 Protein Quantification

Protein concentration was determined by the Lowery method using the DC Protein Assay Kit (BioRad). Known concentrations of bovine serum albumin (BSA) were used as reference. In individual wells of a 96-well plate, 25 µl Reagent A was combined with 5 µl protein sample. 200 µl Reagent B was added to the well and the plate incubated for up to 30 minutes. Sample absorbency at O.D. $_{705}$ was measured and protein concentration calculated using a plate reader.

2.10.7 Tris-Glycine SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

BioRad Proteon II minigel apparatus was employed to separate protein samples through polyacrylamide gel (Laemmli, 1970). The glass plates were wiped with dH_2O, followed by ethanol and placed in gel pouring apparatus. The resolving gel mix of desired acrylamide percentage was assembled and immediately poured between the glass plates (Appendix). 200 µl iso-propanol was placed on top and the gel left to polymerise for 30 minutes. The iso-propanol was discarded and the gel surface washed with dH_2O. Excess dH_2O was removed by draining and blotting. The stacking gel mix was then combined and poured on to the resolving gel. A comb was inserted and the gel left to polymerise for 20 minutes. The gel was transferred to the electrophoresis apparatus and submerged in tank buffer. The comb was removed and unpolymerised acrylamide washed from the wells by flushing with tank buffer using a needle and syringe. The protein sample was combined with sample loading buffer, boiled for five minutes and loaded onto the gel. Rainbow marker (Amersham) was also loaded onto the gel and used as a guide to protein size. Electrophoresis was performed at 150 V until the bromophenol blue dye-front reached the gel bottom (approximately one hour). The gel was removed from the electrophoresis apparatus and the protein visualised by staining with coomassie blue or silver-stain. Alternatively, the protein was transferred from gel to nitrocellulose for western blot analysis.

2.10.8 Tris-Tricine SDS-Polyacrylamide Gel Electrophoresis

Tris-tricine SDS-PAGE was performed when resolution of small proteins (< 10 kDa) was required. The separating and stacking gel solutions (Appendix) were assembled and a minigel poured following the above procedure. The gel was placed in the electrophoresis apparatus and each running buffer, added to the correct electrode. Protein was combined with tricine sample buffer and heated for five minutes at 95°C. Samples were loaded and the gel run for approximately four hours at 30 mA. The gel was then removed from the apparatus and either silver-stained or dried and exposed to a phosphorimager screen.

2.10.9 Coomassie Blue Staining

A minigel was incubated in 50–100 ml coomassie blue stain for approximately one hour. It was then transferred to de-stain solution and incubated until background stain had reduced. The resulting gel was viewed over a light box and photographed.

2.10.10 Silver Staining

The Protein Silver Staining Kit and protocol from Pharmacia Biotech was used. The protein gel was first incubated for 30 minutes in fixative buffer. It was incubated for a further 30 minutes in sensitising buffer. The gel was then washed three times with distilled water, five minutes per wash. The gel was transferred to silver solution and incubated for 20 minutes. It was twice washed with dH_2O for one minute, placed in developing buffer and incubated for five to 30 minutes. Upon adequate staining of the protein, stop buffer was added to the gel. The gel was visualised over a light box and photographed.

2.10.11 Western Blot Analysis

Protein samples were separated by SDS-PAGE. Using western blot apparatus (BioRad), protein was transferred to nitrocellulose membrane. The gel was placed in contact with enhanced chemiluminescence (ECL) nitrocellulose membrane (Amersham) in a cassette, sandwiched between pieces of blotting paper and sponge, pre-wetted in transfer buffer. The cassette was placed in the blotting apparatus with the gel adjacent to the cathode. An ice block was placed in the tank and the cassette, submerged in transfer buffer. A current of 250 mA was applied for one hour. The

apparatus was dismantled and the membrane placed in blocking buffer (TPBS with 5% (w/v) Marvel Powdered Milk) and incubated for one hour at room temperature. The membrane was washed three times with TPBS, 10 minutes per wash. The desired dilution of primary antibody was made in blocking buffer and added to the membrane. The membrane was incubated for one hour at room temperature. The membrane was washed as previously described. Secondary antibody conjugated to horseradish peroxidase (HRP) was diluted in block buffer and applied to the membrane, which was incubated for a further one hour at room temperature. The membrane was again washed and ECL performed. 5 ml each of ECL solutions 1 and 2 were combined and immediately added to the membrane. The membrane was incubated for one minute with shaking. Excess solution was removed by blotting and the membrane, placed in a developing cassette and exposed to film for two to 60 seconds. The film was then developed in a Kodak Xomat machine and antibody-bound protein visualised.

2.11 Protein-RNA Interaction Assays

2.11.1 Radiolabelled RNA Probe Synthesis

The T7-MEGAShortscript™ Kit (Ambion) was used to generate radiolabelled RNA oligonucleotides. Each template DNA contained a T7 promoter sequence and resulting transcripts were between 30 and 50 nucleotides. The template of the specific probe was an oligonucleotide and that of the non-specific probe was a linearised plasmid. Template sequences are listed in the Appendix.

10 x	Reaction Buffer
7.5 mM	CTP, GTP and UTP solutions
3.75 mM	ATP solution
50 μCi	[α-^{32}P]ATP
~1 μg	Template DNA
<u>4 μl</u>	T7 MEGAShortscript™ Enzyme Mix
40 μl	

Reagents were combined and the transcription reaction incubated for five hours at 37°C. Template DNA was removed by digestion with 4 units RNase-free DNase for 30 minutes at 37°C. The reaction was terminated upon addition of formamide gel loading

buffer. The sample was heated for one minute at 95°C, placed on ice for two minutes and loaded onto a 10% acrylamide TBE minigel. Electrophoresis was performed for 45 minutes at 150 V. The gel was removed from the apparatus, placed in a cassette and exposed to film for five minutes. The film was developed and aligned with the gel to identify the position of the radiolabelled RNA. The band was sliced from the gel using a scalpel blade and placed in 1 ml TE buffer (pH 7.4) with 0.5% (w/v) SDS. The RNA was eluted for 16 to 20 hours at 37°C. The eluted RNA was extracted by phenol (pH <6) and by chloroform:IAA extraction. The RNA was precipitated upon addition of 3 M sodium acetate and ethanol. The RNA pellet was dissolved in 60 μl dH_2O. One microlitre was combined with scintillation fluid and the radioactivity measured in a scintillation counter.

2.11.2 Detection of Radiolabelled Products

A phosphorimager (BioRad) was frequently used to detect radiolabelled products. Radiolabelled experimental samples were separated by polyacrylamide gel electrophoresis, the gel was dried under vacuum for one hour at 80°C and exposed to a phosphorimager screen overnight. The resulting image was visualised on the multi-imager. Densitometric analysis was also performed on the multi-imager using BioRad Quantity One software. Densitometry was performed on radiolabelled bands detected by phosphorimagery and on protein bands visualised by western blot and ECL.

2.11.3 Gel Mobility Shift Assay

10 000 counts per minute (cpm) radiolabelled RNA probe was incubated for one minute at 95°C, then transferred to ice for two minutes. It was combined with 0.01-10 μg His-N protein and binding buffer in a final volume of 7.5 μl. The reaction was incubated for 15 minutes at 33°C. One microlitre sterile glycerol was added and the sample loaded onto a 5% acrylamide TBE gel (Hoeffer). Gel loading buffer was placed in a free well and used as a migration marker. Electrophoresis was performed for approximately 2.5 hours at 200 V. The gel was removed from the apparatus, dried at 80°C under vacuum and exposed to a phosphorimager screen overnight. The resulting image was detected by phosphorimaging.

2.11.4 North-Western Blot Analysis

Protein samples were separated by 10% acrylamide SDS-PAGE and blotted onto ECL nitrocellulose membrane. The membrane was washed twice with PBS, placed in PBS containing 5% Marvel and 1 mM DTT and incubated for 16 to 20 hours at room temperature with shaking. The membrane was washed three times with HBB buffer and once with HYB100, 15 minutes per wash. The membrane was then transferred to HYB100 buffer with radiolabelled RNA probe (50 000 cpm ml^{-1}). The membrane was incubated for two hours at room temperature with shaking and was then transferred to 4°C overnight. The membrane was washed three times with HYB100 buffer, 15 minutes per wash, to remove unbound probe. It was then exposed to a phosphorimager screen overnight.

2.11.5 North-Western Blot Analysis: Competition Assay

This assay was devised to compare the affinity of His-N protein for viral and non-viral RNA sequences. It was performed to investigate the possibility of sequence preference by N protein.

Eight wells on a minigel comb were taped together to create a single large well in a 10% acrylamide SDS-PAGE minigel. 16 μg His-N protein was combined with sample loading buffer, incubated for five minutes at 95°C and loaded into the large well. Five microlitres Rainbow Marker was loaded into an adjacent single well. Electrophoresis was performed for one hour at 150 V. Protein was transferred from gel to ECL nitrocellulose membrane as above. The membrane was incubated in PBS with 5% Marvel and 1 mM DTT for 16 to 20 hours at room temperature with shaking.

The region of membrane containing His-N protein was vertically sliced into eight pieces of equal size. By alignment with Rainbow Marker, the approximate position of protein was isolated in each slice. The membrane slices were washed three times with HBB and once with HYB100, 15 minutes per wash. Pieces were then placed in individual 1.5 ml microtubes.

To investigate N protein sequence preference, radiolabelled target RNA was competed out with increasing amounts of unlabelled RNA. Using spectrophotometry (O.D.$_{260}$), the mass and hence, molarity of radiolabelled specific and unlabelled non-specific (i) RNA probes were calculated. RNA probes were combined in increasing molar ratios in

aliquots of HYB100 buffer in a final volume of 500 µl. Radiolabelled probe was maintained at the same concentration (~15 pmol per strip). The ratio of unlabelled probe was extended from 0 to 729 molar excess in a three fold molar ratio dilution series. The membrane was incubated at room temperature for two hours and transferred to 4°C overnight. The membrane pieces were washed three times with HYB100, 15 minutes per wash. They were then exposed to a phosphorimager screen overnight. Densitometric analysis was performed on the resulting bands using Quantity One software on the multi-imager.

The assay was repeated with radiolabelled non-specific (ii) probe, using the same RNA molarity as that used with specific probe. Due to differences in specific activity between the probes, the actual amount of radioactivity (cpm) applied differed slightly. Results from the different RNA probes were normalised to enable their comparison.

2.11.6 Filter Assays

2.11.6.1 Dot Blot Assay

100 to 500 ng His-N protein was combined with 5000 cpm radiolabelled probe and binding buffer in 10 µl. The reaction was incubated for 20 minutes at 33°C. ECL nitrocellulose membrane was placed in dot blot apparatus and a vacuum applied. Each dot was pre-wetted with 200 µl binding buffer. The samples were applied to the membrane then washed with 200 µl 5% (w/v) trichloro-acetic acid (TCA). The vacuum was maintained for a further one minute to dry the membrane. The membrane was investigated by phosphorimaging. Western blot analysis was performed to confirm protein was retained on the membrane.

2.11.6.2 Individual Filter Binding Assay

The dot blot assay procedure was repeated with the exception of the use of glass fibre (Sigma and Millipore) filters in place of ECL nitrocellulose membrane and the dot blot apparatus. The filters were wetted with binding buffer, sample passed through and washed with 5% (w/v) TCA under vacuum. The filters were placed in scintillation fluid and their radioactivity measured in a scintillation counter.

2.11.7 <u>UV Cross-linking Assay</u>

The UV cross-linking assay was performed with purified His-N protein and with radiolabelled specific RNA probe. 0.5–2 μg recombinant protein was combined with 10 000 – 15 000 cpm (~5 ng) RNA and binding buffer in individual wells of a covered 96-well plate. Samples were incubated for 20 minutes at 33°C. The plate was placed on ice, transferred to a Stratalinker UV Cross-Linker 1800 and positioned at a distance of 5 cm from the light source. The lid was removed and the plate exposed to UV light (2500 μjoulescm^{-2} x100). 0.1–1 μg RNase A was added to each sample and the plate incubated for 15 minutes at 33°C. The samples were then transferred to individual 1.5 ml microtubes. 0.1 μg sequencing grade trypsin (Promega) was added to each sample and the digestion incubated for three hours to overnight at 37°C. Additional proteases had been investigated for their ability to digest N protein. Trypsin was the enzyme of choice due to the number and size of digest fragments created and enzyme efficiency (see Chapter 7). The sample was combined with tricine sample buffer, heated to 95°C for five minutes and separated in a 16% acrylamide Tris-tricine gel. The gel was dried for one hour at 80°C, exposed to a phosphorimager screen overnight and the image analysed in a phosphorimager. Upon optimisation of the protocol and visualisation of the digested radiolabelled-protein fragments, fresh samples were prepared for analysis by mass spectrometry. This involved the substitution of labelled- with unlabelled-specific RNA probe. The above protocol was repeated. The samples were then transferred to Slide-A-Lyzer mini-dialysis units (3000 mwco). The units were placed in one litre 0.1% trifluoro-acetic acid (TFA) and incubated for two hours at 4°C with stirring. The buffer was refreshed and dialysis repeated for a further two hours. The samples were transferred to 1.5 ml microtubes, snap-frozen in a dry ice-methanol bath and freeze-dried overnight at -60°C under vacuum. The samples were analysed by mass spectrometry. Analysis was kindly performed by P. Ashton at the Department of Microbiology, University of York.

2.12 Structural Studies

2.12.1 <u>Preparation of Ribonucleocapsid-like Structures</u>

2.12.1.1 Preparation from Transformed Bacteria

The following procedure was used to generate and isolate ribonucleoprotein (RNP)-like structures from bacteria transformed with pET16b-N and mutant derivatives.

50 ml cultures of BL21 (DE3) pLysS *E.coli*, transformed with pET16b-N gene or mutant derivatives, were incubated at $37^{\circ}C$ until an O.D.$_{600}$ 0.8 was obtained. Recombinant protein expression was induced by addition of 1 mM IPTG. The cultures were incubated for a further three hours at $37^{\circ}C$. The cultures were exposed to a centrifugal force of 1700 xg for 15 minutes at $4^{\circ}C$. The supernatant was discarded and each pellet, resuspended in 5 ml TNE buffer with 0.6% (v/v) Nonidet-P40 (N-P40). Bacteria were lysed by three rounds of freeze/ thaw in a dry ice-methanol bath and $37^{\circ}C$ water bath. DNase (units ml^{-1}) and RNase A (200 µg ml^{-1}) were added and the lysate incubated at $37^{\circ}C$ for 30 minutes until viscosity had reduced. The lysate was subjected to a centrifugal force of 3000 xg for 10 minutes at $4^{\circ}C$. The supernatant was removed and placed at $4^{\circ}C$.

2.12.1.2 Preparation from Sf21 cells infected with Bac-N$_{wt}$

RNP-like structures were extracted from Sf21 cells infected with Bac-N$_{wt}$ (Section 2.6.7). The cells were harvested and centrifuged at 3000 xg. The cell pellet was washed twice in NTE buffer and resuspended in NTE buffer with 0.6% N-P40 to lyse the cells. The lysate was treated in the same manner as the bacterial samples.

2.12.2 Caesium Chloride Buoyant Density Gradients

2.12.2.1 Buoyant Density Step and Equilibrium Gradients

A step gradient was created in a 14 ml ultra-centrifuge tube with 2 ml each of 25%, 30% and 40% Caesium Chloride (CsCl) in TNE buffer (Appendix). One millilitre 5% sucrose in TNE buffer was placed at the gradient top by pipette. 1.5-5 ml clarified lysate was laid at the gradient top. The gradient was subjected to a centrifugal force of 250 000 xg for two hours, $15^{\circ}C$ in a TST41 swing bucket rotor.

The resulting gradient was viewed under a direct light source. Using a needle and syringe, the tube was pierced and the visible band in the 30% CsCl ($\sim$1.3 g / cm^{3}) region was removed from the gradient. The volume of the extract was increased to 5 ml with 30% CsCl and the final concentration confirmed by refractometry. The gradient was subjected to a centrifugal force of 155 000 xg for 36 hours, $15^{\circ}C$ in an AH650 Sorvall swing bucket rotor to achieve buoyant equilibrium. The gradient was manually divided into 200 µl fractions from the gradient top to bottom. Individual fractions were analysed by western blot and by electron microscopy (EM).

2.12.2.2 Preformed Buoyant Density Gradient

This gradient type provided a quicker method to isolate RNP-like structures. However, resulting RNP density data may be taken only as an indication as equilibration is not fully achieved in the preformed gradient.

A 20-40% CsCl preformed gradient was created. Steps of 6 ml each of 20% and of 40% CsCl in TNE buffer were placed in an ultra-centrifuge tube suitable for TST41. The tube was placed in a Biocomp Gradient Master and a continuous gradient created (65°; 30 rpm; 2.5 minutes). 0.5 ml 5% sucrose was added at the gradient top. 1.5-4 ml bacterial lysate was laid at the top of the gradient. The gradient was subjected to a centrifugal force of 250 000 xg for 16 hours at 15°C in a TST41 swing bucket rotor.

When viewed under a direct light source, a band was visible at the gradient middle. The gradient was manually divided into 200 μl fractions from the gradient top to the bottom. Fractions were investigated by western blot and EM analysis.

2.12.3 EM Grids

Five microlitres protein sample was placed on a carbon-coated copper grid and incubated for one minute at room temperature. The grid was washed three times with water, one minute per wash. The sample was then stained with 1% (w/v) phosphotungstic acid (PTA) for one minute. Excess stain was gently removed by blotting the grid edge against Whatman #5 filter paper. The grid was allowed to air-dry. The mounted sample was observed under the electron microscope (100 V Phillips) at a magnification of 32000 x. Photographs were kindly developed and images digitally scanned by Dr. D. Bhella.

2.12.4 RNA Extraction from Mutant RNP-like Structures

The RNP-like structures isolated by buoyant density gradient were investigated to establish whether or not they were associated with RNA.

2.12.4.1 Immuno-Precipitation of RNP-Like Structures

Gradient fractions with a density of ~1.3 g / cm^3 and containing N or mutant N protein were analysed by immuno-precipitation. 150 μl of each fraction was combined with 3 μl

anti-His tag antibody (Sigma) and incubated for one hour at 4°C with shaking. 100 µl protein A-sepharose (50% (w/v) in Machamer buffer), 100 µl Machamer buffer and 150 µl dH$_2$O were added and the sample incubated for two hour at 4°C with shaking. The sample was centrifuged for one minute at 6000 xg. The supernatant was removed and the sepharose-protein A pellet was washed with 500 µl lithium chloride buffer. The sample was again centrifuged and the wash twice repeated. The pellet was resuspended in 120 µl dH$_2$O. Twenty microlitres were removed for western blot analysis to confirm presence of His-tag protein.

2.12.4.2 RNA Extraction

RNA was extracted from 100 µl protein A sample with phenol (pH <6) and with chloroform:IAA. The RNA was precipitated by the addition of ethanol and 3 M sodium acetate pH 4.8 and incubation for one hour at -20°C. The resulting RNA pellet was resuspended in 5 µl dH$_2$O and 1 µl RNasin.

2.12.4.3 End-Labelling of RNA

RNA isolated from the RNP-like structures was radiolabelled with ^{32}P. The RNA sample was combined with polynucleotide kinase (PNK), PNK exchange buffer and 2.5 µCi [γ^{32}P]-ATP in a volume of 10 µl. The phosphorylation reaction was incubated for 30 minutes at 37°C. Formamide loading buffer (T7-MEGAShortscript kit, Ambion) was added and the sample was heated at 95°C for one minute then placed on ice for two minutes. The sample was separated in a large 15% acrylamide TBE gel (Hoeffer) for approximately 2.5 hours at 200 V. The gel was dried for 1.5 hours, exposed to a phosphorimager screen overnight and analysed on the multi-imager.

3 Growth of RSV

3.1 Introduction

Before addressing the aims of the project, a few preliminary experiments were performed to study the interactions of the N protein within RSV-infected cells. The RSV strain A2 was used in each experiment. It was routinely cultured in CV-1 cell monolayers as this cell-type produced the highest yield of infectious virus compared with other cell lines (J. Murray, personal communication). Results include the determination of virus titre by plaque assay. The expression of viral proteins was analysed by western blot and by immuno-precipitation, the latter demonstrating the occurrence of certain protein-protein interactions. The intracellular distribution of RSV proteins was observed by immunofluorescent microscopy. This allowed visualisation of structures in the cell relating to viral replication. Also included in this chapter is data, where relevant, accrued by other members of the laboratory. Some of the data presented herein was used in the publication of Murray et al. (2001).

3.2 Preparation of RSV Stock

RSV stock was routinely generated by infecting CV-1 cell monolayers, at 70 to 80% confluency, for three to four days. The majority of RSV is cell-associated, therefore, to release infectious virus, the cells were disrupted with glass beads in a volume of 10 ml growth medium. Cell debris was removed by centrifugation and viral stocks stored at -70°C.

Similarly, sub-confluent monolayers were infected at low multiplicity for three to four days prior to preparation of cell lysates for western blot and for RIPA. These conditions were identified as ideal for the preparation of virus (J. Murray, personal communication).

3.3 Titration of RSV by Plaque Assay

CV-1 cells, cultured as monolayers in six-well plates, were infected with RSV in a 10 fold dilution series. Cells were overlaid with agar-DMEM and incubated for five days prior to fixing and giemsa staining. The plate was viewed under the light microscope and the number of plaques visible in each well counted. RSV plaques were observed as small gaps (< 2 mm diameter) in the cell monolayer.

Virus Dilution	Plaque Number
Neat	TNTC
10^{-1}	TNTC
10^{-2}	TNTC
10^{-3}	TNTC
10^{-4}	250
10^{-5}	47
0	0

Table 3.1: Titration of RSV in CV-1 cells using the plaque assay.

(TNTC – Too numerous to count)

Table 3.1 lists results from a typical plaque assay. In this instance, the virus titre was calculated to be 4.7×10^6 plaque forming units per ml (pfu ml^{-1}). The yield of RSV is low when compared with other members of the *Mononegavirales*. For example, a yield in excess of 1×10^8 pfu ml^{-1} was routinely obtained with measles virus, prepared in a similar manner (P. Yeo, personal communication).

3.4 Western Blot

Whole cell lysates from RSV-infected CV-1 cells or from mock-infected cells were separated by SDS-PAGE and transferred to nitrocellulose. The membrane was divided into strips of 0.5 cm width and each strip probed with one of a number of antibodies to different RSV proteins (Figure 3.1). Antibodies employed had been raised against recombinant RSV proteins expressed and purified from *E.coli* (see Chapter 2). Each antibody detected a single band in the strips containing RSV antigen. The estimated molecular weights of the detected protein bands corresponded to their predicted size, as determined from amino acid sequence information (reviewed by Collins *et al*, 1996) with the exception of P. The N protein, present in lanes 1, 2, 4, 5 and 6, migrated as a band of 42 to 43 kDa which corresponded to the calculated mass of 43.4 kDa. The M protein (lane 8) was observed at 28 kDa, corresponding to it's calculated mass of 28.7 kDa. M2-1 (lane 9) was present at 22 kDa, again close to it's calculated mass of 22.2 kDa. The P protein (lanes 3 and 7) appeared as a band of approximately 38 kDa, considerably larger than it's calculated mass of 27.2 kDa. This phenomenon has

Figure 3.1: Western blot analysis of RSV-infected CV-1 cell lysates.

RSV-infected (a) and mock-infected (b) CV-1 cell lysates were separated by 10% polyacrylamide SDS-PAGE and transferred to nitrocellulose membrane. The membrane was blocked with 5% Marvel in TPBS, divided into strips and probed with a variety of RSV-specific antibodies. Protein bands were visualised by ECL and autoradiography. Membrane strips were probed with RSV antibodies as follows: **1** αN014, **2** αN015, **3** αP658, **4** αN3, **5** αN9, **6** αN15, **7** αP001, **8** αM, **9** α22k.

The viral proteins are expected to migrate as the following masses: N (43.4 kDa), P (38 kDa), M (28.7 kDa) and M2-1 (22.2 kDa).
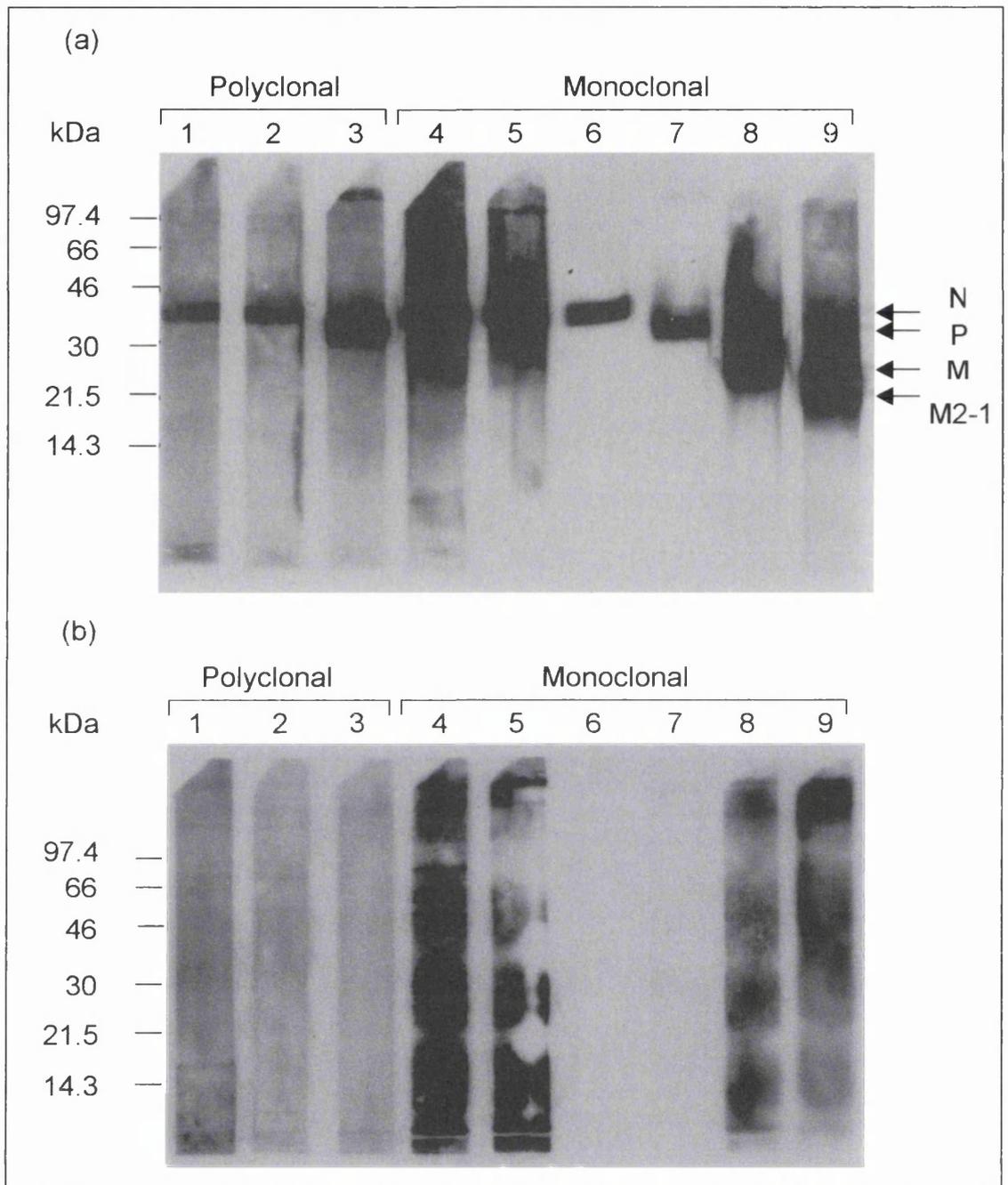
previously been reported and is caused by the acidity of the protein affecting it's interaction with SDS. The mass/charge ratio is disturbed resulting in the retarded migration of P through the gel (Caravokyri & Pringle, 1992).

Bands were absent from the majority of strips containing mock-infected CV-1 cell lysates with the exception of those probed with polyclonal antisera αN014 and αN015 and with monoclonal antibody αN3. A faint band of 50 to 55 kDa was present in each of these strips possibly generated by cross-reactivity of the primary or secondary antibody to cellular protein. These bands were absent from strips containing infected lysate possibly due to saturation of antibody binding to the target viral protein.

3.5   Radiolabelled Immuno-Precipitation Assay (RIPA)

Cell extracts in RIPA buffer, radiolabelled with [^{35}S]-methionine, were prepared and incubated with RSV antibodies for one hour at 4°C. The antibody-protein complexes were precipitated using Protein A-sepharose beads. Each reaction was washed three times with cold RIPA buffer before the addition of sample loading buffer and separation by 12% polyacrylamide SDS-PAGE. The gel was dried and exposed to a phosphorimaging screen for 10 to 12 hours at room temperature before being scanned by a phosphorimager (Figure 3.2).

Antibodies against proteins N, P, M and M2-1 were used to precipitate RSV antigen. The majority of antibodies caused the precipitation of more than one viral protein. Polyclonal antisera αN014 and αN015, and MAbs αN3 and αN9 each precipitated a protein of the expected size of N (43.4 kDa). In each lane, an additional band migrated to a position expected of the P protein (38 kDa). The N-P interaction has previously been documented for RSV and other members of *Mononegavirales,* for example, Sendai virus (Curran *et al.,* 1995; Garcia-Barreno *et al.,* 1996). In each sample, the N band was of greater intensity than that of P. This may have arisen due to a difference in protein radiolabelling – the N protein contains 14 methionines compared with 7 methionines in the P protein. Alternatively, it may represent the stoichiometry of N-P protein association. The N antibodies also precipitated a protein similar in size to that of M (28.7 kDa). A faint band with a size corresponding to that of M2-1 (22.2 kDa) was detected when N was precipitated by αN014, αN015 and αN9.

Figure 3.2: Immuno-precipitation from RSV-infected [^{35}S]-methionine-labelled CV-1 cell lysates.

The [^{35}S]-methionine-labelled lysates of RSV-infected (i) and mock-infected (m) CV-1 cells were incubated with various RSV antibodies for 1 hour at 4°C. Antibody-protein complexes were precipitated with Protein A-sepharose beads. The beads were washed, combined with sample loading buffer, boiled and separated by 12% polyacrylamide SDS-PAGE. Radiolabelled bands were visualised by phosphorimagery. The viral proteins N (43.4 kDa), P (38 kDa), M (28.7 kDa) and M2-1 (22.2 kDa) are indicated in the figure.

There are four protein bands of unknown origin precipitated from RSV-infected cell lysates. Bands are identified as follows: 1 * (> 66 kDa), 2 * (30 kDa), 3 * (25 kDa) and 4 * (< 14.3 kDa). (Marker, M).
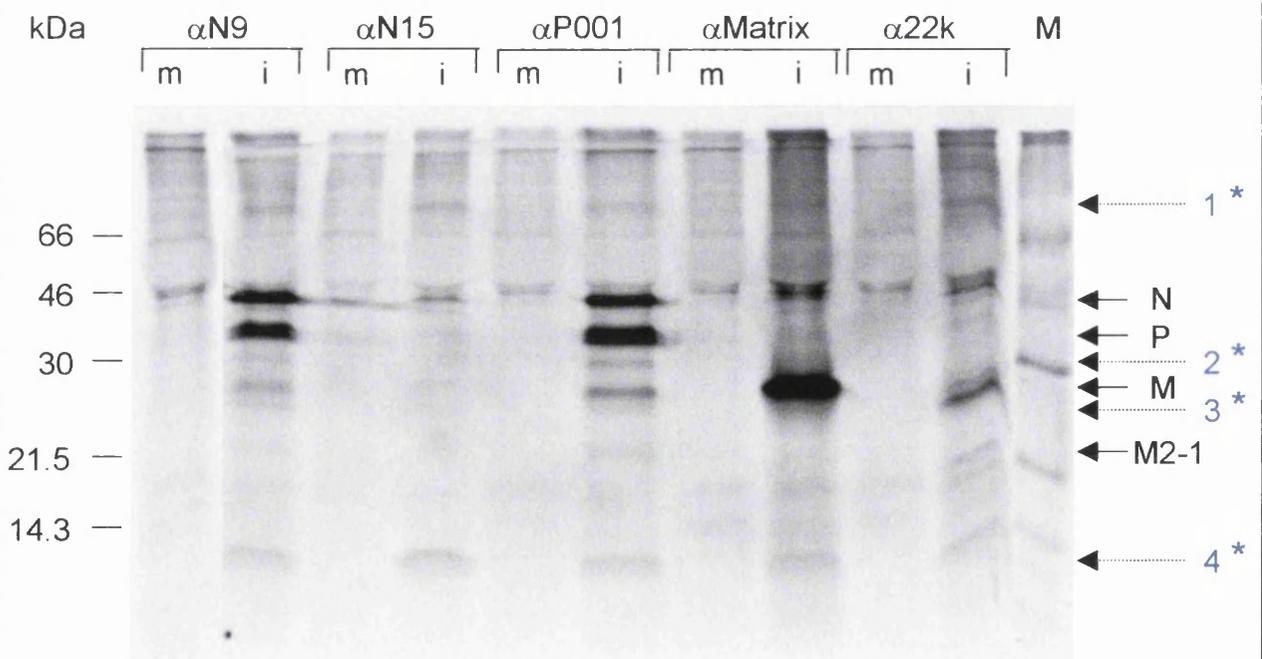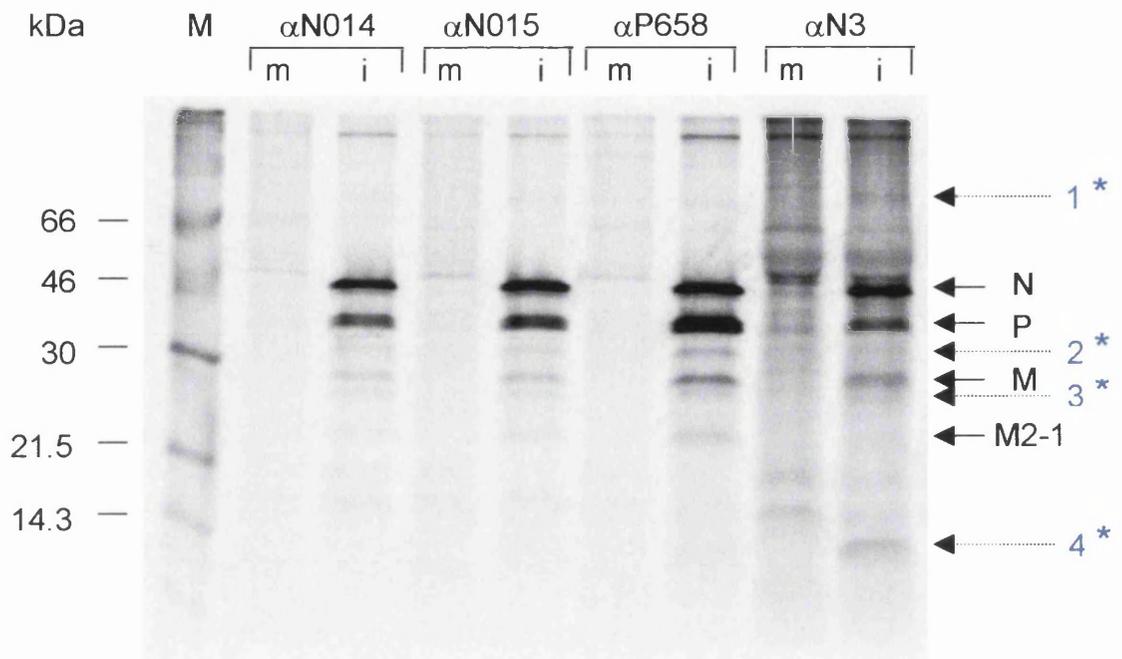
Monoclonal antibody αN15 precipitated only a small amount of N, forming a faint protein band. The P protein was notably absent from this sample. The epitope recognised by this antibody has been mapped, using a blocking peptide assay, to a site on the N protein inaccessible when complexed with P (Murray et al., 2001). Therefore, the faint band of N, precipitated by αN15, may represent either free-protein (N protein not complexed with P, prior to assembly into the nucleocapsid) or N protein, assembled with RNA into the nucleocapsid. When compared with the other αN MAbs, the band intensities suggested that the majority of N is in soluble form and is complexed with P protein. In experiments where recombinant N was expressed in the absence of P, MAb αN15 precipitated protein as efficiently as do αN3 and αN9 (P. Yeo, personal communication). Thus, the failure to precipitate N efficiently from RSV-infected cell lysates was not due solely to the affinity of the antibody.

Polyclonal antiserum αP658 and MAb αP001, directed against P, co-precipitated N with the P protein. The band intensities were the inverse of those obtained with the N antibodies, that is, P was more intense than N. The stoichiometry of the N-P interaction has been estimated from 1N: 2P to 1N: 4P in related viruses (the reverse has been estimated in VSV – 2N: 1P (Green et al., 2000)). However, the data here does not allow estimation of the numbers of P monomers interacting with N. Protein bands corresponding in size to M and very small amounts of M2-1 were also present within these lanes.

Anti-Matrix precipitated M and small amounts of N. Anti-22k antibody recovered the proteins M2-1, M and N. The M2-1 protein band is very faint, as observed upon its co-precipitation with other viral proteins. This may reflect either low M2-1 concentration within the infected cell lysate or poor radiolabelling of the protein. The co-precipitation of N, M and M2-1 by α22k and by N-specific antibodies and of N and M by αMatrix suggests a possible relationship exists between these proteins. The N-M interaction has been postulated for related viruses and would be important in the assembly of the mature virion prior to and during budding (Coronel et al., 1999).

A number of bands occurred in precipitates from infected cell lysates that were not readily identified as viral proteins. A 30 kDa band (2 *) was present in several lanes derived from infected cell lysate. A band of 25 kDa (3 *) was detected in lanes of infected lysate precipitated by N and P antibodies. A band greater than 66 kDa (1 *)

and another less than 14.3 kDa (4 *) occurred in infected samples precipitated with monoclonal antibodies. Each band, absent from mock samples, may represent protein of either viral or cellular origin. If viral, the species with a mass of less than 14.3 kDa corresponds approximately with the size of either NS1 (15.6 kDa), NS2 (14.7 kDa) or partially processed SH (13 to 15 kDa). However, the identity and origin of this protein is unknown. The RIPA illustrates interactions that occur between viral and, potentially, cellular proteins. The unambiguous association of N with P has been demonstrated and possible associations between N, M and M2-1 proteins suggested.

3.6   Immunofluorescence

CV-1 cells, seeded on 13 mm-diameter glass cover-slips, were infected with RSV at an m.o.i. of 0.5 and incubated for 16 hours. Cells were fixed with ice-cold methanol and probed with primary antibodies directed against viral proteins. Antibodies were used either individually or in combination. Viral proteins were detected with the appropriate anti-species secondary antibody, containing a fluorescent conjugate, and viewed using a Nikon fluorescent microscope (Figure 3.3 a to e). For co-localisation experiments, images were obtained on a Zeiss LSM510 laser confocal microscope (Figure 3.3 f to i; courtesy of J. Murray). Several distinct features were apparent, most notably that N, P and M2-1 (Figure 3.3 a, b, and c respectively) share a common staining pattern. All showed diffuse cytoplasmic staining with concentration in what appeared to be large aggregates of protein (arrowed). These structures, referred to as cytoplasmic inclusion bodies (CIs), appear to contain many of the components necessary for transcription and replication of the genome. Confocal images showed the co-localisation of N and P (panel f), N and M2-1 (panel g) and P and M2-1 (panel h) within the CI bodies. The presence of L within CIs has yet to be confirmed.

A second region of viral protein accumulation was in spike-like structures around the periphery of the cell, presumably representing regions of virion assembly. The spike structures are most clearly seen in Figure 3.3 e (polyclonal αN015), the sizes of which range from 5 to 10 nm. The structures are analogous to the filamentous form of human Parainfluenza virus type 2 (hPIV-2) described by Yao & Compans (2000). The spikes are actin-dependent structures and are composed of detergent insoluble membranes or lipid rafts (J. Murray and G. Henderson, personal communication).

Figure 3.3: IF images of primary RSV infection in CV-1 cells.

CV-1 cells, grown as a monolayer on cover-slips, were infected with RSV and incubated overnight at 33°C. Cells were fixed and probed with antibody specific to viral antigens. Cells were then probed with secondary antibody containing a fluorescent label and viewed by fluorescent microscopy (a to i). Panels (f) to (i) were viewed under a confocal microscope to visualise protein co-localisation.

Individual antibodies employed in panels (a) to (e) are as follows: polyclonal α-N015 (a) & (e), polyclonal α-P658 (b), monoclonal α-22k (c) and monoclonal α-Matrix (d).

The antibody combinations used in panels (f) to (i) are as follows: polyclonal α-N015 & monoclonal α-P001 (f), polyclonal α-N015 & monoclonal α-22k (g), polyclonal α-P658 & monoclonal α-22k (h) and polyclonal α-N015 & monoclonal α-Matrix (i).

Arrows indicate inclusion bodies (CIs) visible within the cytoplasm upon staining for N, P and M2-1. At least at this stage in infection, the M protein is absent from CIs. Each protein has a diffuse pattern of staining in the cytoplasm. N, P and M also occur in spikes protruding from the cell membrane.

In (f), N and P co-localisation is observed. N and M2-1 (g) and P and M2-1 (h) co-localise in the CIs but not in the membrane spikes. Little obvious co-localisation of N and M (i) is observed in this instance.
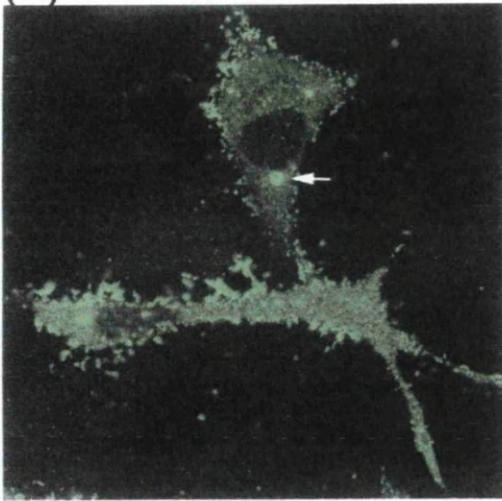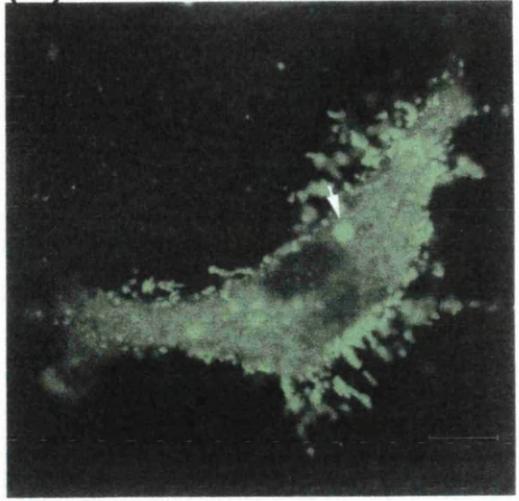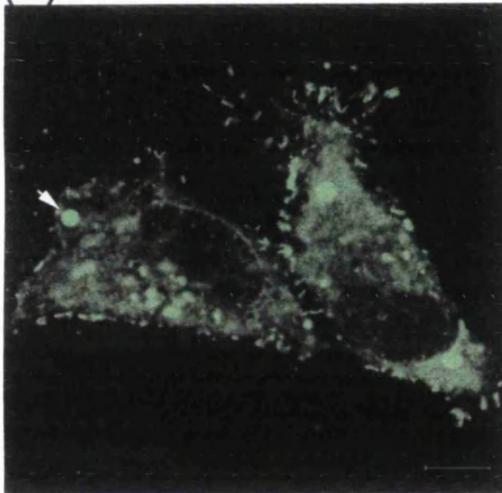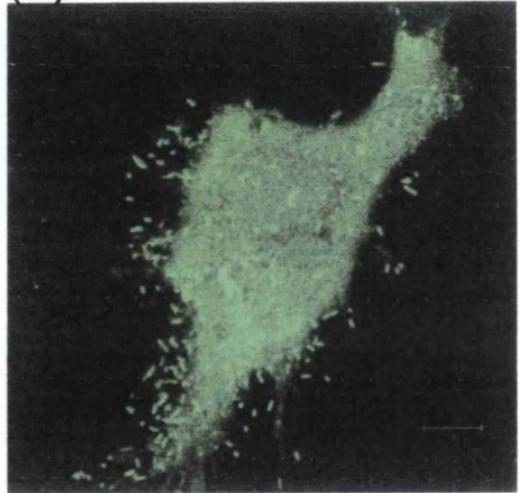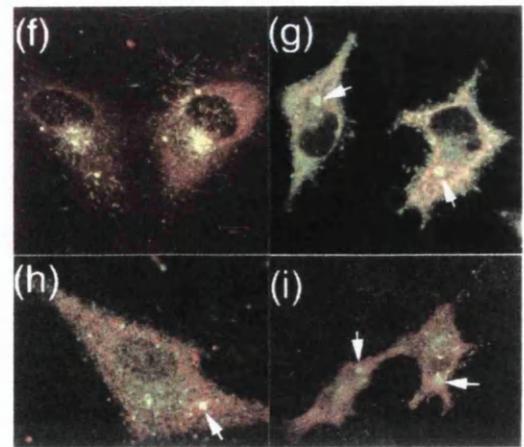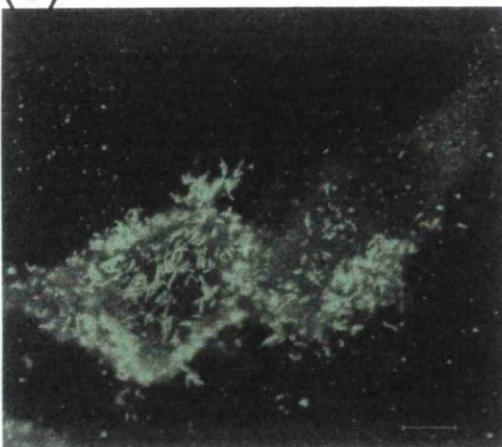
The distribution of M within infected cells was also analysed. Notably, the protein showed a diffuse cytoplasmic staining pattern and was detected in the spikes (Figure 3.3 d). However, at least in this stage of infection, M did not accumulate within the CIs (Figure 3.3 panel i). Thus, it is probable that the main site of N and M interaction is at the plasma membrane. An interaction between the RSV N and M proteins may be demonstrated *in vitro* (G. Henderson and C. Loney, personal communication). Coronel *et al.* (2001) demonstrated that the M protein of hPIV-1 interacts with the carboxy-terminus of the nucleocapsid protein to aid packaging of the nucleocapsid into virus-like particles.

The distribution of N in infected cells was investigated by IFA using MAbs that map to three antigenic sites (Figure 3.4). The MAbs were mapped by a peptide scanning methodology; the peptide positions within N are indicated in Figure 3.4. MAbs αN3 and αN9 showed very similar patterns of staining, similar to the pattern obtained by polyclonal antisera (Figure 3.3 a). In contrast, MAb αN15 had a different staining pattern. As did αN3 and αN9, staining with αN15 detected protein within CIs and membrane spikes, but the diffuse cytoplasmic stain was no longer evident. This suggested that αN15 detected a form of N different from that identified by αN3 and αN9. Subsequently, it was shown that αN15 competes with P for binding to N protein (Murray *et al.*, 2001), thus by inference, it is probable that the cytoplasmic pool of N is mostly complexed with P.

3.7 Discussion

Experiments presented within this chapter confirm several established observations of RSV growth. The plaque assay demonstrated *in vitro* growth of the virus to relatively low titre. The figure obtained of 4.7×10^6 pfu ml^{-1} is representative of RSV infection in tissue culture and is lower than viruses of the same taxonomic order. Despite the low yields, the expression of viral proteins may be detected by western blot, RIPA and IFA. The size of the RSV proteins observed by both western blot and RIPA agree with both the published observations and calculated molecular weights. In accordance with previous observations, the P protein migrated as a larger-sized protein than expected from its calculated molecular weight.

Figure 3.4: IF images of primary RSV infection in CV-1 cells probed with monoclonal antibodies directed against the N protein.

CV-1 cells, grown as a monolayer on cover-slips, were infected with RSV and incubated overnight at 33°C. Cells were fixed and probed with N protein-specific monoclonal antibodies, α-N9, α-N15 and α-N3. Cells were then probed with secondary antibody containing a fluorescent label and viewed by fluorescent microscopy. The monoclonal antibodies are specific to the following antigenic sites within N protein (391 aa): 16-35 aa (α-N9), 331-345 aa (α-N15) and ~346-365 aa (α-N3).

α-N9 (a) and α-N3 (c) demonstrate a similar staining pattern detecting CIs, membrane spikes and diffuse cytoplasmic-distribution. In contrast, α-N15 (b) does not detect this cytoplasmic staining, recognising only protein present in the CI bodies and the spikes.

MAb

αN9   αN15   αN3
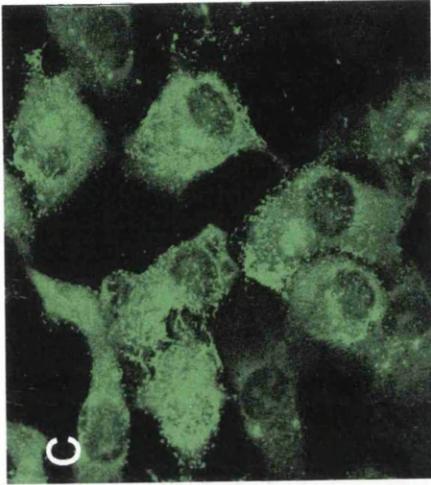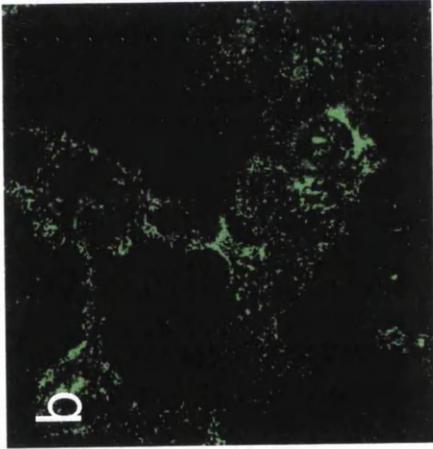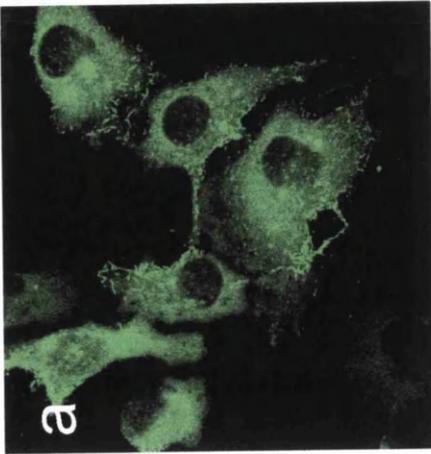
The RIPA revealed protein interactions that occur within RSV-infected cells. With the exception of MAb αN15, each N and P antibody strongly precipitated their reciprocal partner and to a lesser extent, M and M2-1 proteins. Antibodies directed against the M protein appeared to co-precipitate M and N. Anti-22k appeared to precipitate M2-1, N and M. The interaction between N and P has been characterised for a number of viruses though the exact relationship between the proteins needs to be further refined. Possible relationships are suggested between proteins N and M and M2-1, though their precise interaction may not be elucidated from this assay. Antibodies αMatrix and α22k do not co-precipitate the P protein indicating that P interacts either weakly, or not at all, with M and with M2-1. However, this may be a reflection of the antibodies used in this assay as a strong P-M2-1 interaction may be detected *in vitro* using purified recombinant proteins and the yeast two-hybrid system (C. Loney, personal communications). Hengst & Kiefer (2000) demonstrated that N and M2 interact directly but were unable to detect a P-M2 interaction in a similar yeast two-hybrid assay. Four proteins of sizes greater than 66 kDa, 30 kDa, 25 kDa and less than 14.3 kDa were routinely co-precipitated with the viral proteins. Both the identity and origin of each protein are unknown.

The MAb αN15 detects only N protein dissociated from P. This specificity results from P protein and the antibody sharing a common binding domain upon N, preventing their simultaneous association (although the epitopes are distinct, Murray *et al.*, 2001). Therefore, only free-N or that assembled into nucleocapsid may be recognised by this antibody. No additional proteins were precipitated and the presence of RNA was not investigated.

The co-localisation of viral proteins within infected cells was visualised by IFA. N and P proteins were distributed within the cytoplasm and at the cell membrane. Garcia-Barreno *et al.* (1996) showed that the co-expression of RSV N and P was required for CI body formation and that the carboxy-region of the P protein was important for this interaction. In contrast, the expression of Measles virus N resulted in the formation of CIs in the absence of P (Spehner *et al.*, 1997). Thus, the CIs may be regions of active transcription/replication and nucleocapsid assembly, or regions where genomes undergo quiescence with a possible lead into persistence. RSV is able to persist *in vitro* in macrophages (Guerrero-Plata *et al.*, 2001) whether it does so in the host is not known.

4 Production and Purification of the RSV Nucleocapsid Protein

4.1 Introduction

The principal direction of research was to investigate the interaction between RNA and the N protein of RSV. To address this aim, a variety of *in vitro* assays were selected, creating a requirement for relatively large amounts of purified N protein. This was achieved by cloning of the N protein open reading frame (ORF) and subsequent expression of recombinant protein. A bacterial expression system was selected as the N protein has no requirement for directed post-translational modification, thereby negating the need for a taxonomically higher expression system. Also, the bacterial system was selected due to expertise present in the lab at the time of research. Recombinant N protein was expressed with 10 histidine residues fused at its amino terminus, providing a tag for ease of identification and of purification.

4.2 Cloning of the RSV nucleocapsid protein gene ORF

The ORF of the N gene was cloned by RT-PCR using RNA isolated from RSV infected CV-1 cells. Total RNA was isolated from the infected cell layer by phenol extraction (Section 2.6.6). Recovery of RNA was confirmed by separation and visualisation in an ethidium bromide-stained formaldehyde agarose gel under ultra violet (UV) light (results not included). cDNA was prepared by reverse transcription from the RNA using AMV reverse transcriptase with random hexamer primers (Section 2.9.1). The newly synthesised cDNA was used as a template for PCR amplification of the RSV N protein ORF (Section 2.9.2). Primers complementary to sequences at the gene termini were used to amplify full-length N, including transcription initiation and stop codons. Each primer introduced a *Bam* HI restriction sequence to facilitate the cloning of N (Appendix). The amplification products were separated and visualised in an ethidium bromide-stained agarose gel under long wave UV light. In each reaction, a single PCR product of the desired size was obtained (1244 base pairs, bp). The resulting DNA band was excised from the gel and extracted from the agarose by Gene Clean (BIO 101).

Gene Clean DNA was digested with *Bam* HI and subsequently extracted from solution by ethanol precipitation. The resulting DNA was viewed by ethidium bromide-stained

agarose gel electrophoresis (Figure 4.1). Figure 4.1 shows the N ORF fragment and plasmid pGEM 3Z (Promega), digested with *Bam* HI and ethanol precipitated.

4.2.1   Cloning of N ORF into pGEM 3Z

The digested N gene and plasmid pGEM 3Z were combined in a ligation reaction. The N gene was introduced into pGEM 3Z as the plasmid was readily available in the laboratory making it a good candidate for use to increase familiarity with cloning techniques. The reaction products were used to transform competent DH5 cells by electroporation which were selected on L-broth agar with ampicillin (Section 2.7.6.1). A number of colonies were picked with sterile tooth-picks, placed individually in 10 ml selective 2YT broth and incubated overnight at 37°C. Plasmid DNA was extracted from 1 to 5 ml of the overnight culture by Qiagen Miniprep.

Each plasmid mini-preparation was digested with *Bam* HI to confirm the presence of the fragment in the plasmid (Figure 4.2). In Figure 4.2, two bands were present in lanes b, e, f and i. The upper band was linearised plasmid (2743 bp) and the lower was N gene insert of the expected size (ca. 1220 bp). Lanes g and h possessed a band of linearised plasmid DNA and a very faint DNA band of less than 500 bp in length of which the identity is unknown. The single DNA bands in lanes a, d and j were larger than the linearised plasmid. Their size was approximately that of ligated plasmid and insert, which suggested the plasmid had been linearised by digestion at only one restriction site.

4.2.2   Cloning of N ORF into pET16b

N insert was isolated from the clone in lane i, Figure 4.2. The DNA band was excised from agarose gel, Gene Cleaned and subsequently cloned into pET16b. Due to the design of primers used in the PCR of the original fragment, the N ORF was introduced in-frame with an amino-terminal histidine-tag sequence present in the vector. DH5 cells were transformed by electroporation and selected on L-broth agar. Individual bacterial colonies were picked, grown in selective liquid culture and plasmid DNA extracted by Miniprep. Plasmid DNA was digested with the restriction enzymes *Bam* HI, to confirm presence of insert, or *Xba* I, to determine N ORF orientation relative to the T7 promotor (Figure 4.3). The N ORF possesses an *Xba* I site approximately 400 bp from the AUG codon.
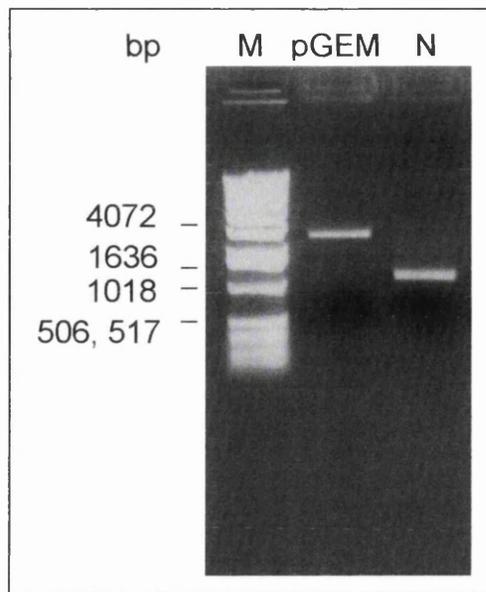
Figure 4.1: N and pGEM 3Z DNA digested with *Bam* HI and ethanol precipitated.
DNA was separated in a 1% agarose TAE gel, stained with ethidium bromide and visualised by UV light. Digested N is 1220 bp and pGEM 3Z is 2743 bp in length. (Marker, M).
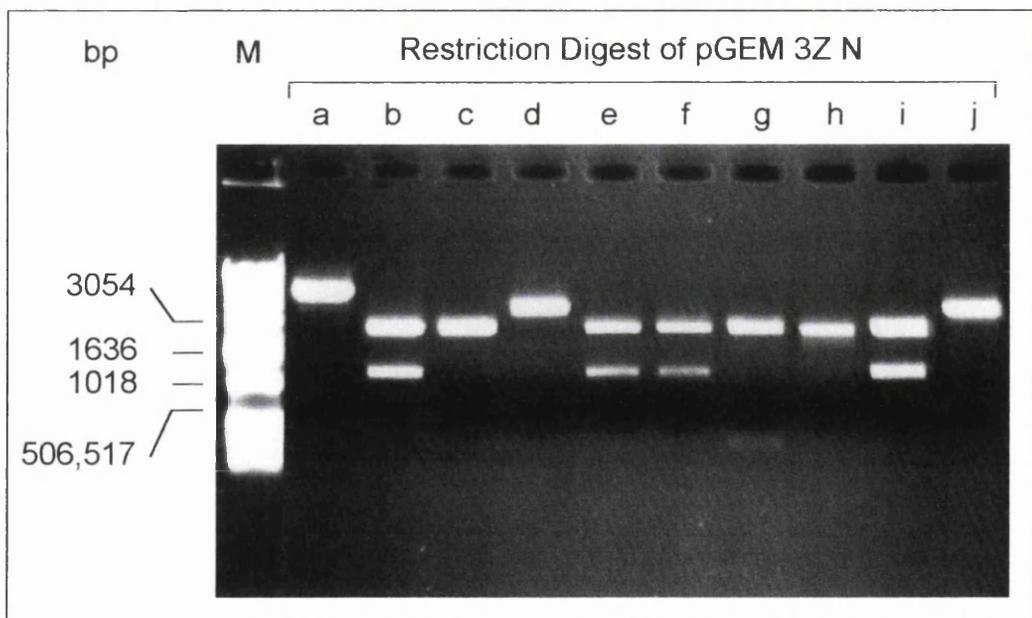
Figure 4.2: Digestion of pGEM 3Z N with *Bam* HI.
Digested plasmid samples a to j were separated by 1% agarose TAE gel, stained with ethidium bromide and visualised by UV light. Linearised plasmid and N insert of the correct size are present in lanes b, e, f and i. (Marker, M).
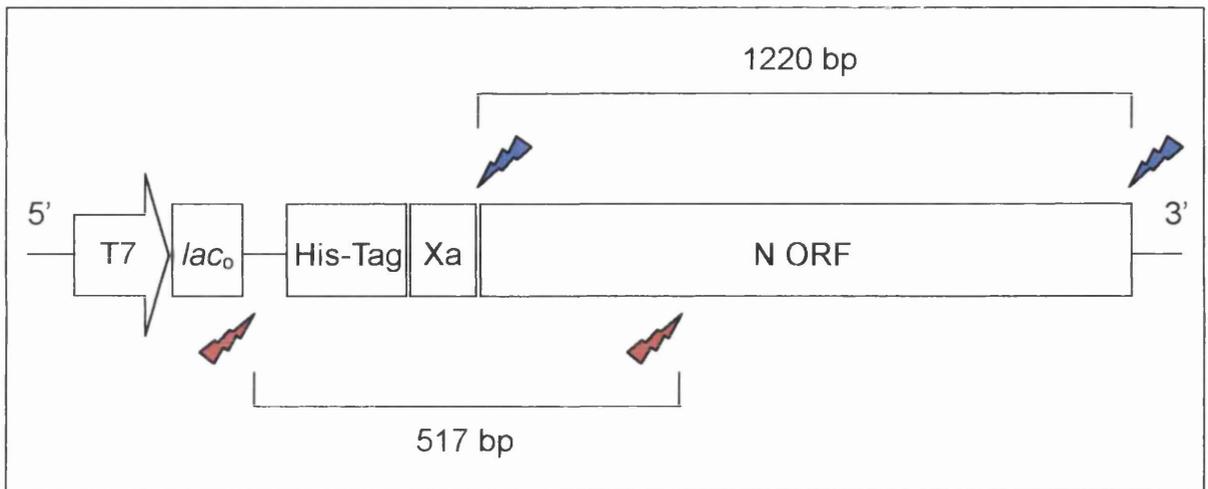
Figure 4.3: A schematic of the RSV N ORF in plasmid vector pET16b.

Transcription is controlled by the T7 promoter, adjacent to a *lac* operator, and commences upstream of the Histidine-tag. The His-tag is separated from the N ORF by the protease Factor Xa cleavage site. *Bam* HI endonuclease sites are indicated in blue and *Xba* I sites, in red. The size and position of restriction enzyme cleavage products are marked.

The *Xba* I orientation digests of five pET16b N clones are shown in Figure 4.4. Samples 1 and 3 possessed a band of approximately 510 bp that corresponds with excision of N in the correct orientation with the T7 promoter and His-tag. The digest patterns of the remaining clones represented N in an incorrect orientation (lanes 2 and 4) or a mixed plasmid population (lane 5).

To confirm N insert orientation in pET16b N samples 1 and 3 (Figure 4.4), DNA sequencing was performed both manually and by an automated system (Section 2.9.10). Primers to T7 promoter and terminator regions in the plasmid and to internal sequences within N were employed (Appendix). In addition to the correct orientation of insert in the plasmid, results confirmed the sequence of N was in agreement with that published in Genbank (accession code M74568) and that N insert was present in-frame with the histidine-tag.

4.3 Recombinant N Protein (His-N) Expression

pET16b N (Figure 4.5) was introduced by heat shock into competent BL21 (DE3) pLysS and selected on L-broth agar with appropriate antibiotics. Individual colonies were picked from the plate, placed in 10 ml selective 2YT medium and incubated overnight at 37°C. 100 ml selective 2YT medium was inoculated with overnight culture to create a 1 / 200 dilution. The transformed bacteria were incubated until an O.D.$_{600}$ of 0.8 was obtained. Protein expression was then induced by addition of 1 mM IPTG and incubation for a further three hours at 37°C.

The expression of recombinant His-tag N protein was investigated by western blot of bacterial cell lysates (Figure 4.6). Antibodies directed against RSV and the histidine-tag were employed to confirm expression. At the time of investigation, antibodies specific to N protein had not been generated therefore a commercial anti-RSV antibody mix was used (Figure 4.6 a). Three bands were visible in the positive control. The upper faint band was that of N (43.4 kDa), smaller than bacterially-expressed protein due to the absence of a histidine-tag. Lower bands were the viral proteins P (38 kDa) and M (28.7 kDa). A single band of approximately 46 kDa was detected in lanes with lysates prepared from induced bacteria. The size corresponded to that of N (43.4 kDa) plus approximately 2 kDa derived from the His-tag fusion.
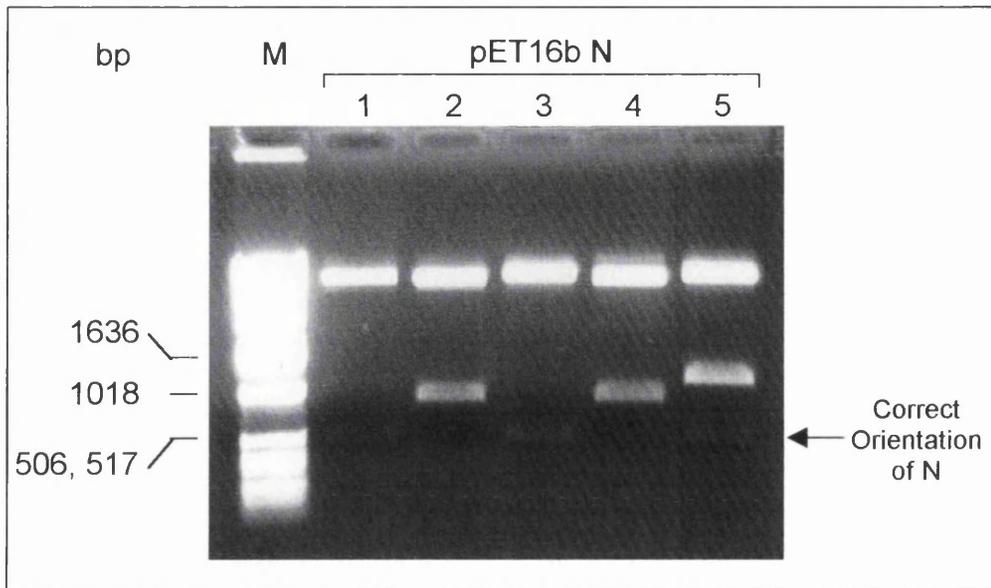
Figure 4.4: Orientation digestion of pET16b N with restriction enzyme *Xba* I.
Digested clones were separated in a 1% agarose TAE gel, stained with ethidium bromide and visualised by UV light. The upper DNA band in each sample is linearised pET16b with a fragment of the N gene. Lower bands in lanes 1 and 3 (~510 bp) represent *Xba* I excision of N, present in the correct orientation in the plasmid. (Marker, M)
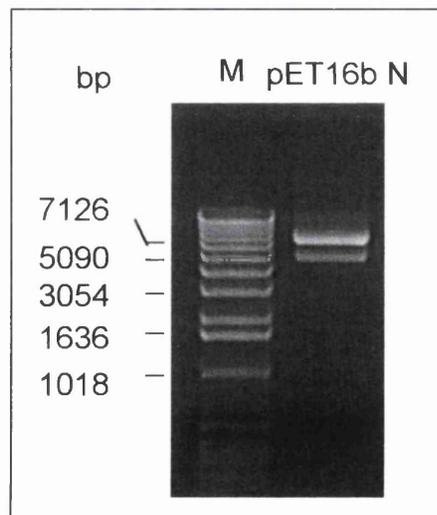
Figure 4.5: pET16b N plasmid DNA.
Plasmid DNA (6934 bp) was separated in a 1% agarose TAE gel, stained with ethidium bromide and visualised by UV light. DNA bands represent the plasmid in super-coiled and open-circle conformations. (Marker, M).
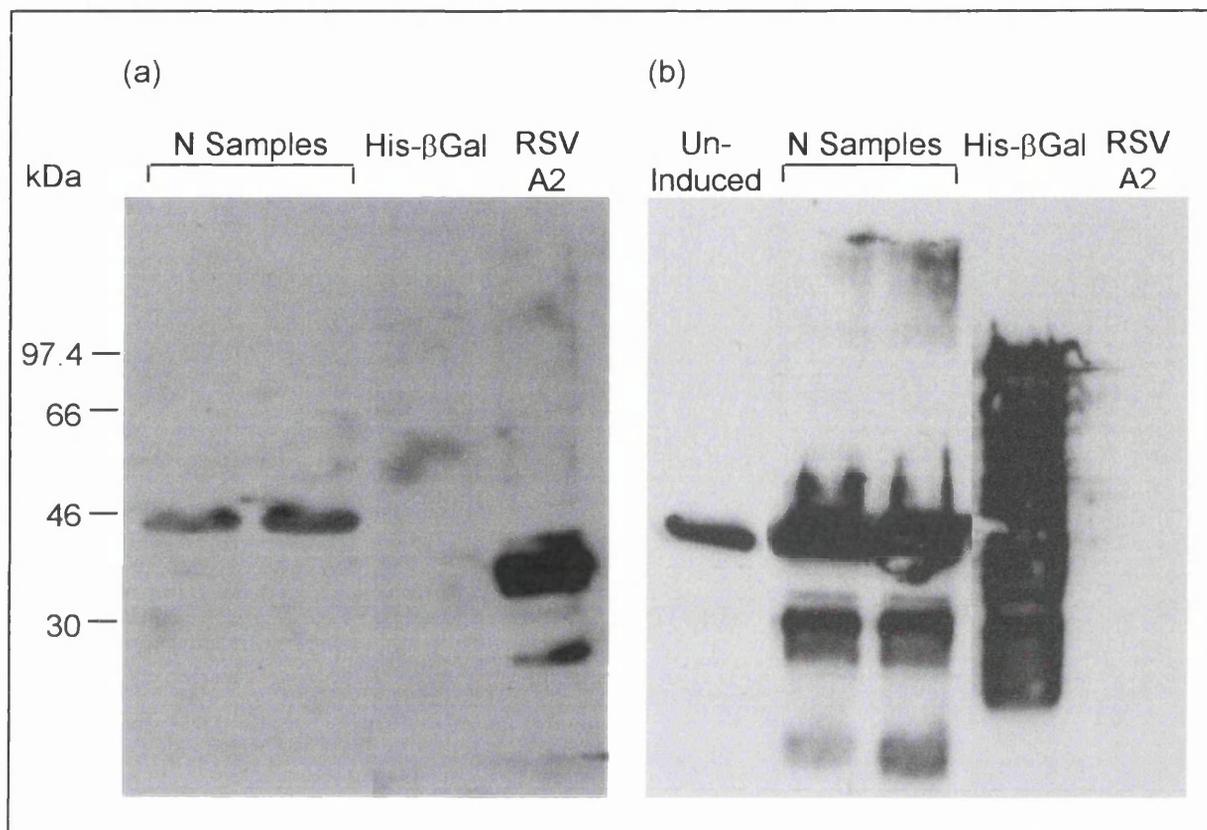
Figure 4.6: Western blot analysis of transformed bacterial cell lysates.

Whole lysate was separated by 10% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane, blocked with 5% Marvel in TPBS and probed with anti-RSV (a) and with anti-His-tag (b) antibodies. Membranes were probed with an appropriate secondary antibody conjugated to HRP and the protein bands visualised by ECL and autoradiography. Lysates of transformed bacteria expressing recombinant histidine-tagged β-galactosidase (His-βGal) and of CV-1 cells infected with RSV strain A2 were used as controls. In (b), His-N protein is detected in the un-induced bacterial lysate.

A band of 46 kDa was also detected in bacterially-derived N samples when probed with anti-His-tag antibody (Figure 4.6 b). An additional band of approximately 30 kDa was also observed. This smaller band is unlikely to be of bacterial origin as it was absent from the un-induced sample. More probably, it is an N protein breakdown product. Detection by the anti-His MAb suggested it represents the amino-terminus of His-N, possibly after some proteolytic event that has removed the C-terminal region. A band of full-length His-N was present in the un-induced sample suggesting the *lac* operon control of protein expression is not as stringent as was expected.

4.4 N Protein Purification by Metal Chelate Liquid Chromatography

The experimental techniques selected to address the aims of this project required the production of purified protein. The histidine-tagged N protein was purified by means of metal ion chelate affinity chromatography.

The recombinant protein was expressed in bacteria as previously described. The bacteria were harvested by pelleting, lysed by freeze/thaw and the resulting solid debris removed by centrifugation. The clarified lysate was applied to a HiTrap column (Pharmacia Biotech) and washed with buffers of increasing imidazole concentration. The flow-through was retained and investigated by SDS-PAGE and coomassie blue staining. An example of the manual purification of His-N is shown in Figure 4.7. His-N protein (46 kDa) was eluted at an imidazole concentration of 300 mM and greater. Protein was noticeably retained in the wells of the protein gel and at the interface between the stacking and separating gels, possibly as a result of N protein aggregation. After fractions 1 and 2 at 300 mM, the protein appeared to be relatively devoid of bacterial protein contamination. Considerable amounts of His-N appeared in the column flow-through. A faint band of approximately 30 kDa was visible in the initial 300 mM fractions, potentially the same band detected by the His-tag antibody in Figure 4.6.

In the latter stages of the project the AKTA FPLC apparatus was used in the purification of His-N. Figure 4.8 shows the output of the FPLC in the form of a chromatogram. Subsequent SDS-PAGE analysis confirmed the presence of purified His-N in the second peak particularly fractions 10 to 16 (results not included). The nucleic acid trace mapped closely to that created by protein. This may indicate protein-
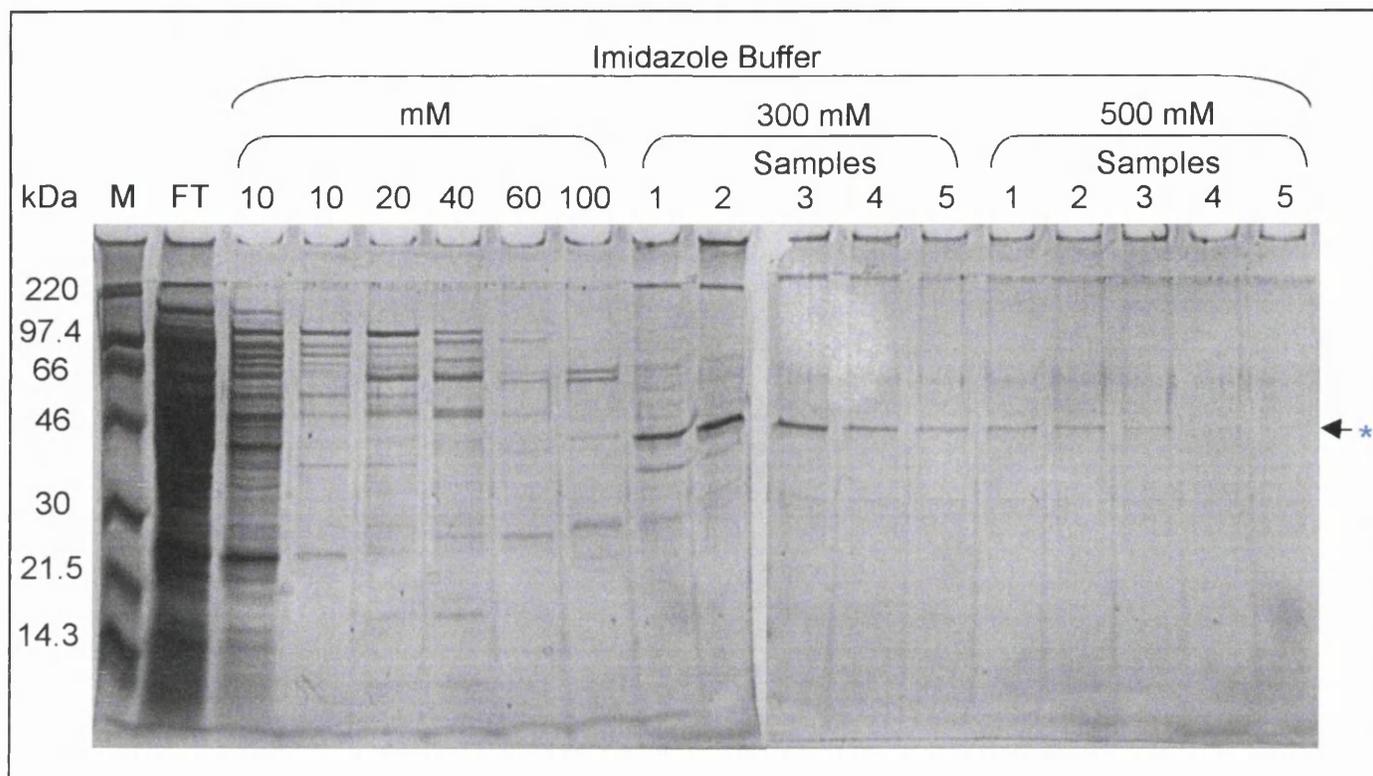
Figure 4.7: Manual purification of Histidine-tag N protein.

The clarified and filtered lysate of bacteria expressing recombinant His-N protein was manually applied to a HiTrap metal ion affinity column (Pharmacia Biotech). The column was washed with 5 ml aliquots of buffer of increasing imidazole concentration. The flow-through of buffers with 300 mM and 500 mM imidazole was collected in 1 ml fractions. 20 μl each fraction was separated by 10% polyacrylamide SDS-PAGE. Proteins were visualised by staining the gels with coomassie blue. The position of His-N within the gels is marked by the symbol '*'. (Marker, M; Flow-through, FT).
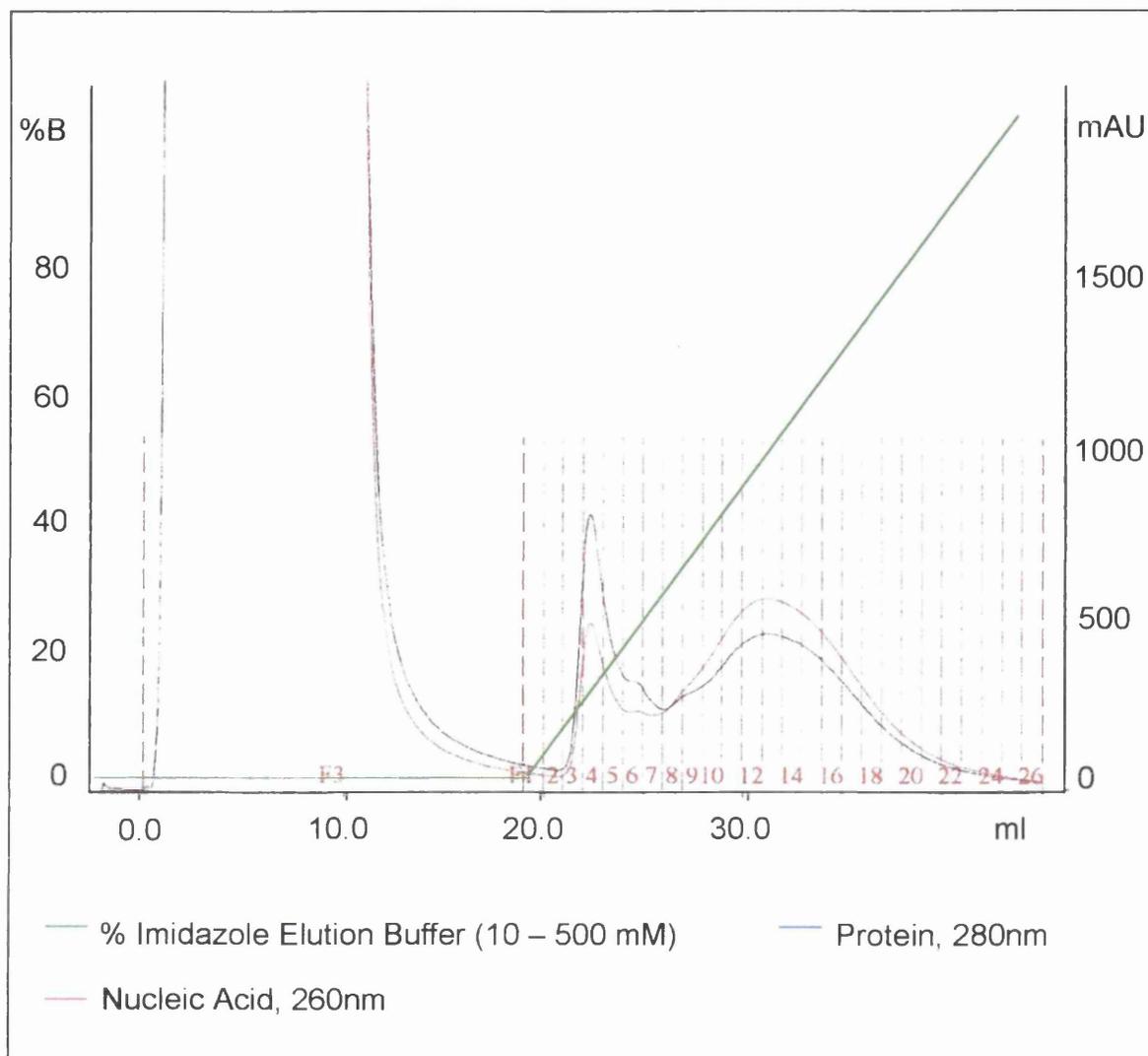
Figure 4.8: Metal ion chelate liquid chromatography of Histidine-tag N protein by FPLC.

9 ml clarified and filtered transformed bacterial cell lysate was loaded on to a HiTrap metal ion affinity column and washed with 10 ml 10 mM imidazole buffer. An elution buffer gradient of 10 mM to 500 mM imidazole was applied to the column and the flow-through collected in 1 ml fractions. Buffer volume applied to the column is followed on the x-axis. Absorbency (absorbence units $x10^{-3}$, mAU) and buffer gradient concentration are monitored on the y-axes.

nucleic acid association, especially in terms of the N protein peak. The purity of N protein obtained in this manner was similar to that obtained manually.

Following purification, His-N fractions were concentrated and buffer exchange into PBS achieved using Centricon 10 spin concentrators (Millipore). A yield of approximately 2.5 mg L^{-1} purified His-tag N protein was routinely obtained.

The purity of recombinant His-N was further investigated by the separation of protein by SDS-PAGE and it's visualisation by silver-stain (Figure 4.9). Silver-staining is more sensitive than staining with coomassie blue (1 to 5 ng compared with 40 to 50 ng, respectively). Therefore, silver-stain was used to investigate the possible presence of low concentrations of contaminating proteins. In Figure 4.9, His-N may be seen as an over-stained protein band at 46 kDa in each lane. Also present were bands of smaller proteins. These may be residual bacterial proteins, surviving purification or they may be N protein breakdown products. In particular, there was a band present at 30 kDa, the same size as the potential amino-terminus breakdown product of N. Neither the number nor intensity of these smaller bands altered with the age of the protein sample, suggesting they were not accumulating post-purification.

The level of purity achieved by metal ion chelate affinity chromatography was considered acceptable to enable subsequent investigation of the RNA-binding properties of the N protein.

4.5 Discussion

The nucleocapsid protein ORF of RSV strain A2 was successfully cloned into the vectors pGEM 3Z and pET16b. Digestion with *Xba* I demonstrated the correct orientation of N insert with respect to the T7 promoter within pET16b. This was validated by sequence analysis. In addition, this data confirmed the N amino acid sequence as identical to that published in Genbank (M75468) and that insertion occurred in-frame with the histidine-tag. The His-tag is present at the amino terminus of the recombinant fusion protein.

Expression of histidine-tag N protein was performed in *E. coli* strain BL21 (DE3) pLysS. This bacterial system was selected as it should allow relatively stringent control of
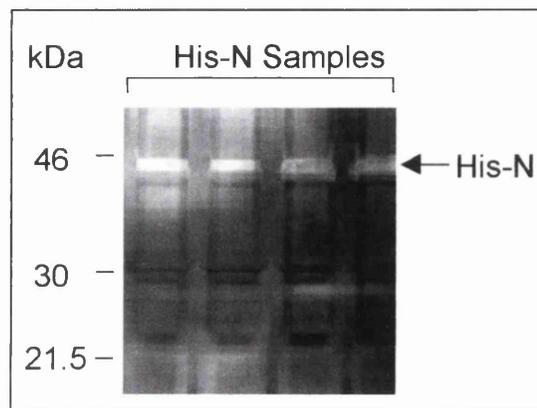
Figure 4.9: Histidine-tag N protein purified by metal ion chelate FPLC.
Purified His-N protein samples were separated by 10% polyacrylamide SDS-PAGE and visualised by silver-staining.

recombinant protein expression. A system with such control was selected due to the possible toxicity of the N protein to the host cell.

The RNA polymerase gene of bacteriophage T7 is present within the bacterial chromosome under *lacUV5* promoter control. Upon transcription by host RNA polymerase and subsequent expression, the T7 RNA polymerase effects the transcription of N, itself controlled by T7 promoter. The expression system selected affords three opportunities of transcriptional control. *Lac I*, encoding the *lac* repressor protein, is present in the host chromosome. In the absence of an inducer, the *lac* repressor binds the operator sequence present in the *lac* promoter controlling the T7 RNA polymerase gene, thereby preventing transcription. Should low levels of transcription and hence protein expression occur, an additional control is available. The plasmid pLysS present in the bacteria expresses low levels of T7 lysosyme, a natural inhibitor of T7 RNA polymerase. Lysosyme binds and inactivates any escaping T7 polymerase protein. As a final level of control, transcription of recombinant protein gene is controlled by the T7 promoter immediately adjacent to a *lac* operator sequence. Again in the absence of inducer, the repressor, expressed from *Lac I* present in pET16b, binds the operator sequence disrupting N transcription. Repressor control at the *lac* promoter is overcome upon addition of the inducer IPTG enabling expression of T7 RNA polymerase and hence recombinant protein. The increase in T7 polymerase level following IPTG addition overcomes the level of restriction offered by T7 lysosyme. Despite such controls being in place, the expression of His-N protein was detected by western blot in un-induced samples indicating that the system was not effectively regulating control. Despite this expression, the bacteria grew well in culture suggesting that the protein was not especially toxic to host survival.

A relatively large amount of recombinant protein was detected in the pellets of solid bacterial lysate (data not shown). Accumulation of His-N within this insoluble fraction is probably due to protein aggregation in inclusion bodies. However, this was not further investigated as sufficient recombinant protein was present in the soluble fraction to enable successful purification. 2.5 mg L^{-1} purified protein was routinely obtained from the bacterial expression system.

Examination of His-N protein by SDS-PAGE and silver-stain showed the presence of faint bands of unidentified protein. These proteins may be of bacterial origin or

truncated His-N protein. A band of approximately 30 kDa, visible by silver stain, may be the same as that detected by western blot. The proteins detection with antibody directed to the histidine-tag suggests that it is the truncated amino terminus of His-N, cleaved approximately two-thirds along its length. The number of additional proteins and their intensity was sufficiently low that the level of His-N protein sample purity was considered acceptable for use in subsequent experiments.

5 The Specificity of RNA-Binding by RSV N Protein

5.1 Introduction

The interaction between RNA and the N protein of RSV was investigated using a variety of *in vitro* assays. Other N proteins from the *Mononegavirales* have shown a lack of specificity when expressed as individual proteins (Masters & Banerjee, 1988; Spehner *et al.*, 1991; see Introduction Section 1.12), thus it was of interest to determine whether RSV N acted similarly. For this purpose, histidine-tag N protein, expressed in a bacterial system and purified by NTA-nickel affinity chromatography (Chapter 4), and radiolabelled RNA probes (Section 2.11.1) were employed. Two probe sequences were selected to represent viral and non-viral RNA. One was described as 'specific RNA' and represented the first 47 nucleotides of the positive sense, RSV leader sequence. A second probe was described as 'non-specific RNA' and was transcribed from the *Sal* I digest of Bluescript plasmid, creating a 32 nucleotide sequence. Radiolabelled probes were used to investigate the specificity in RNA-binding of the N protein and to determine the existence of RNA sequence preference. A second objective of this work was to determine conditions that would prove favourable for the *in vitro* binding of RNA for future studies on locating the RNA-binding domain of N. Thus, assays had to meet the criteria of reliability and reproducibility.

5.2 Filter-Binding Assays

The association between RNA and the N protein was investigated using filter-binding assays. Purified histidine-tag N protein and radiolabelled specific probe were combined and allowed to associate, as described in Section 2.11.6. Samples were passed through nitrocellulose membrane present in dot blot apparatus (Figure 5.1). The figure demonstrates that N protein is able to capture RNA and that binding is observed only after incubation at 33°C for 20 minutes. Passing RNA over N protein, previously captured by nitrocellulose, showed only a marginal increase of signal above background. The specificity of this RNA-protein interaction with His-N was apparent when it was observed that the negative control, His-tag β-galactosidase, failed to bind RNA under identical conditions. There was little difference in dot intensity upon increase in N protein concentration from 0.1 to 0.5 μg perhaps suggesting that the RNA was saturated with protein.
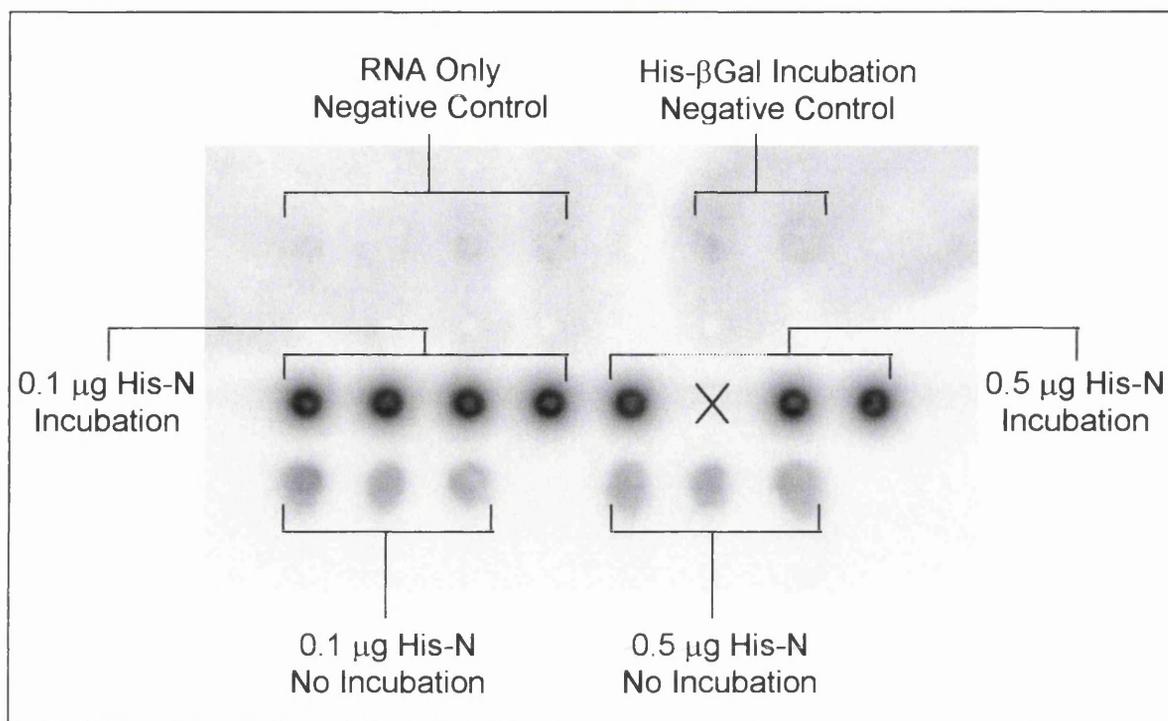
Figure 5.1: A dot blot filter-binding assay with His-tag N protein and specific RNA.

0.1 µg and 0.5 µg His-N protein samples were incubated with 5000 cpm radiolabelled specific RNA for 20 minutes at 33°C. Samples were applied under vacuum to pre-wetted nitrocellulose membrane in dot-blot apparatus. Certain samples were not incubated prior to application to the membrane. In such instances, His-N protein was applied to the membrane followed by radiolabelled RNA. The membrane was washed with 5% TCA and examined by phosphorimagery.

RNA only and RNA with histidine-tag β–galactosidase (His-βGal) were employed as negative controls. Controls confirm that RNA alone is not captured by the membrane and that RNA is not retained due to the presence of histidine-tag non-specific protein.

An alternative capture protocol, employing glass fibre filters, was not successful (Figure 5.2). Samples were prepared in the same manner as for the dot blot assay, captured onto individual filters and the radioactivity measured in a scintillation counter. The radioactivity of N protein-RNA samples increased only slightly with an increase in protein concentration. Levels of radioactivity in N protein samples were very low, ranging between 1.2 and 1.6% of the total radioactivity added. This may result from only weak protein-RNA association, poor retention of complexes by the filters or may indicate protein saturation by only small quantities of RNA. Radioactivity levels obtained with N protein were only slightly above background.

Results obtained by dot blot and by glass filter-binding assays proved un-reproducible over a number of attempts. Frequently in the dot blot assay, no difference was observed between negative controls and reaction samples. On occasion, a negative control would be visible as an intense dot possibly as a result of dust on the membrane. The individual filter-binding assays were similarly erratic. High levels of radioactivity were frequently retained in the negative controls. Conversely, low radioactivity levels occurred in the N protein samples, with no observable correlation with N protein concentration. In each assay, retention of protein by the membrane was demonstrated by western blot (data not included). This suggested that the absence of radioactivity from the membrane was due to N protein-RNA complexes not being formed.

The filter-binding assays produced reasonable results upon occasion. However, the ability of RNA and His-N to associate *in vitro* was confirmed and investigated more satisfactorily using subsequent techniques. Therefore, the filter-binding assays were abandoned.

5.3 Yeast Three Hybrid System

The RNA sequence specificity of N was investigated using the yeast three-hybrid system (SenGupta *et al.*, 1996); the components of the system were kindly provided by M. Wickens (State University of New York, USA). The system has strong analogies to the yeast two-hybrid system, which is used to investigate protein-protein interactions. Within the yeast three-hybrid system, two recombinant proteins, when in close proximity to one another, form an active transcriptional activator that induces
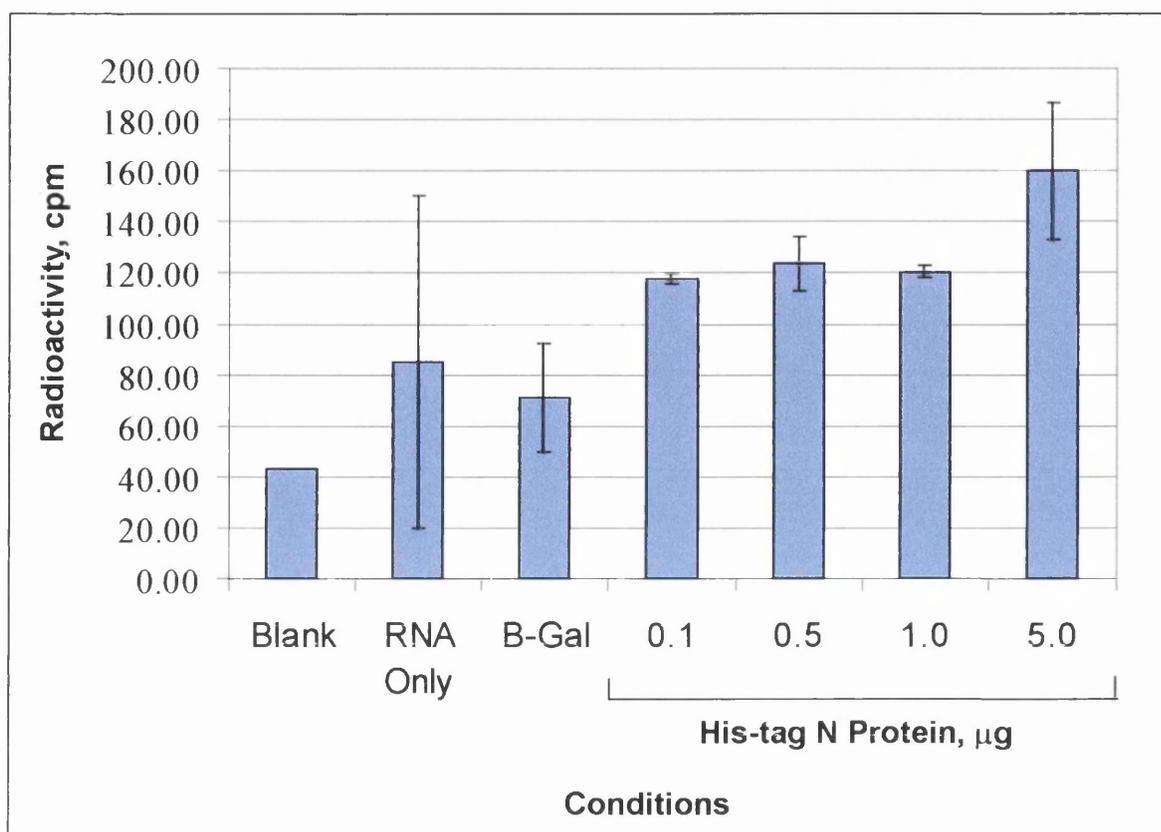
Figure 5.2: A graph of data obtained from a filter-binding assay with His-tag N protein and specific radiolabelled RNA.

0.1 µg to 5.0 µg His-N protein samples were incubated with 10 000 cpm radiolabelled specific RNA for 20 minutes at 33°C. Samples were applied under vacuum to pre-wetted individual filters. Filters were washed with 5% TCA and retained radioactivity measured in the scintillation counter.

RNA only and His-tag β-galactosidase (B-Gal) provided negative controls. Experimental conditions were performed in triplicate.

expression of reporter genes *LacZ* and *His3*. Unlike the yeast two-hybrid system, the proteins do not interact directly but via a third moiety, a hybrid RNA (see schematic in Figure 5.3). It is the ability of protein to bind RNA that is measured by the reporter gene activity. The yeast strain L40-coat expresses the coat protein of bacteriophage MS2 fused to Lex A DNA-binding protein. This fusion protein binds the Lex A operon in the genome and provides an anchor for the recombinant RNA. The latter is the MS2 coat protein RNA-binding site fused to the RNA sequence of interest. In this assay, the positive-sense RSV genome leader sequence was used as 'specific RNA' (designated L3-2 RSV) and the 3' non-coding region of the hepatitis C virus genome was used as 'non-specific RNA' (designated L40 HCV) (provided by J. Wood, MRC Virology Institute). N proteins were expressed as carboxy-terminal fusions of the GAL4p transcription activation domain from plasmids pACT2 (A2 strain N protein, provided by P. Yeo) and from pGAD424 (RSS-2 strain N protein, provided by A. Easton, University of Warwick).

Controls were employed to confirm the system was working as anticipated. L40-coat and L40-coat transformed with vector pACT2 lacking insert were used as negative controls. Control success was varied, as may be seen in Figure 5.4. With L40 coat (i.e. no RNA), the colony-lift filter assay with RSS-2 N (zero positive results out of five repeat experiments), PpACT2 (zero out of one) and pACT2 (zero out of one) were negative as expected. However, more than half of those performed with RSV strain A2 N (four out of seven) were positive. With pACT2 (i.e. no N), the colony-lift filter assays with L40-Coat (zero out of one) and L40-HCV (zero out of four) were negative. However, one third of those performed with L3-2-RSV (two out of six) were positive.

Despite varying success of the controls, the N and P proteins of RSV were investigated using the yeast three-hybrid system. Duplicate experiments were performed with the P protein of RSV strain A2 in L3-2 RSV and in L40 HCV. The colony-lift filter assays were negative suggesting that P protein does not bind RNA. The N protein of RSV strain A2 produced positive results when expressed in L3-2 RSV (seven out of 11) and L40 HCV (six out of six). Should this be representative of a true protein-RNA interaction, the results suggest that N protein is able to bind both specific and non-specific RNA. However, N pACT2 also produced positive results when present in L40 coat (four out of seven). Upon transfection with PpACT2 or pACT2, L40-coat produced only negative results. This would suggest positive results obtained with N pACT2 were due to the N
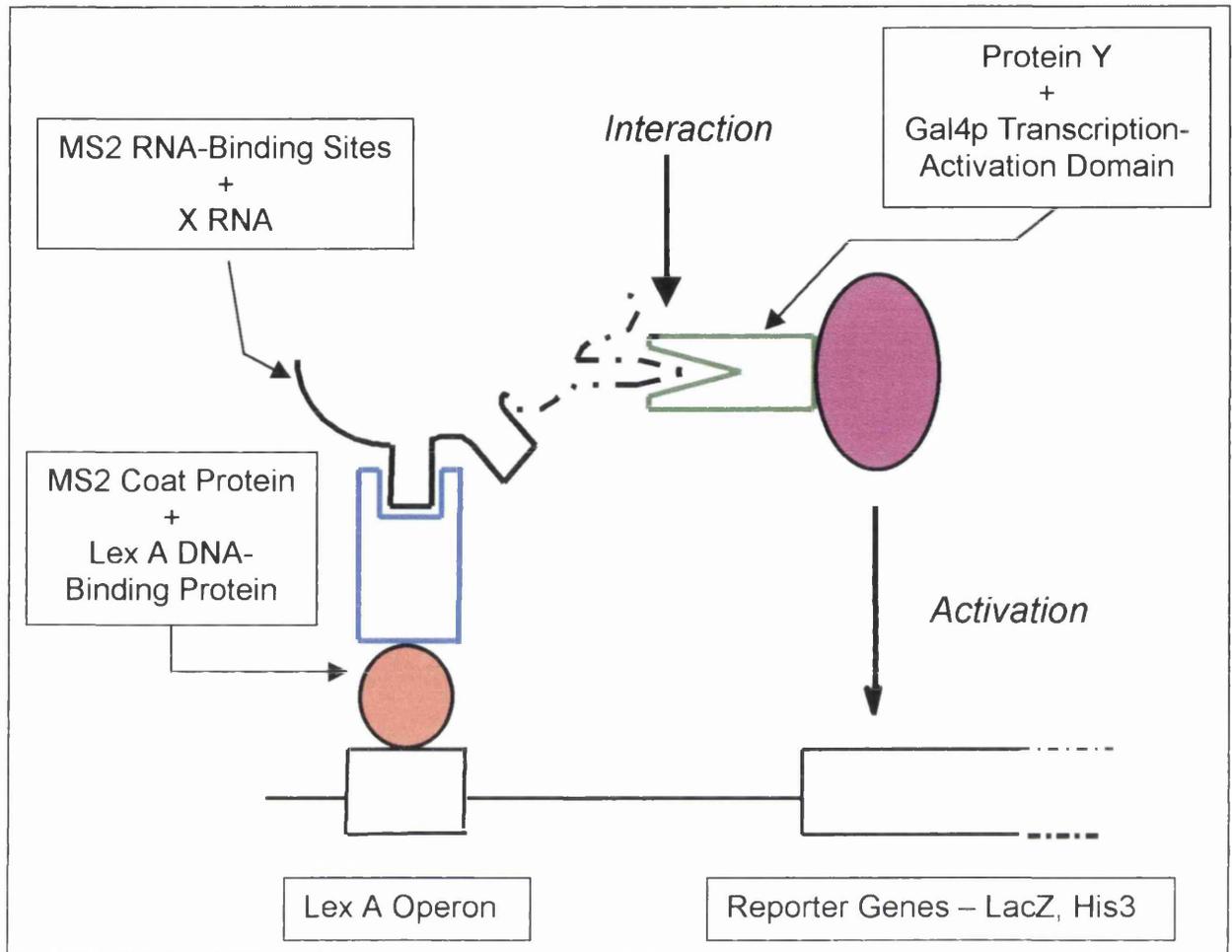
Figure 5.3: A cartoon depicting the yeast three hybrid system.

In the yeast three-hybrid system, two recombinant proteins form a transcriptional activator, when in close proximity to one another, that induces expression of reporter genes *LacZ* and *His*3. The proteins do not interact directly but via a third moiety, a hybrid RNA. It is the ability of protein to bind RNA that is measured by the reporter gene activity.

The yeast strain L40-coat expresses the coat protein of bacteriophage MS2 fused to Lex A DNA-binding protein. This fusion protein binds the Lex A operon in the genome and provides an anchor for the recombinant RNA, MS2 coat protein RNA-binding site fused to the RNA sequence of interest (X). The second fusion protein is the Gal4p transcription-activation domain fused to the protein sequence of interest (Y), expressed from plasmid DNA. Direct interaction between RNA X and protein Y enables reporter gene expression.

Figure 5.4: A scatter graph of data obtained with the yeast three hybrid system.
The transformed yeast strains L3-2-RSV and L40-HCV each carry a plasmid encoding different fusion RNAs. L3-2-RSV transcribes fusion RNA derived from RSV. L40-HCV encodes fusion RNA derived from hepatitis C virus. These yeast and their parent strain L40-Coat were transformed with plasmids encoding different fusion proteins. Interaction between the recombinant fusion protein and fusion RNA results in expression of the reporter gene β-galactosidase. The expression of β-galactosidase was investigated using the colony-lift filter assay and the extent of expression was estimated on a scale of between 0 and 3.

Fusion proteins were expressed from plasmids pGAD424 and pACT2. The N protein of RSV strain RSS2 was present in pGAD424. The N and P proteins of RSV strain A2 were present in pACT2. The plasmid pACT2 alone provided a negative control.

protein fusion. These observations may result from N protein association with the Lex A DNA-binding fusion protein or perhaps the bridging of both proteins with a non-specific RNA. Both situations would bring the transcription activation domain fused to N into close proximity with the reporter gene promoters. Upon expression in bacteria and subsequent purification, the N protein was observed to adhere to surrounding proteins. This would lend credibility to the explanation of non-specific protein-protein associations causing the observed positive results. Indeed, in a yeast two-hybrid screen with A2 N, performed concurrently with these experiments, it was noted that most of the interactions observed were non-specific in nature (C. Loney, personal communication).

In contrast, the N protein of RSS-2 produced predominantly negative results upon expression in yeast. Only one out of nine experiments was positive with L3-2 RSV and one out of six, positive with L40 HCV. The sequences of the N proteins of strains A2 and RSS-2 are very similar (A. Easton, personal communication). However, without knowing the position of the important domains on N it is difficult to conclude which residues contribute to binding RNA. Primary protein differences are unlikely to provide an explanation for the contrasting results. A possible explanation is that there was a difference in the level of protein expression. This was not investigated directly. However, in the yeast two-hybrid system, an interaction was demonstrated between the A2 P and the N proteins of both A2 and RSS-2 indicating an acceptable expression level of both proteins with no discernable differences in reactivities (C. Loney, personal communication).

The erratic nature of the results and the unreliability of the controls led to the assay being pursued no further. This system appears to preferentially enable the study of RNAs with a high degree of structure, such as the HCV IRES. The RSV leader has no significant structure. Recently the role of secondary structure was discounted by work on hPIV3 (Hoffman & Banerjee, 2000). In addition, there is the possibility that both the leader and trailer sequences must be present for functionality. These points may contribute to the lack of success observed in this assay. However, the data obtained suggested that the N protein of RSV is able to bind RNA in a non-sequence specific manner.

5.4 Gel Mobility Shift Assay

The interaction between His-N protein and radiolabelled RNA was investigated by gel mobility shift assay (GMSA). In a similar manner to the filter-binding assays, recombinant protein and RNA were combined and allowed to associate (Section 2.11.3). Samples were analysed by non-reducing polyacrylamide gel electrophoresis and visualised by phosphorimagery. When bound by protein, migration of RNA through the gel would be retarded. Therefore, a shift may be observed from free radioactive RNA at the gel bottom to a higher position within the gel. The assay was initially performed using bacterial lysates of BL21 (DE3) pLysS expressing His-N protein from pET16b N (Figure 5.5). Bacteria lacking the plasmid were employed as a negative control. Samples 1 to 5 represented increasing volumes of bacterial culture, ranging from 0.1 to 25 ml. As may be seen in Figure 5.5, lanes containing N protein and control bacterial lysates shared several features. In each group, there was an increasing shift of radioactivity from the gel bottom to the upper gel as culture volume increases. This represented the reduction in free RNA and accumulation of RNA complexed with protein. In each sample there were three such bands visible in the upper region of the gel. Migration of each of these bands was progressively retarded as culture volume increased. This was caused by additional proteins associating with the established RNA-protein complex, shifting the band step-wise up the gel. The presence of the bands in both groups identified the responsible proteins as being bacterial in origin. In those lanes with N protein there was greater radioactivity retained within the wells, potentially the result of N protein-RNA complexes too large to allow migration into the gel. However, this hypothesis is questionable due to the radioactivity trapped in the well of the lane with probe alone. There was a very faint band visible just below each well in N protein lanes 3 to 5. These bands could potentially represent RNA associated with N protein of a size able to pass into the gel. These bands were absent from the negative control. Additional controls used were Bsc cell lysate as a positive control and purified His-β–galactosidase as a negative control. The shift observed with the Bsc cell lysate was routinely obtained but the cell protein involved is unknown. No bands are visible in the lane with purified His-tag β–galactosidase.

The assay was performed using purified His-tag N protein in place of bacterial lysates. 10 000 cpm radiolabelled RNA was combined with samples of N protein ranging from 10 ng to 10 µg (Figure 5.6). There were no bands visible in lanes containing the N protein. However, there was a reduction in radioactivity at the gel bottom (free probe)
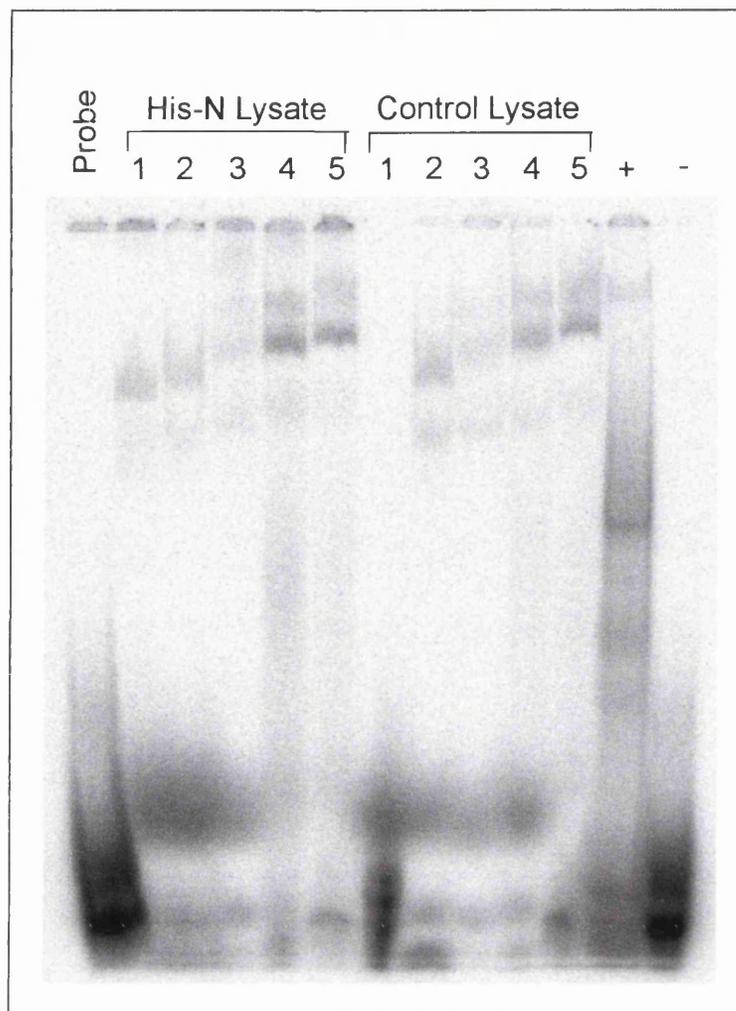
Figure 5.5: A gel mobility shift assay with bacterial cell lysates and specific radiolabelled RNA.

Lanes 1 to 5 represent volumes of bacterial culture, ranging from 0.1 ml to 25 ml. Volumes were centrifuged, the supernatants discarded and the bacterial pellets each resuspended in 500 µl PBS. Bacteria were lysed by freeze / thaw, chromosomal DNA shredded by passing through a needle and syringe and the lysate clarified by ultra-centrifugation. Bacterial RNA was removed from the clarified lysate by digestion with 75 units micrococcal nuclease in the presence of 1 mM $CaCl_2$. Ca^{2+} ions were chelated and hence enzyme activity halted upon addition of EGTA. 38 µl of each sample was incubated with 10 000 cpm radiolabelled RNA for 15 minutes at $33^{\circ}C$. Samples were separated in a 5% polyacrylamide TBE gel. The gel was dried and radiolabelled bands visualised by phosphorimagery.

Bsc cell lysate provided a positive and His-tag β-galactosidase, a negative control.
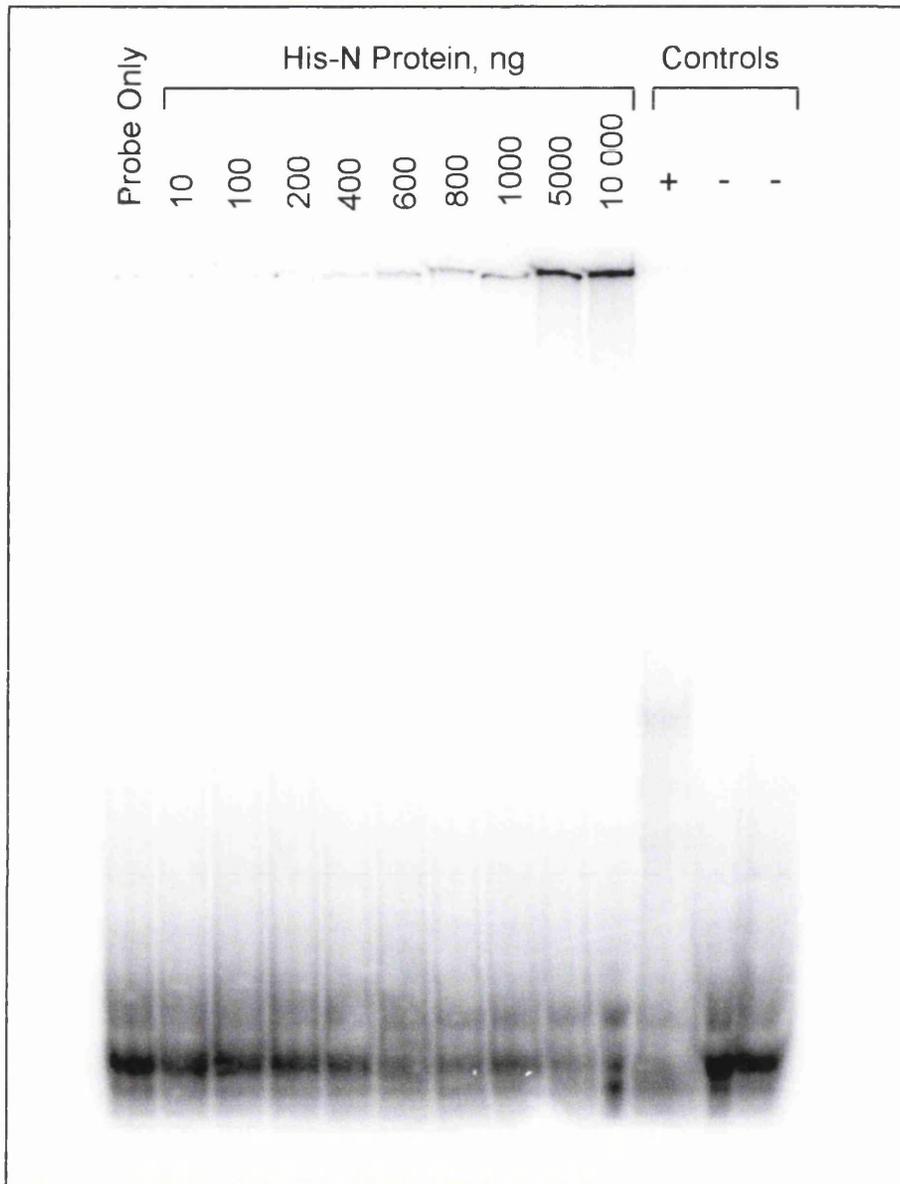
Figure 5.6: A gel mobility shift assay with purified His-tag N protein and specific radiolabelled RNA.

10 ng to 10 000 ng purified His-N protein was incubated with 10 000 cpm radiolabelled specific RNA for 15 minutes at 33°C. Samples were separated on a 5% polyacrylamide TBE gel. The gel was dried and radioactive bands visualised by phosphorimagery.

Bsc cell lysate provides a positive control and His-tag β-galactosidase and imidazole provide negative controls.

as protein concentration increased. In contrast, there was a gradual increase of radioactivity in the wells. This movement of radiolabelled RNA from the gel bottom to the wells was affected by N protein concentration. The N protein-RNA complexes were apparently too large to pass into the gel. This may be the result of the primary N protein-RNA complexes being bound and bridged by additional N proteins, either by protein-RNA or protein-protein associations, creating large aggregates. This phenomenon of protein-RNA retention in the wells has also been observed with HCV core protein and with the hantavirus nucleocapsid (J. Wood personal communication; Severson et al., 1999).

The positive and negative controls described in Figure 5.5 were duplicated and behaved as expected. Imidazole was included as an additional negative control due to trace amounts being present in the purified protein. No bands were present in this control.

The assay was repeated several times in the attempt to pass radiolabelled RNA-protein aggregates from the wells into the gel. Various gel compositions were investigated. The acrylamide percentage was varied between 3.5 and 5% and agarose was used in place of acrylamide. Despite altered matrix conditions, the radioactive complexes would not migrate into the gel. Therefore, the question of RNA sequence specificity by the N protein was not addressed using this assay. The gel mobility shift assay again confirms the ability of recombinant N protein to bind RNA in vitro.

5.5 North-Western Blot Assay

North-western blot analysis provided an alternative approach in the investigation of RNA-binding specificity by His-tag N protein (Section 2.11.4). In brief, the protein sample was separated by SDS-PAGE and transferred to nitrocellulose membrane. The membrane-bound protein was renatured by incubation in HBB buffer and then probed with radiolabelled RNA. The membrane was washed and bound radioactivity visualised by phosphorimagery.

The assay was performed on bacterial lysates of BL21 (DE3) pLysS expressing His-N (Figure 5.7). There were several bands common to both His-N protein and control samples, as observed in the gel mobility shift assay (Figure 5.5). Lanes containing N
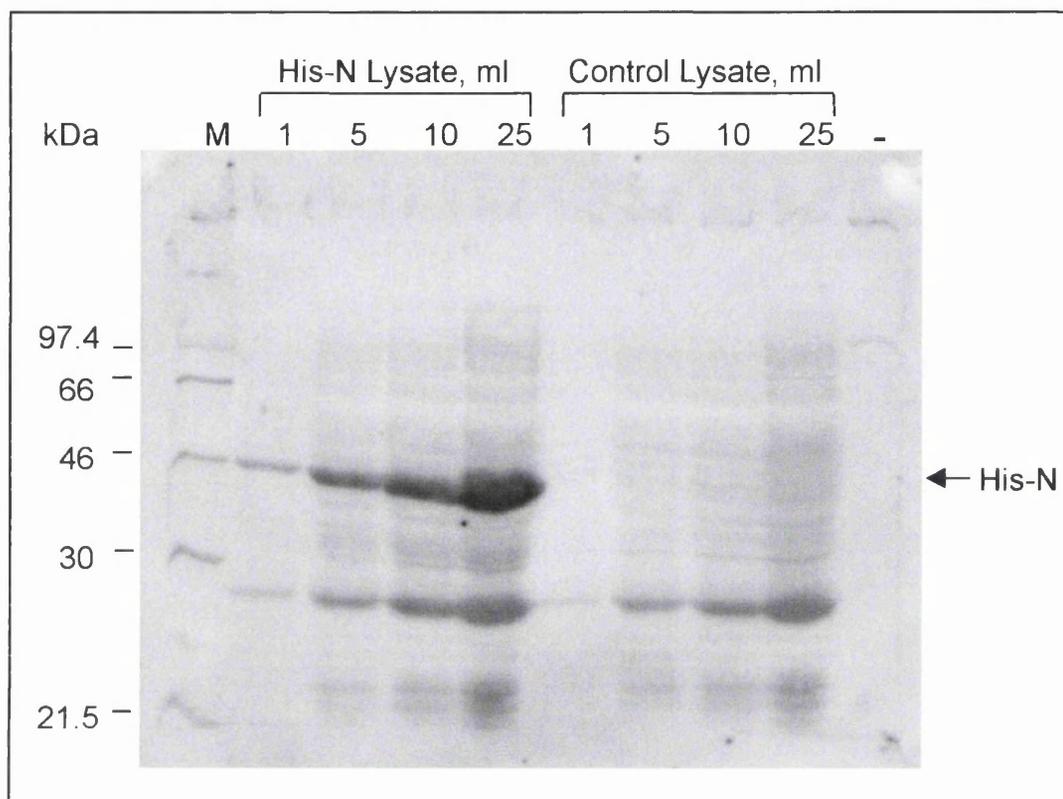
Figure 5.7: A north-western blot with bacterial cell lysates and specific radiolabelled RNA.

Volumes of 1 ml to 25 ml bacterial culture were centrifuged, the supernatants discarded and the bacterial pellets each resuspended in 500 µl PBS. Bacteria were lysed upon addition of sample buffer and boiling. 30 µl of each bacterial lysate was separated by 10% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane, blocked with PBS containing 5% Marvel and 1 mM DTT and probed with 50 000 cpm ml^{-1} radiolabelled specific RNA. The membrane was washed with HYB100 buffer and radiolabelled bands visualised by phosphorimagery.

Purified His-tag β-galactosidase provided a negative control and may be seen as a very faint band at 116 kDa. β-galactosidase does not bind RNA. The band may result from either weak binding of RNA to the histidine-tag or from the assay being inherently sticky. In this instance, the protein bands of His-N are more intense than other protein bands, perhaps indicating the occurrence of a specific RNA interaction.

protein lysates also possessed a band of 46 kDa, the expected size of His-tag N. The intensity of these bands increased with increasing culture volume. Bands of this size were absent from the control samples.

Following the success of this preliminary experiment, north-western analysis was used to investigate the ability of purified His-N to bind RNA. Varying amounts of purified protein were separated by SDS-PAGE and transferred to nitrocellulose membrane. The membranes were probed with 50 000 cpm ml^{-1} of either specific or non-specific radiolabelled RNA. In Figure 5.8 (a), His-N protein was present in a range from 10 ng to 10 µg and was probed with specific RNA. In Figure 5.8 (b), His-N was present in a range of between 100 ng and 2.4 µg and was probed with the non-specific RNA transcript. In each blot, bands of 46 kDa in size were visible, representing His-N bound by radiolabelled RNA. In both (a) and (b), band intensity increased with the increase in protein concentration. The presence of such bands in each blot indicated that His-N was able to bind both specific and non-specific RNA. This confirmed the observation from the yeast three-hybrid experiments that N protein bound RNA in a non-sequence specific manner.

A north-western blot was performed with a titration of His-tag β-galactosidase and specific radiolabelled RNA to confirm protein suitability as a negative control and to determine the inherent stickiness of the assay (Figure 5.8 c). No bands of radioactivity were visible in lanes with His-β-galactosidase. BSA provided an additional negative control and may be seen as a very faint band at 66 kDa. Upon comparison with (a) and (b), the blot in (c) demonstrated that only a low level of background binding occurred in the north-western assay. This reinforces the specificity of the interaction occurring between His-N and RNA.

Figure 5.9 confirms the non-sequence specificity of N protein-RNA association. The concentration of purified protein present in the strips of each blot was uniform. The strips were exposed to varying amounts of either specific (a) or non-specific (b) radiolabelled RNA. As radioactivity was increased from 80 to 50 000 cpm, a step-wise increase in band intensity was observed in both blots. The kinetics of RNA capture appear similar for both RNA species suggesting that the N has a similar affinity for both i.e. N protein lacks sequence specificity. However, this experiment serves only as an

Figure 5.8: A north-western blot of purified Histidine-tag N protein with specific (a) and non-specific (b) radiolabelled RNA.

0.01 µg to 10 µg purified His-N protein was separated by 10% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane, blocked with PBS containing 5% Marvel and 1 mM DTT and probed with specific (a) and non-specific (b) radiolabelled RNA. Membranes were washed with HYB100 buffer and radioactivity visualised by phosphorimagery.

A blot with purified His-tag β-galactosidase (His-β-Gal) and specific radiolabelled RNA provided a negative control (c). Bsc cell lysate was used as positive control (a) and His-β-gal (a), imidazole (a) and BSA (b & c) were used as negative controls. In (c), bands were absent from His-β-Gal lanes, however, a faint band of 66 kDa appeared in the BSA control. The band may be visible compared with His-β-Gal due to the higher protein concentration.

(a) His-N Protein, µg

(b) His-N Protein, µg

(c) His-β-Gal, µg
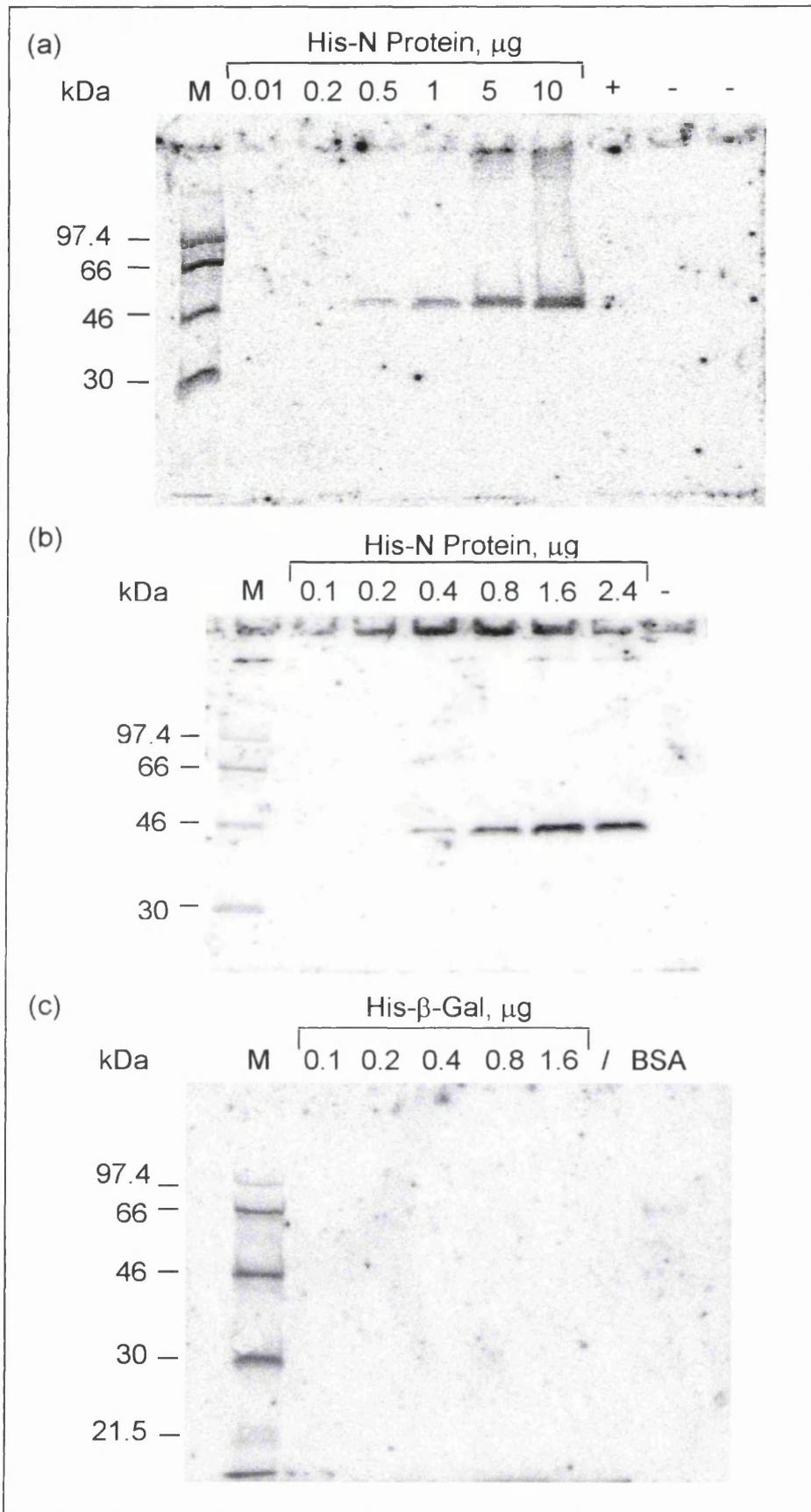
Figure 5.9: A north-western blot with Histidine-tag N protein and varied amounts of radiolabelled specific (a) and non-specific (b) RNA.

Purified His-N protein was separated by 10% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane and blocked with PBS containing 5% Marvel and 1 mM DTT. The membrane was then vertically sliced resulting in 3 µg His-N protein per strip. Each strip was probed with 80 to 50 000 cpm specific (a) or non-specific (b) radiolabelled RNA. Membrane strips were washed with HYB100 buffer and radioactivity visualised by phosphorimagery. His-N protein may be observed in each strip at approximately 46 kDa. A second band of approximately 30 kDa in size is present in strips with 50 000 cpm radiolabelled RNA. This may represent truncated His-N previously observed during protein expression and purification.

His-tag β-galactosidase with 50 000 cpm radiolabelled RNA provided a negative control. In the control lane probed with specific RNA, a very faint band is detected.
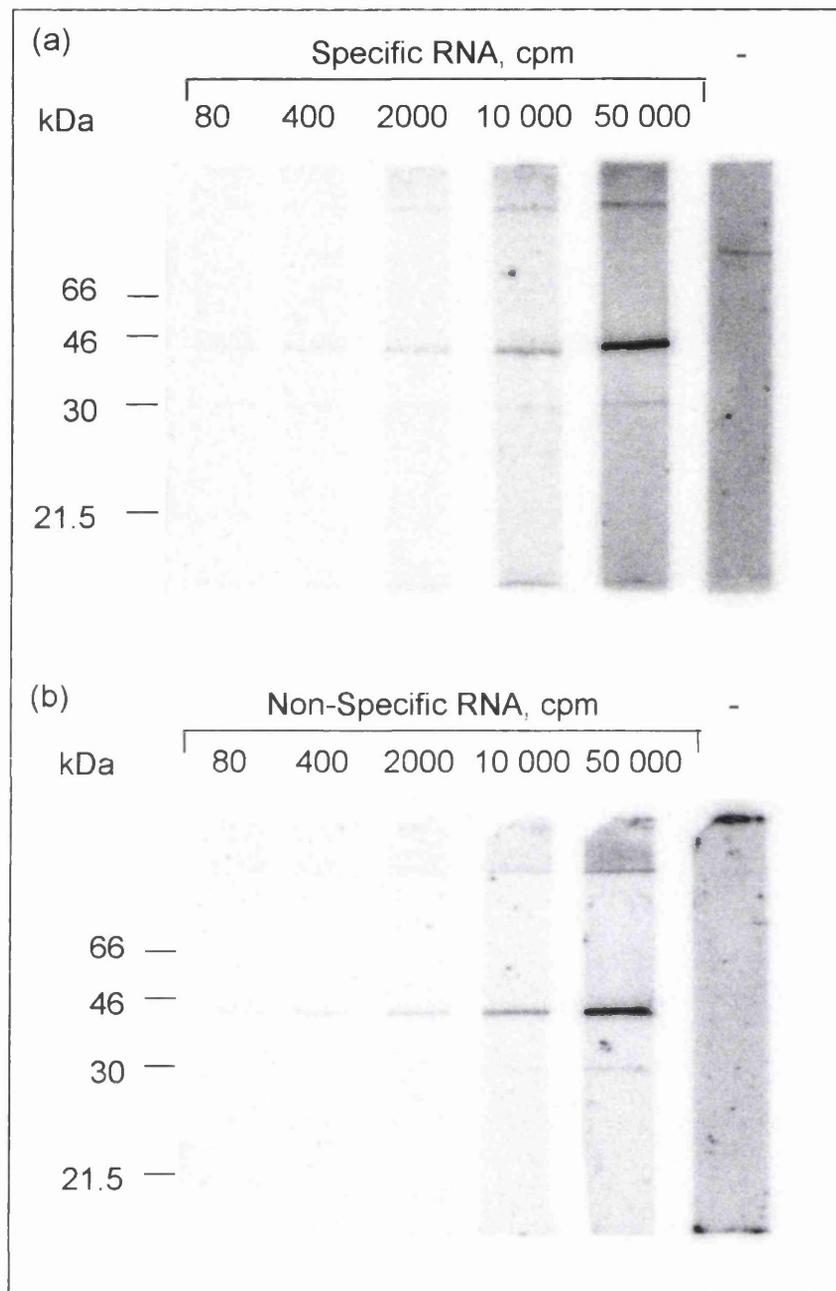
indication as it considers RNA radioactivity and not molarity. The specific activity of each probe was not calculated for this experiment.

5.5.1 North-Western Competition Assay

The north-western competition assay compares the molar ratios at which specific and non-specific radiolabelled RNA are removed from membrane-bound N protein by a different non-specific RNA cold competitor (Section 2.11.5). To perform this analysis, a modified north-western protocol was devised. Such studies are normally performed in solution and analysed by GMSA. However, as previously observed, this assay may be unsuitable due to protein aggregation.

In the modified north-western assay, individual pieces of nitrocellulose with His-N protein were incubated in a solution containing radiolabelled RNA with cold (un-labelled) competitor. Different ratios of competing species were used, and the bound radioactivity determined by phosphorimaging and densitometry. The north-western competition assay compares the molar ratios at which specific and non-specific radiolabelled RNA are removed from membrane-bound N protein by the cold RNA competitor (Figure 5.10). Radiolabelled RNA molarity was the same for each piece of membrane. The molar ratio of competitor RNA was increased from 0 to 729. The radiolabelled band intensity decreased as the concentration of cold competitor RNA increases, a feature common to both specific and non-specific radiolabelled RNA. The faint band obtained upon probing His-tag β-galactosidase with specific RNA disappeared at a low ratio of cold RNA competitor; the band was no longer detectable by the ratio 1:27 compared with a ratio of 1:729 for His-N.

N protein band intensities differ greatly with specific and non-specific radiolabelled RNA. This is the result of differences in probe specific activity. ^{32}P-labelled CTP was used in the synthesis of each probe. Nine C residues are present in specific RNA compared with 11 C residues in non-specific RNA. Therefore, non-specific RNA potentially has a higher specific activity than specific RNA, which accounts for the difference in band intensity observed in the assay.

The purpose of the competition assay was to investigate whether or not N protein possessed RNA sequence preference. To address this question, densitometric analysis using Quantity One software (BioRad) was performed on each band in the
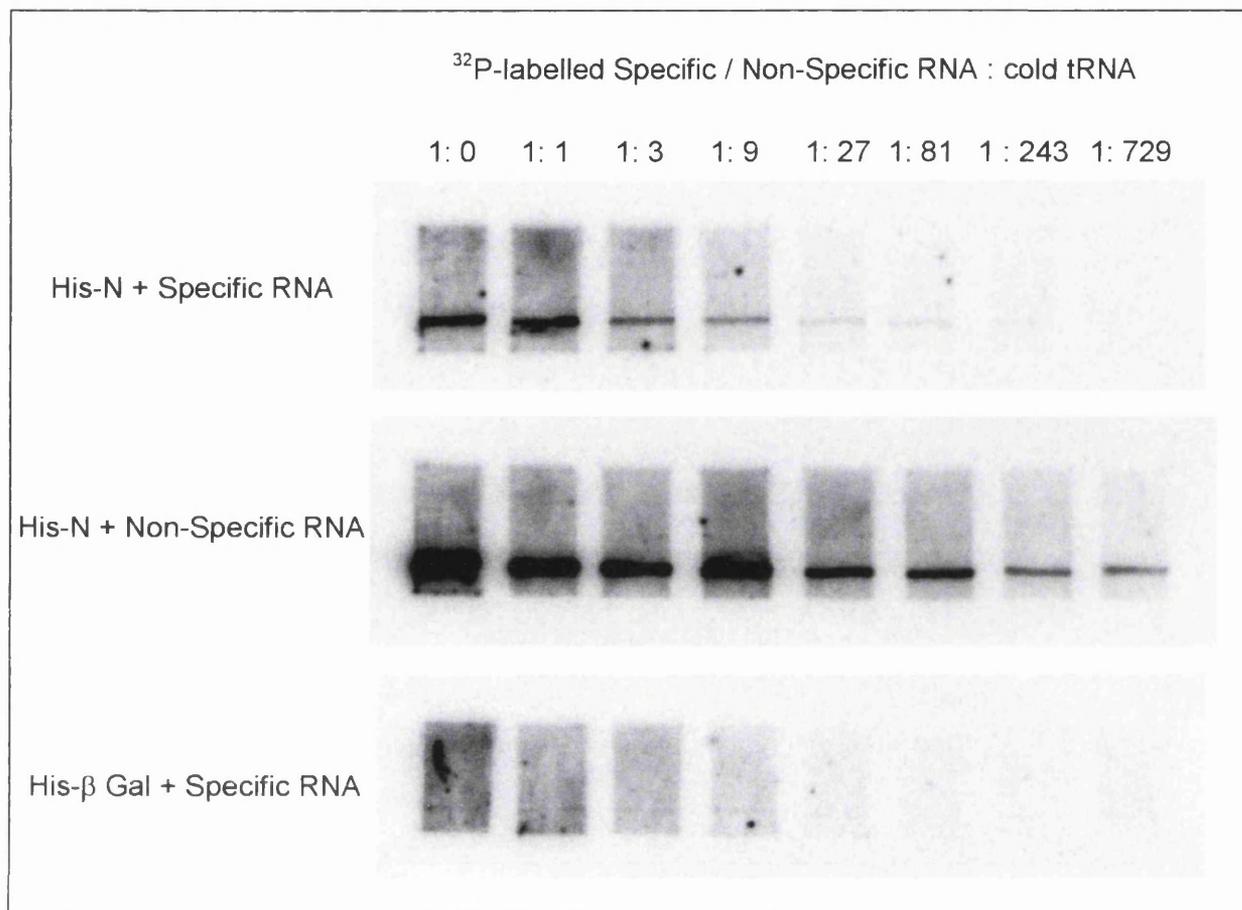
Figure 5.10: A north-western blot competition assay with Histidine-tag N protein and specific and non-specific radiolabelled RNA.

Purified His-N protein was separated by 10% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane and blocked with PBS containing 5% Marvel and 1 mM DTT. The membrane was then sliced into strips and the protein band isolated by alignment with Rainbow marker resulting in 2 μg His-N protein per membrane piece. Each membrane slice was probed with specific or non-specific radiolabelled RNA in the presence of increasing molar ratios of cold competitor tRNA. Membrane slices were washed with HYB100 buffer and radioactivity visualised by phosphorimagery.

His-tag β-galactosidase with specific radiolabelled RNA provided a negative control.

resulting blots. Both actual measurements and percentages are listed in Table 5.1 and Figure 5.11. Analysis showed that the competition between specific or non-specific RNA and the cold competitor followed similar patterns. The similarity in rates at which specific and non-specific radiolabelled probe were competed from His-N by cold RNA suggested that the strength of association for each RNA was similar. The decrease in binding is rapid initially but a small amount of radiolabelled RNA appears to have resisted being dislodged by the cold competitor. It is possible that the majority of the association was not due to RNA binding to a particular site on His-N but rather that the acidic RNA bound basic residues on the protein surface. RNA bound in this manner may have been easily dislodged by the competitor. However, a small amount was captured by the genuine RNA-binding site on N, the binding to which may be irreversible, at least *in vitro*.

5.6 Discussion

The association between purified N protein and RNA was investigated *in vitro* by a number of assays. The approaches selected to address the question of RNA sequence specificity of the N protein had varying success. However, collectively, the assays demonstrated that recombinant histidine-tagged N protein possessed the ability to bind RNA *in vitro*.

Filter-binding and gel mobility shift experiments demonstrated the association between radiolabelled RNA and N protein. Frequently, the former assay was not greatly successful with little radioactivity captured on the membrane. Retention of protein was confirmed by western blot. Therefore, it was concluded that the RNA-protein complex frequently dissociated under assay conditions. The gel mobility shift assay similarly demonstrated association between N and RNA. In Figure 5.6, the protein-RNA complex was retained within the wells of the gel, an inescapable feature of this assay with the N protein. Such retention arises from protein-protein and protein-RNA interactions creating aggregates too large to migrate into the gel. From subsequent experiments, it may also be considered that bacterial RNA, remaining from recombinant protein expression, exacerbates this situation. Similarly, the presence of contaminating bacterial RNA could explain poor protein-radiolabelled RNA association in the filter-binding assays. Protein-RNA interactions observed by these assays were confirmed as true by inclusion of histidine-tag β-galactosidase as negative control. The

Molar Ratio of Radiolabelled RNA to Cold Competitor	Band Intensity		% of Total Radioactivity	
	Specific RNA	Non-specific RNA	Specific RNA	Non-specific RNA
1:0	782	4184	100	100
1:1	675	2364	86.32	56.5
1:3	253	1403	32.35	33.53
1:9	149.7	1574	19.14	37.62
1:27	43.45	703.7	5.56	16.82
1:81	32.42	568.4	4.15	13.59
1:243	24.65	226.7	3.15	5.42
1:729	19.14	160.7	2.45	3.84

Table 5.1: A table of band intensities and percentages of radioactivity obtained from a north-western competition assay of His-N protein with specific and non-specific RNA sequences.

Densitometric analysis of the bands obtained in the north-western competition assay in Figure 5.10 was performed using Quantity One software (BioRad). Band intensities, normalised to background, are listed above.

With each radiolabelled probe, the band created in the absence of cold competitor (i.e. ratio 1:0) was considered 100%. Measurements of subsequent bands were calculated as a percentage of this intensity.

Figure 5.11: Graphs of percentages of total radioactivity obtained from the north-western competition assay.

The graphs represent north-western blot competition assay data presented in Figure 5.10 and Table 5.1. They depict the percentage of total radioactivity of each His-N band upon increasing molar ratio of cold competitor RNA.

All competitor ratios are present in (a). Competitor ratios 1 to 243 only are present in (b). Competitor ratios represent the molar excess of cold competitor with respect to radiolabelled RNA.
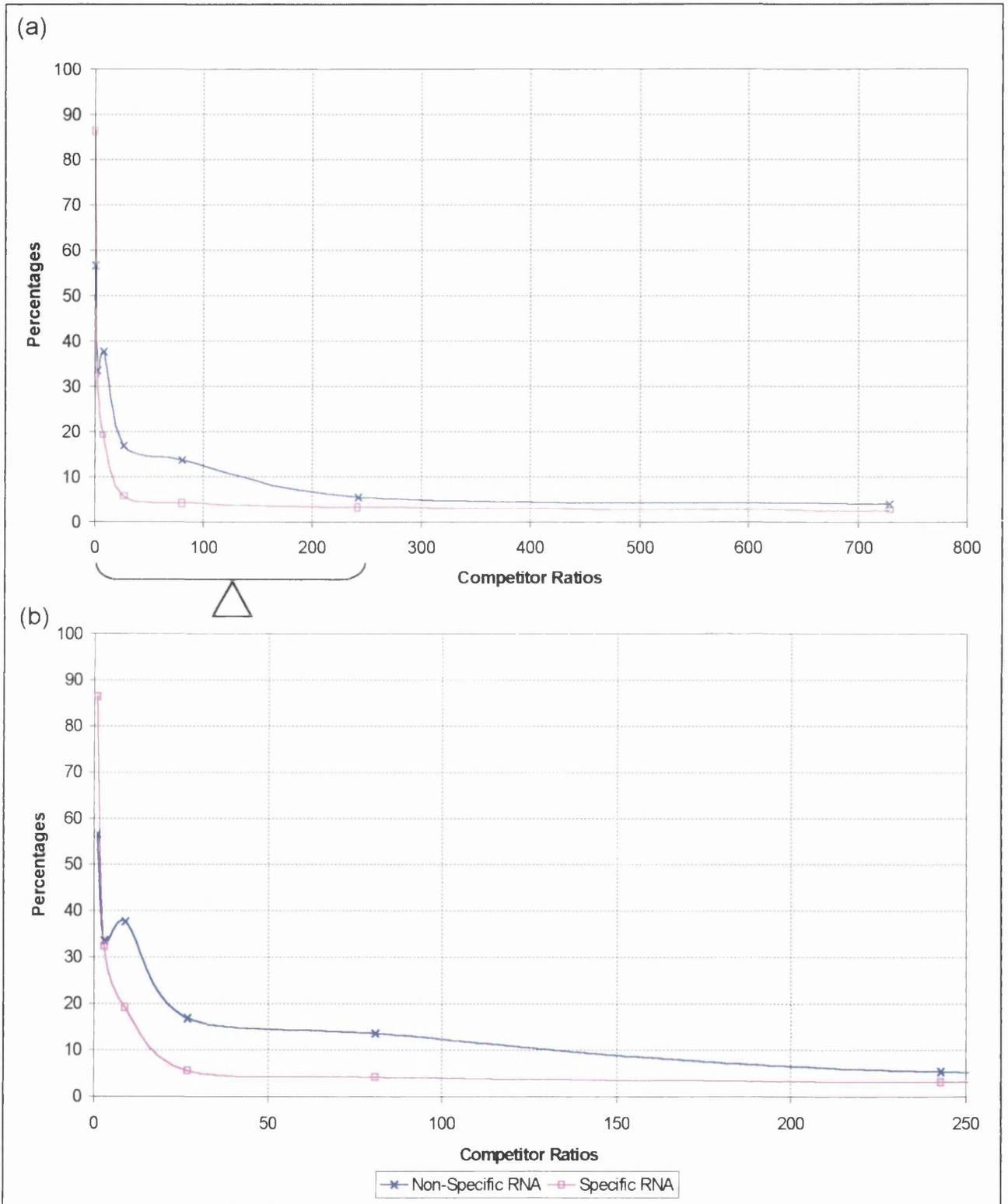
protein confirmed the specificity of the His-N-RNA interaction and the absence of contribution by the His-tag.

The ability of His-N protein to bind sequences other than the RSV leader was demonstrated by the yeast three-hybrid system and north-western blot analysis. The yeast system met with varied success. Though the controls were not absolute, the assay strongly indicated that the N protein of RSV strain A2 was able to bind with a non-specific RNA, in this instance, the 3' non-coding region of the HCV genome. The N protein of RSV strain RSS-2 failed to bind RNA. Previously, expression and biological activity of this clone had been demonstrated by its ability to bind P in the yeast two-hybrid system (C. Loney, personal communication). Therefore, the lack of success of RSS-2 N in the yeast three-hybrid system is probably due to reduced RNA-binding ability of the clone.

North-western blot analysis provided the most convincing evidence for the ability of His-N protein to bind both specific and non-specific RNA. In this procedure, protein was denatured upon separation by SDS-PAGE. Therefore, any bacterial RNA remaining from recombinant protein expression was dissociated from His-N. Following transfer to nitrocellulose membrane, protein was renatured by incubation in HBB buffer. Therefore, unlike several of the *in vitro* assays so far employed, north-western blots investigated the renatured protein, free from contaminating RNA.

During these studies, a protein band of approximately 30 kDa was identified as binding RNA (Figure 5.9; strips probed with 50 000 cpm). This band was the same size as that observed during the purification of His-N in Chapter 4. The 30 kDa band was identified as representing the amino-terminus of the N protein. Its detection in this assay suggests that the RNA-binding domain may be present within the amino terminus of the N protein (see Chapter 6).

Modification of north-western blot analysis led to development of the competition assay. This assay enabled semi-quantitative analysis of the preference of RNA binding by N. Results indicated that the N protein does not posses RNA sequence preference, confirming observations made from the yeast three-hybrid system.

From the studies presented in this chapter, it was concluded that recombinant His-tag N protein was able to bind RNA *in vitro*. In addition, the protein possessed the ability to bind RNA sequence of both viral and non-viral origin and displays an apparent lack of sequence preference. Also various conditions were determined that may be of use in future studies on the N-RNA interaction. A novel assay to overcome the inherent difficulties of working with N proteins was devised and successfully employed.

6   A Study of the RNA-Binding Domain of N Protein (I)

6.1   Introduction

The purpose of this study was to identify regions in N protein responsible for interacting with RNA. The first approach selected to identify this RNA-binding domain, or domains, required the creation of a series of deletion mutants of N protein. The ability of mutant protein to bind RNA was investigated by the use of established assays such as north-western blotting. If interesting domains were identified, the mutant proteins would be further refined, with the ultimate aim of focussing on individual amino acids.

6.2   Mutant N Protein Synthesis

6.2.1   Mutant His-N Generation by Site-Directed Mutagenesis

Two histidine-tag N protein mutants with internal deletions were created using the site-directed GeneEditor mutagenesis kit (Promega) using pET16b N as template. The deletion mutants NΔ1-200 and NΔ121-160 were generated by this method. Mutant DNA excised from plasmid pET16b by digestion with *Bam* HI is presented in Figure 6.1. The mutant sizes of NΔ1-200 and NΔ121-160 were of the expected lengths (620 bp and 1100 bp, respectively). The deletions, and maintenance of the ORF, were confirmed by manual and automated DNA sequencing.

Attempts were made to expand the repertoire of mutants generated by site-directed mutagenesis however cloning difficulties prevented extension of this mutant range.

6.2.2   Mutant His-N generation by PCR

N protein deletion mutants were generated by PCR to compliment those created by site-directed mutagenesis. Internal mutation primers were designed to introduce a premature stop codon and restriction site into the N protein gene creating a series of carboxy terminus deletions. Truncations were made in increments of 50 amino acids.

PCR was performed with pET16b N template DNA and mutation primers following the method described in Section 2.9.2. DNA amplification was confirmed by ethidium bromide-stained agarose gel electrophoresis (data not included). Following limited exposure to long wave UV light, the DNA bands were located, excised and extracted
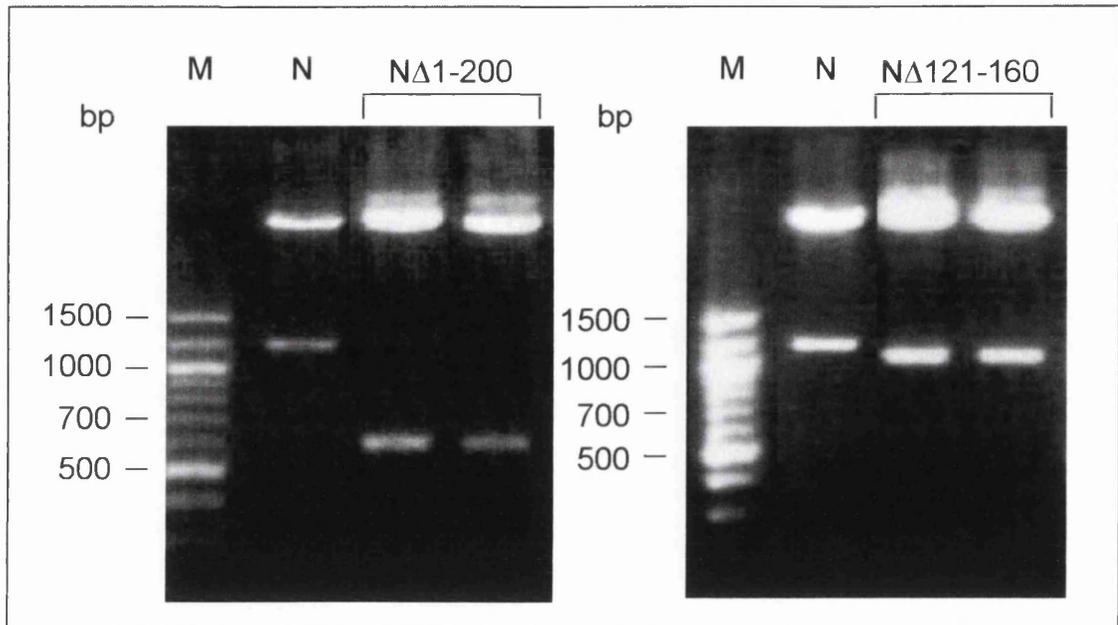
Figure 6.1: *Bam* **HI digestion of N mutants with internal deletions generated by Site-Directed GeneEditor mutagenesis.**

Samples were separated in a 1% agarose TAE gel, stained with ethidium bromide and visualised under UV light. Full-length N is 1220 bp in size. Mutants NΔ1-200 and NΔ121-160 are approximately 620 bp and 1100 bp, respectively. Linearised pET16b (5711 bp) is present in each lane. (Marker, M).

from agarose gel by Gene Clean. DNA was digested with *Bam* HI to enable subsequent ligation into pET16b. The products of digestion were visualised by ethidium bromide-stained agarose gel electrophoresis (Figure 6.2 a). A single discreet band was present within each sample lane. Each band was of the expected size for the particular mutant, ranging from 1070 bp for NΔ342-391 to 170 bp for NΔ42-391.

Each PCR product was combined with *Bam* HI-digested plasmid pET16b in a ligation reaction. The ligation products were ethanol precipitated followed by introduction into DH5 cells by electroporation. Transformed bacteria were selected on L-broth agar. Individual colonies were picked and cultured in selective liquid growth medium. Amplified plasmid DNA was prepared and insert size and orientation determined by restriction digestion patterns in ethidium bromide-stained agarose gel electrophoresis. The digestion products of NΔ342-391 to NΔ142-391 may be seen in Figure 6.2 (b). *Bam* HI DNA fragments excised from plasmid were of the size expected for each particular deletion, ranging from 1070 bp for NΔ342-391 to 470 bp for NΔ142-391. *Hind* III was used to digest pET16b NΔ242-391 to determine insert orientation. An excised band of 702 bp was produced demonstrating the correct orientation of insert in plasmid. All other plasmids were digested with *Xba* I with which a band of 509 bp was expected if the mutant gene was present in the correct orientation. Bands of similar size were excised from NΔ292-391, at least one clone of NΔ192-391 and from NΔ142-391. Therefore in this example, mutant clones were present in the correct orientation in all but NΔ342-391. Upon re-ligation, plasmids carrying NΔ342-391 and NΔ92-391 in the correct orientation were obtained (data not included).

6.3 Mutant Protein Expression

Mutant DNA clones, generated by site-directed mutagenesis and by PCR, were introduced into *E.coli* strains BL21 (DE3) pLysS and NovaBlue (DE3) by electroporation. The eight mutants successfully created are depicted in Figure 6.3.

The temperature and length of incubation for optimal mutant protein expression in BL21 (DE3) pLysS were investigated. Culture temperatures of 30°C and 37°C were compared and the length of incubation following induction of protein expression was varied from three hours to overnight. It was determined that conditions previously identified as optimal for full-length N protein expression were suitable for expression of

Figure 6.2: Restriction endonuclease digestion of gene-cleaned mutant N PCR products (a) and of N PCR mutants in pET16b (b).

DNA samples were separated by 2% agarose TAE gel electrophoresis, stained with ethidium bromide and visualised by UV light. *Bam* HI was used to digest gene-cleaned mutant N PCR products (a). It was also employed in the plasmid-insert digestions visible in (b). In both (a) and (b), *Bam* HI digestion created bands of mutant DNA of the following approximate sizes: NΔ342-391 (1070 bp), NΔ292-391 (920 bp), NΔ242-391 (770 bp), NΔ192-391 (620 bp), NΔ142-391 (470 bp), NΔ92-391 (320 bp), NΔ42-391 (170 bp). In (b), orientation digestions were performed with *Xba* I and with *Hind* III (NΔ242-391). With the exception of NΔ342-391, each mutant was present in the correct orientation within pET16b. Linearised plasmid (5711 bp) is present in each sample lane. (Marker, M).

(a)

(b)
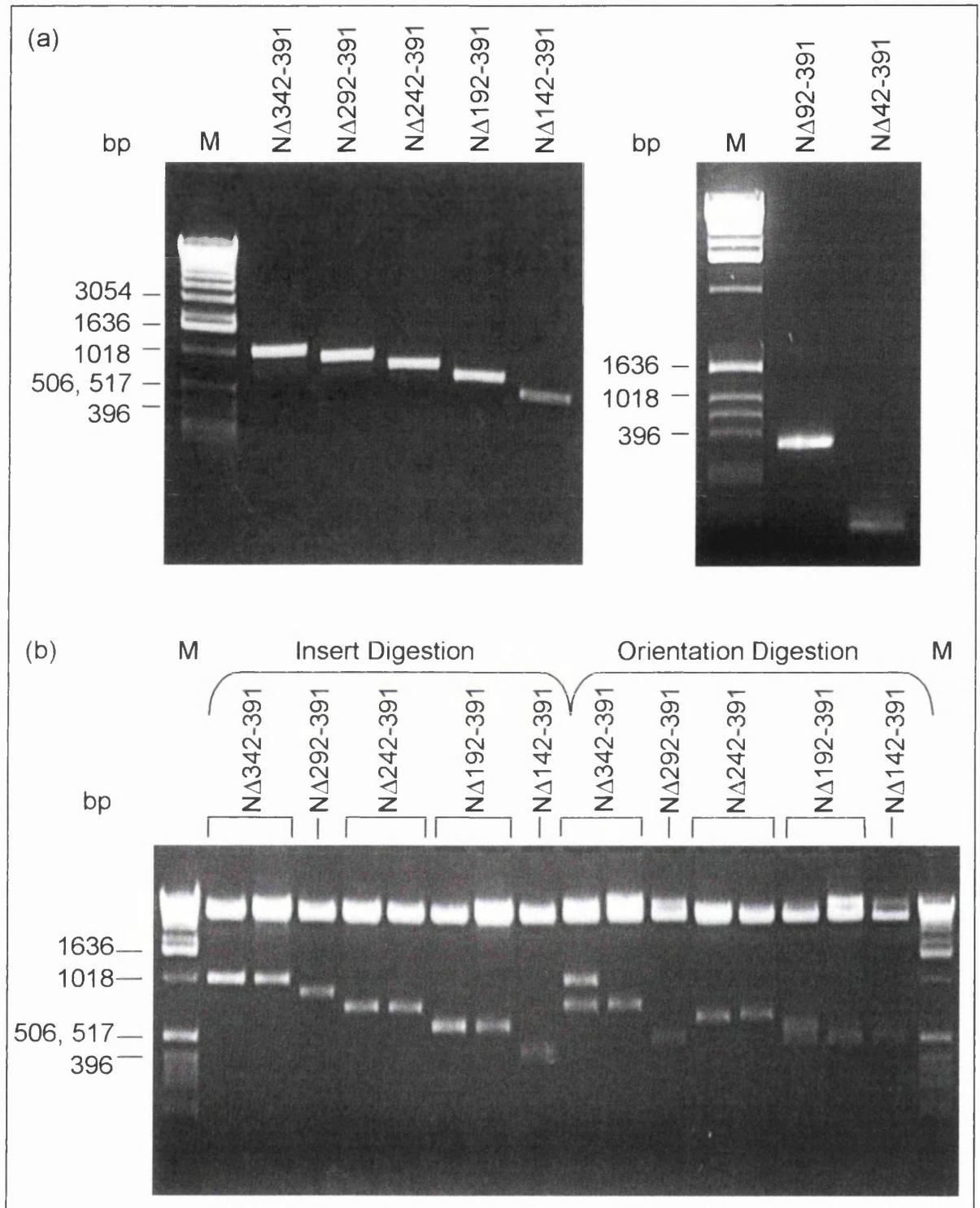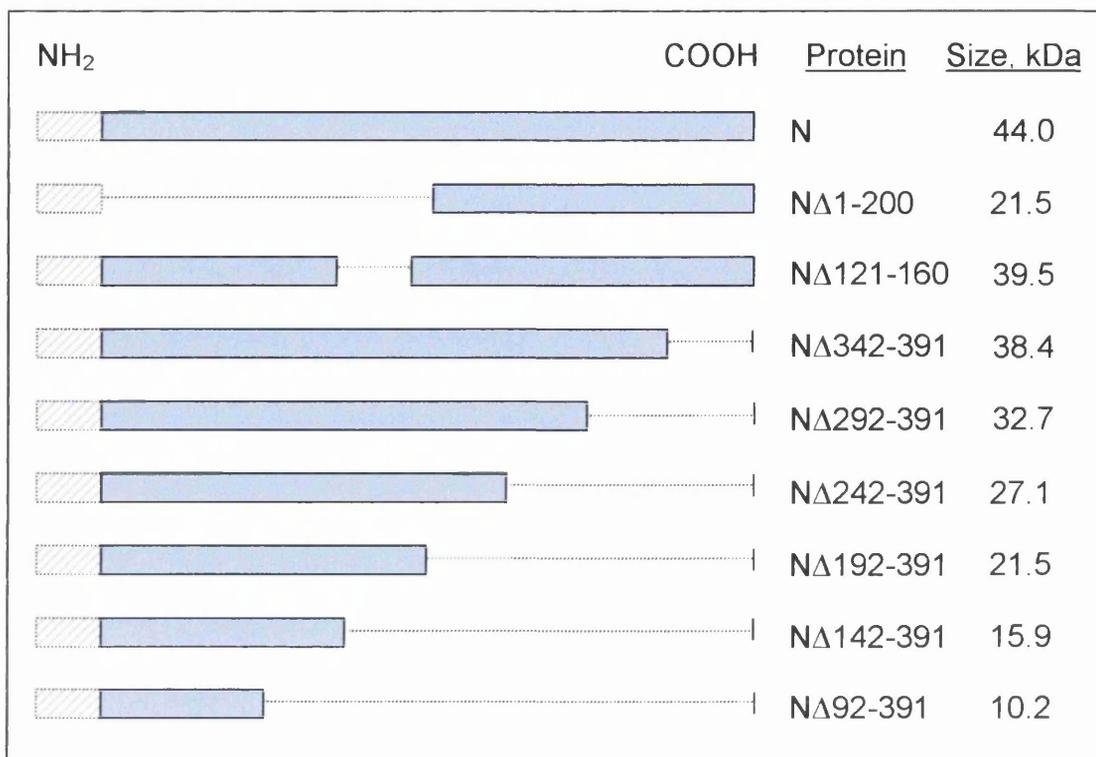
Figure 6.3: A cartoon of Histidine-tag N protein mutants.
Mutants are compared with full-length protein and their deletions indicated by dotted line in place of the coloured box. The amino terminus box surrounded by dotted line represents the histidine-tag. Expected sizes of mutant protein (kDa) are approximate and are calculated disregarding the His-tag.

mutant proteins. Therefore, bacterial cultures were incubated at 37°C until an O.D.$_{600}$ 0.8 was obtained. Recombinant protein expression was induced with 1 mM IPTG for a further three hours at 37°C.

Mutant protein expression from BL21 (DE3) pLysS and NovaBlue (DE3) was compared under these culture conditions. Expression was reasonable from NovaBlue (DE3) bacteria but was more uniform and at a higher level in BL21 (DE3) pLysS. Therefore, the latter was routinely used for mutant protein expression.

Bacterial expression of the histidine-tag N mutant proteins was confirmed by western blot (Figure 6.4). Purified His-N was used as a positive control and was visible as a band of approximately 46 kDa. A band of approximately 90 kDa was also present within this lane possibly representing His-N protein dimers. Mutant protein of the expected size was present in each lane (sizes are listed in Figure 6.3). There was considerable variation in the level of expression which resulted in some lanes being overloaded, particularly those with the amino-terminal deletion mutants NΔ1-200 and NΔ121-160. In addition to the full-length mutant proteins, breakdown products of approximately 21.5 kDa and 30 kDa were visible in lanes with NΔ121-160 and NΔ342-391, respectively. The additional band, present in NΔ121-160, was similar in size to NΔ1-200 and may result from sample overflow from the adjacent lane. Similarly, the extra band in NΔ342-391 was similar in size to NΔ292-391 and may represent contamination from the adjacent lane. Alternatively, it may be the amino-terminal truncation product previously observed in samples of full-length N protein, resulting from proteolytic degradation.

The mutant proteins were extremely prone to degradation. Storage conditions were varied in temperature (-70°C, -20°C, +4°C), by the addition of protease inhibitors and of stabilisers such as glycerol but with little effect. The proteins could be maintained in bacterial lysates or in purified state at 4°C for up to two weeks but use of freshly synthesised protein proved much more reliable.

6.4   RNA-binding properties of N protein mutants

The ability of His-N deletion mutants to bind RNA was investigated by north-western blot analysis. Initial experiments were performed with clarified bacterial lysates (Figure

Figure 6.4: A north-western (a) and associated western blot (b) of clarified lysates of bacteria expressing histidine-tag N mutants.

The bacterial lysates were separated by 12% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane, blocked with 5% Marvel in PBS and probed with either 50 000 cpm ml^{-1} specific radiolabelled RNA (a) or HRP-conjugated antibody directed against the Histidine-tag (b). The north-western blot (a) was washed with HYB100 buffer and the radiolabelled bands visualised by phosphorimagery. The western blot (b) was washed with TPBS and protein bands detected by ECL and autoradiography.

In (a), untransformed bacterial lysate (1) and purified His-β-galactosidase (2) provided negative controls. Purified His-N (~46 kDa) (1) and His-N bacterial lysate (2) were used as positive controls. In (b), untransformed bacteria provided a negative control and purified His-N and His-β-galactosidase (116 kDa) were used as positive controls.
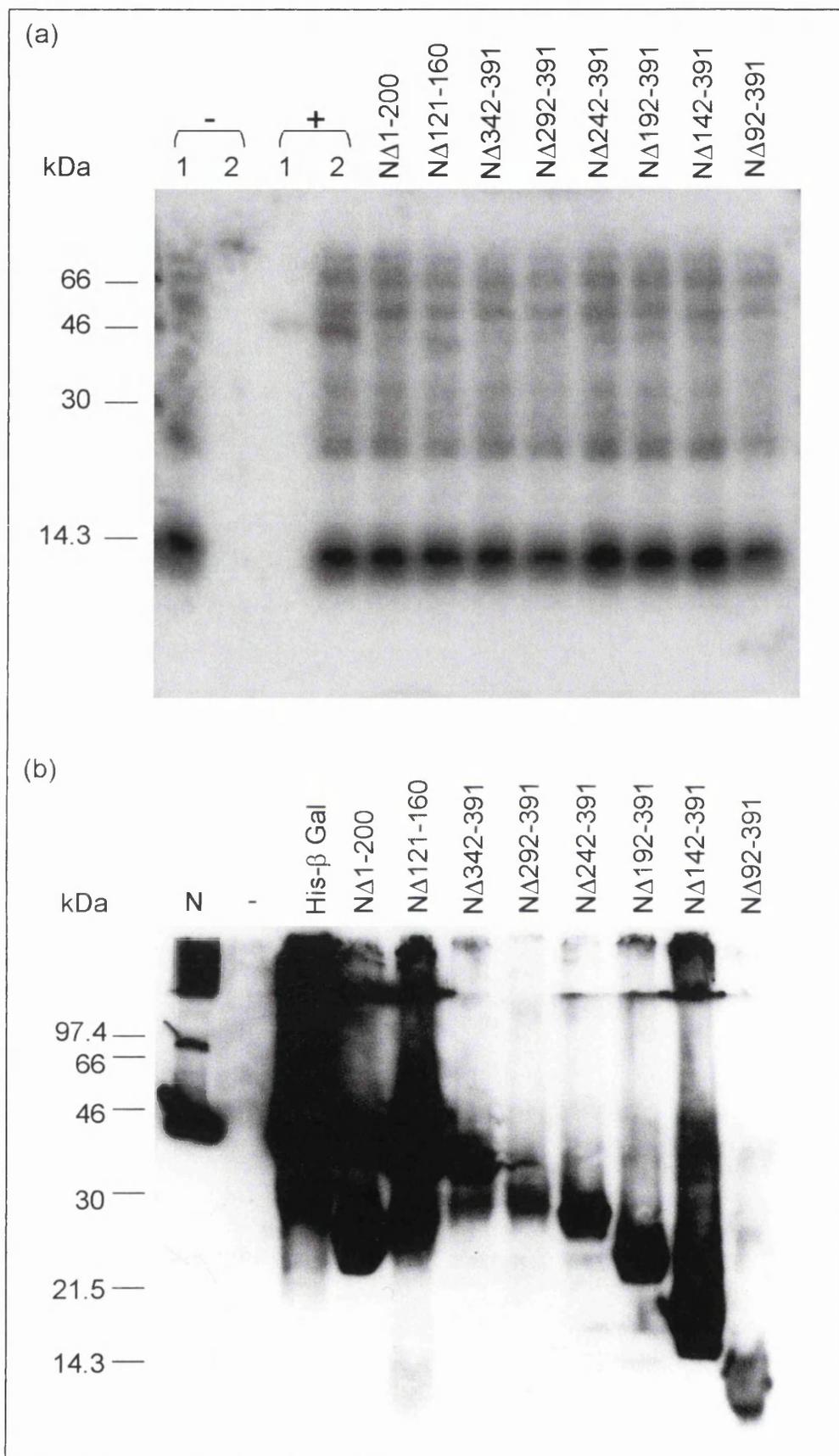
6.4). However, considerable background radioactivity was obtained by this method due to bacterial contaminants. Background radioactivity obscured the signal of potential bands created by the association of RNA with mutant N protein. However, possible bands were observed in some sample lanes providing encouragement for further investigation.

Mutant proteins were purified from bacterial lysate in an attempt to reduce background in the resulting north-western. The most successful method of His-tag mutant purification was by addition of nickel-NTA agarose to clarified bacterial lysates with incubation for 90 minutes at $4^{\circ}C$. The agarose was pelleted by centrifugation and washed successively with 10 mM, 20 mM, 40 mM, 60 mM and 100 mM imidazole phosphate buffer. The agarose was then combined with sample buffer to remove bound His-tag protein, centrifuged and the supernatant investigated by SDS-PAGE. Gels stained with coomassie blue demonstrated that bacterial contaminants were considerably reduced and western blot confirmed recovery of the His-tag mutant proteins (data not included).

North-western blot analysis was performed using purified proteins. However, considerable background radioactivity was still obtained. To counteract this, the number and duration of membrane washes was increased and an unlabelled competitor RNA was added.

A north-western assay was performed with purified mutant proteins and with radiolabelled specific RNA (i.e. leader RSV) and un-labelled non-specific competitor RNA at a molar ratio of 1: 5 (Figure 6.5 (a)). Recovery of mutant proteins following purification was confirmed by western blot (Figure 6.5 b). In Figure 6.5 b, the protein band intensities were quite varied ranging from the total absence of NΔ292-391 and low concentrations of NΔ342-391 to high concentrations of full-length N protein, the amino-terminus deletion mutants and NΔ142-391. In Figure 6.5 (a), the background radioactivity was improved by the addition of competitor RNA but was still high and the few protein-RNA bands observed were diffuse. Banding similar to that in the negative control (purified lysate of bacteria lacking His-tag protein) was present in each lane of purified protein. A strong band of approximately 46 kDa was visible with purified full-length N. There was a distinct band observed with NΔ121-160 and potential bands in lanes with NΔ1-200 and NΔ142-391 though these were of lower intensity. Each band

Figure 6.5: A north-western (a) and associated western blot (b) of purified mutant N proteins.

Purified His-N mutant proteins was separated by 12% polyacrylamide SDS-PAGE. Proteins were transferred to nitrocellulose membrane and probed with specific radiolabelled RNA and unlabelled bovine tRNA in a molar ratio of 1:5 (a) or with HRP-conjugated antibody directed against the histidine-tag (b). The north-western blot (a) was washed with HYB100 buffer and the radiolabelled bands visualised by phosphorimagery. The western blot (b) was washed with TPBS and protein bands detected by ECL and autoradiography.

In both (a) and (b), purified His-tag N protein provided a positive control. Purified lysate of bacteria lacking His-tag protein was used as negative control. In (b), faint bands at 46 kDa, >30 kDa and >14.3 kDa are present in the negative control. There are potential RNA-protein bands in the lanes His-N, NΔ121-160, NΔ1-200 and NΔ142-391 in the north-western blot (a).

(a)

(b)
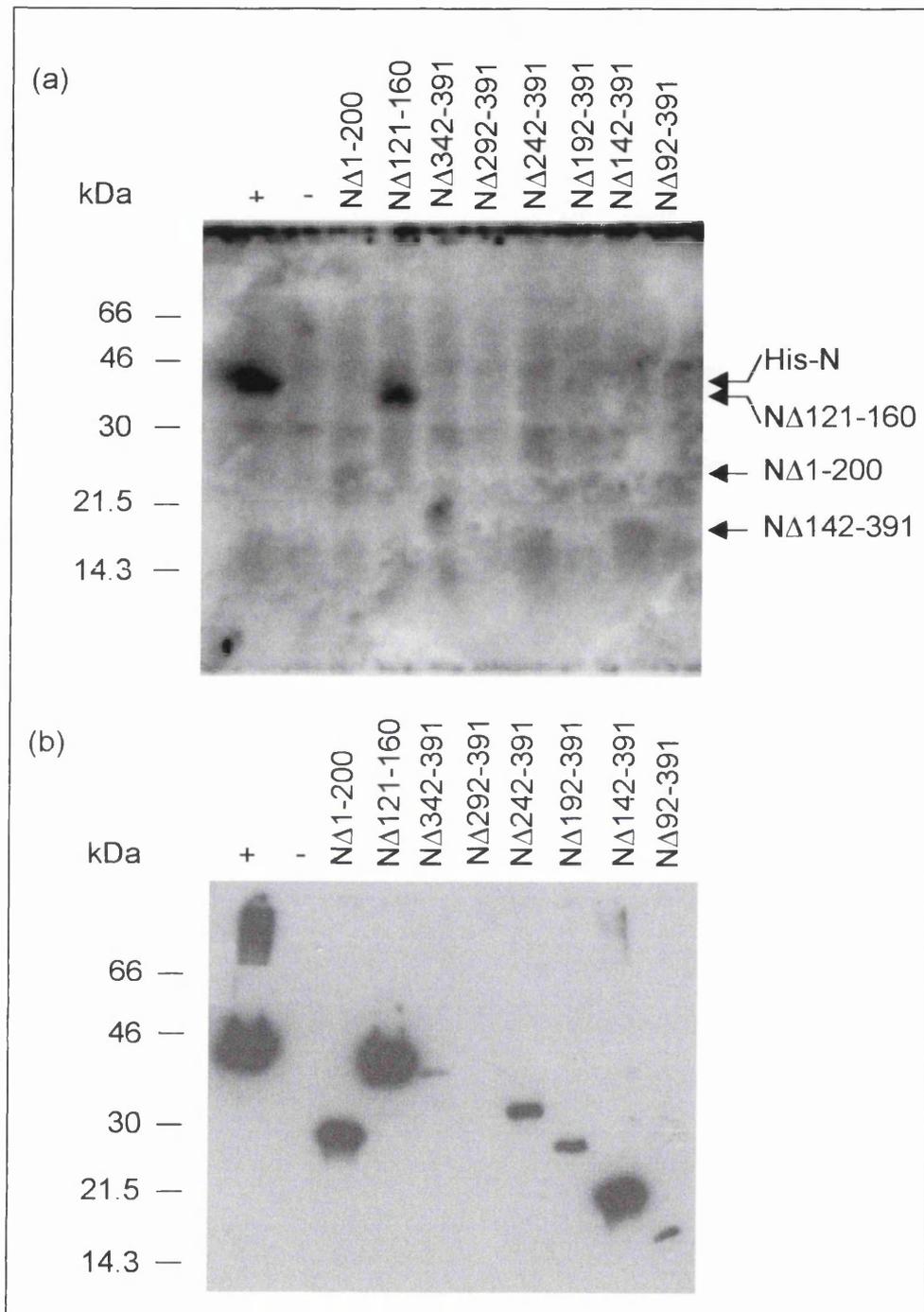
was of the expected size for the mutant protein. As demonstrated in the western blot, these proteins were present at high concentrations. The faint bands observed with NΔ1-200 and NΔ142-391 may be the result of non-specific association of RNA to high concentrations of protein, a feature previously observed in north-western blots. However, the intensity of the band present in NΔ121-160 suggests the occurrence of a specific protein-RNA interaction. No bands could be distinguished in the remaining lanes.

Despite numerous repeats of the assay, background radioactivity remained too high to reliably determine which mutants retained the ability to bind RNA. Further protein purification was attempted using HiTrap chromatography. This gained little advantage in the resulting north-western blot. Therefore, the assay was discontinued having only concluded that NΔ121-160 and full length N protein bind RNA.

6.5 Discussion

Having previously established the general nature of association between RNA and the N protein of RSV, the next stage of investigation was to identify domains within the protein responsible for such interaction. In this chapter, this aim was addressed by the generation of N protein mutants and analysis of their ability to bind radiolabelled RNA.

Eight histidine-tag N protein mutants were successfully created. A series of six carboxy-terminal deletion mutants, with truncations in increments of 50 amino acids, were generated by PCR. Two mutants with internal deletions in the amino terminus were created by site-directed mutagenesis using the GeneEditor kit (Promega). Further attempts to create more mutants by site-directed mutagenesis were unsuccessful. This was probably the result of putting too great a demand on the system by attempting to create internal deletions of too large a size. The size and orientation of mutants within the plasmid pET16b were confirmed by restriction enzyme digestion. Partial sequencing of each mutant sequence confirmed the nature of each mutation.

Conditions optimal for the expression of full-length His-N from BL21 (DE3) pLysS were found to be suitable for the expression of the mutant proteins. Protein production in NovaBlue was compared with that from BL21 (DE3) pLysS. Mutant expression was more consistent from the latter host. Unlike the full-length protein, mutants of His-N

were found to be very sensitive to degradation enabling only short-term storage. This lack of stability may result from the absence of specific amino acid sequences or of larger protein domains, thereby creating less stable protein conformations and causing proteins to fall apart with greater ease. In Figure 6.4 (b), N protein mutants in clarified bacterial lysates are visualised in a western blot. In addition to the mutant protein, lanes containing NΔ121-160 and NΔ342-391 posses smaller bands of 21.5 kDa and 30 kDa, respectively. Additional bands could represent the overflow of protein from adjacent wells that would create bands of the appropriate sizes. Alternatively, they may result from mutant protein breakdown. Bands were visible by western blot analysis with antibody specific to the His-tag therefore the amino terminus of N protein was present in each. The band present in NΔ342-391 is similar in size to the amino terminus breakdown product observed during full-length His-N purification (Chapter 4). If this band also represents a breakdown product of mutant protein, it would suggest that the N is rather susceptible to degradation in a region approximately two-thirds along its length.

Purification was reasonably successful with recovery of most mutant proteins (Figure 6.5 a). The protein concentration recovered appeared to reflect the abundance of the original protein present in the clarified lysate.

The success of mutant protein analysis by north-western blot was disappointing. Despite the use of purified protein, an increased number of membrane washes and the addition of unlabelled RNA to saturate non-specific RNA association, high background radioactivity was routinely obtained (Figure 6.5 b). An intense band of the expected protein size was obtained in the NΔ121-160 lane representing a specific interaction between mutant protein and radiolabelled RNA. Faint bands of the correct protein size were also present in lanes with NΔ1-200 and NΔ142-391. However, visualisation of the latter proteins, which were present at high concentrations, may result only from non-specific association with radiolabelled RNA, although steps were taken to counteract this effect. Bands were not visible in the remaining lanes. This may indicate that the remaining mutants are unable to associate with RNA. However, it is probable that potential bands were obscured by high background radioactivity. Indeed, it is probable that several remaining mutants posses the ability to bind RNA as the RNA-binding site in the N protein of related viruses occurs within the amino-terminal two-thirds of the protein, for example, in Sendai virus (Buchholz *et al.*, 1993). The RNA-binding ability of

the 30 kDa fragment, identified in earlier chapters, should also be considered. It has been demonstrated to represent the amino-terminal region of the N protein and to retain the ability to bind RNA. Thus, like Sendai virus, the N protein of RSV appears to have an amino-terminal for nucleocapsid formation and a carboxy-terminal for association with the P protein (Murray *et al.*, 2001).

Despite repeated efforts to employ His-N deletion mutants to identify the RNA binding domain, the inherent difficulties in working with N were all too apparent. It would appear that the mutant proteins were unstable and prone to degradation. In addition, the deletions may affect the conformation of the remaining protein rendering it inactive. Concurrent work within the laboratory using N deletion mutants in the yeast two-hybrid system also encountered the problem that some deletions would prevent the N: P interaction. However, larger deletions, that encompassed smaller ones, regained the ability to bind P thus making any deduction speculative (C. Loney, personal communication). Indeed, Garcia-Barreno *et al.* (1996) were unable to map regions on N protein responsible for the formation of CI bodies using deletion mutants, yet could readily do so with P. This suggests that the tertiary structure of N might be quite complex. Similarly, data from Krishnamurthy & Samal (1998) would indicate that deletions within N protein may have other effects than the simple removal of binding sites.

Despite the reliability and success of the north-western blot assay with full-length N, it proved an unsuitable approach for the analysis of the mutant proteins. This may be due to the limited level of purity obtained with the mutant proteins. However, this is unlikely as successful north-western blot analysis was performed on transformed bacterial cell lysates (Chapter 5).

An alternative approach such as gel mobility shift analysis may have been employed to investigate the RNA-binding ability of the mutant proteins. Though by this method, protein purity may have presented a problem. However, due to time restraints, further analysis of the mutants was halted in preference of an alternative approach to investigate the N protein RNA-binding domain.

7 A Study of the RNA-Binding Domain of N Protein (II)

7.1 Introduction

The interaction between N protein and RNA was investigated by crosslinking studies and subsequent manipulation. His-N and radiolabelled RNA were crosslinked by exposure to UV radiation and the product separated by SDS-PAGE and visualised by phosphorimagery. Proteolytic and nucleic acid digestion enabled those peptides associated with radiolabelled nucleotides to be visualised. Analysis then progressed to the use of unlabelled RNA to enable investigation of peptides by mass spectrometry. The purpose of such analysis was to identify the sequences of N peptides modulated by the presence of RNA and thereby indicate regions of N protein involved in binding RNA.

To enable execution of crosslinking studies, greater purity of His-N than that so far achieved by affinity chromatography was required (Section 2.10.1). Therefore, further methods of protein purification were investigated using nickel affinity purified His-N as the starting material.

7.2 N Protein Purification by Ion-Exchange Liquid Chromatography

7.2.1 Determination of His-N Protein Iso-Electric Point

Ion-exchange liquid chromatography was chosen to further purify recombinant N protein. This required knowledge of the iso-electric point (pI) of His-N. The pI is defined as the pH value at which an amphoteric protein has a net zero charge.

The pI of His-N was investigated using the Phast System (Pharmacia Biotech). The pre-formed gels of Phast Gel IEF media are composed of homogenous polyacrylamide and contain Pharmolyte® Carrier Ampholytes, able to generate a stable, linear pH gradient within the gel. Proteins migrate under an electric field, relatively unhindered by the porous gel, to a point in the pH gradient that corresponds to their pI. His-N was separated in a Phast gel (pH 3 to 9) and visualised by silver-stain (Figure 7.1). The position within the gel to which His-N migrated suggests its pI is approximately 7.35 – 7.5. This pI range was confirmed by IEF software available in the SEQLab package (GCG). The protein sequences of His-N and native N were entered into the programme
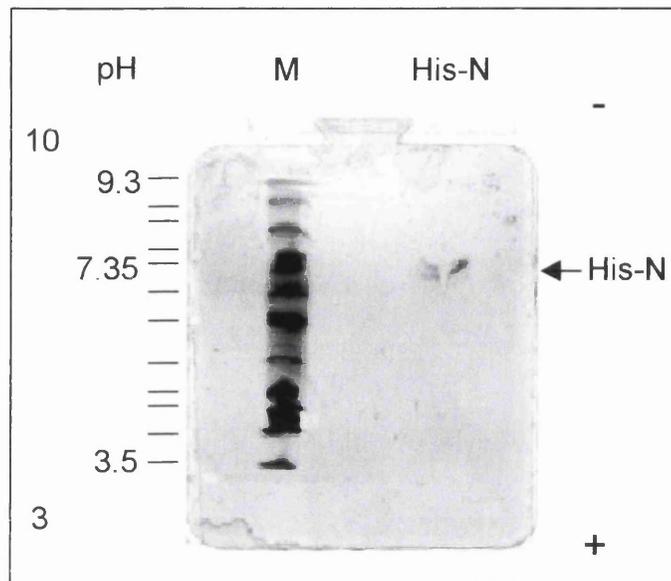
Figure 7.1: An iso-electric focussing gel of Histidine-tag N protein.

100 ng purified His-N protein was separated by electrophoresis in a pre-formed gel of Phast IEF media (Pharmacia Biotech). His-N was visualised by silver-stain and its pI value estimated between 7.35 and 7.5. (Marker, M).

and their theoretical pI values calculated. That of His-N was 7.63, close to that obtained experimentally. The theoretical pI of native N was 8.37; thus, the addition of the polyhistidine-tag, and inter-linking amino acids, appeared to have made the recombinant protein more acidic than unmodified N protein. However, despite the slightly more acidic nature of His-N it was still able to bind RNA, suggesting that the over-all basic charge of N is not required for binding RNA.

7.2.2   N Protein Purification by Cation-Exchange FPLC

The basic pI of His-N suggested the use of a cation-exchange resin. The Mono-S cation-exchange column (Pharmacia Biotech) contains the negatively-charged group methyl sulphonate (SP) bound to beaded polystyrene resin. This was selected for use in the purification of His-N. The recombinant protein was dialysed into 0.2 M Tris HCl (pH 6.5), loaded onto the SP column and eluted with a sodium chloride gradient (up to 1 M). Western blot analysis revealed the presence of His-N in the flow-through only, prior to application of the elution gradient. This suggested that His-N did not bind to the column. Rather than attempt to optimise this calculated purification step, the use of different ion-exchange resins was investigated.

7.2.3   Investigation of Ion-Exchange Resins

Sepharose-bound ions SP (strong anion) and DEAE (weak cation) were investigated to establish which was better able to bind His-N.

Purified His-N protein was combined with each resin in 0.2 M Tris (pH 6.5) and briefly shaken. The mixture was centrifuged for one minute at 6000 xg and the supernatant, removed and retained. His-N protein was eluted from the ion-exchange beads with a series of sodium chloride steps ranging from 0 to 2.5 M in 0.2 M Tris (pH 6.5) (Figure 7.2). With SP-resin, His-N protein was detected in each sample. A considerable amount of protein was present in the flow-through fraction. A major peak in protein elution occurred at 1 M NaCl and above, although a significant amount of His-N remained associated with the SP-resin. In contrast, with DEAE-resin only a small amount of protein was present in the flow-through fraction. Protein elution occurred at 1 M sodium chloride but again a small amount of His-N remained bound to the resin. Other protein bands were visible in the majority of lanes. A band of approximately 80 kDa may represent His-N protein dimers. Additional larger bands are visible in lanes overloaded with His-N demonstrating protein oligomerisation, able to survive the
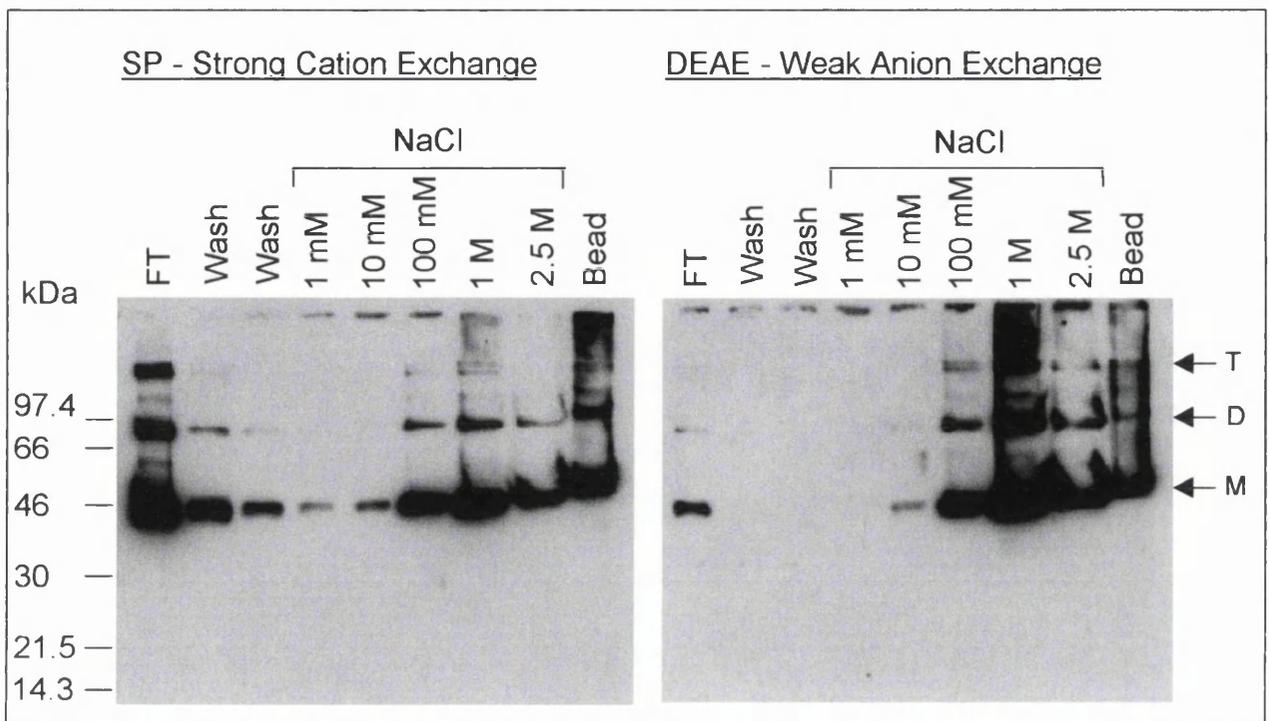
Figure 7.2: A comparison in the ability of ion exchange resins DEAE- and SP-agarose to associate with Histidine-tag N protein.

10 μg purified His-N was combined with each resin in a volume of 100 μl 0.2 M Tris (pH 6.5) (wash buffer). The resin was centrifuged and the supernatant removed and retained (flow-through, FT). His-N was eluted with a gradient from 0 to 2.5 M sodium chloride (NaCl) in wash buffer. Elutants were analysed by western blot with HRP-conjugated antibody directed against the Histidine-tag. Protein bands were visualised by ECL and autoradiography.

In addition to monomeric His-N (M), larger protein bands are present in several lanes, possibly representing His-N dimer (D) and trimer (T).

reducing conditions of SDS-PAGE. On the basis of these observations DEAE was used for further purification of His-N.

7.2.4 N Protein Purification by Ion-Exchange FPLC

A DEAE column was connected to an Äkta purifier chromatography system and washed with 0.2 M Tris buffer (pH 6.5) in preparation for His-N purification. The protein sample was applied to the column, washed with start buffer and eluted with a salt gradient from 0 to 1 M sodium chloride in 0.2 M Tris (pH 6.5). In Figure 7.3 (a), a single, tight protein peak was obtained in elution fractions 3 to 6. Unexpectedly, the peak occurred at an elution buffer concentration of 0.2 M sodium chloride, considerably lower than that observed in Figure 7.2. As observed in metal ion chelate chromatography, the nucleic acid trace imitated that of protein (Figure 4.8).

Protein in fractions 3 to 8 was separated by SDS-PAGE and visualised by silver-stain and by western blot (Figure 7.3 b). In the silver-stained gel, His-N was present as a band of 46 kDa in size and the highest protein concentration occurred in samples 3 to 5. In fractions 4 and 5, a larger band is visible, possibly resulting from N protein oligomerisation. In fractions with a high concentration of N, smaller protein bands were visible. A similar distribution of protein band intensity was observed by western blot, performed with antibody directed towards the histidine tag. Smaller protein bands (< 30 kDa) were again observed. This suggests that they result from His-N breakdown and possess the amino terminus of the protein. This figure suggests that recombinant N protein is free from contaminating proteins but possesses self-breakdown products, more evident since ion-exchange purification. To further improve His-N protein purification, size-exclusion liquid chromatography was performed.

7.3 Size-Exclusion Liquid Chromatography

A Superdex 6 column (Pharmacia Biotech) was used to further purify DEAE His-N protein. The column and Äkta purifier chromatography system were connected and washed with Tris-based RNA binding buffer, used in the gel mobility-shift and UV cross-linking experiments. The His-N protein sample was applied and washed through the column with RNA binding buffer. The elution profile in Figure 7.4 (a) showed a relatively large peak in fraction 13 followed by several small peaks. The large peak may represent full-length His-N and subsequent peaks, the smaller breakdown products of

Figure 7.3: DEAE ion exchange FPLC of Histidine-tag N (a) and subsequent analysis of the purified protein (b).

His-N protein, purified by HiTrap affinity chromatography, was applied to the DEAE column, washed with 0.2 M Tris (pH 6.5) and eluted with a NaCl gradient (a). Fractions were separated by 10% polyacrylamide SDS-PAGE and analysed by silver-staining and by western blot with HRP-conjugated antibody directed against the Histidine-tag (b).

(a)

(b)
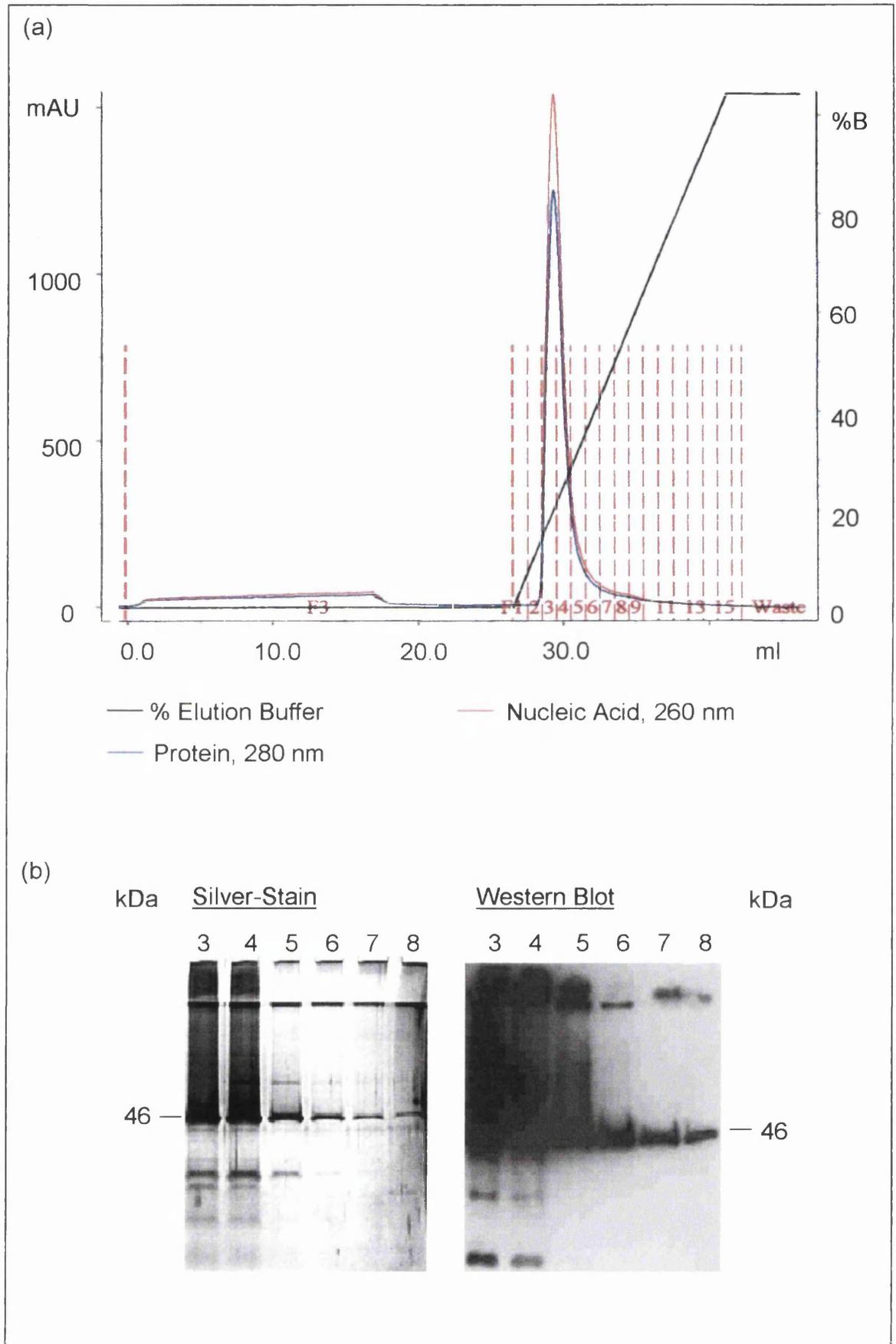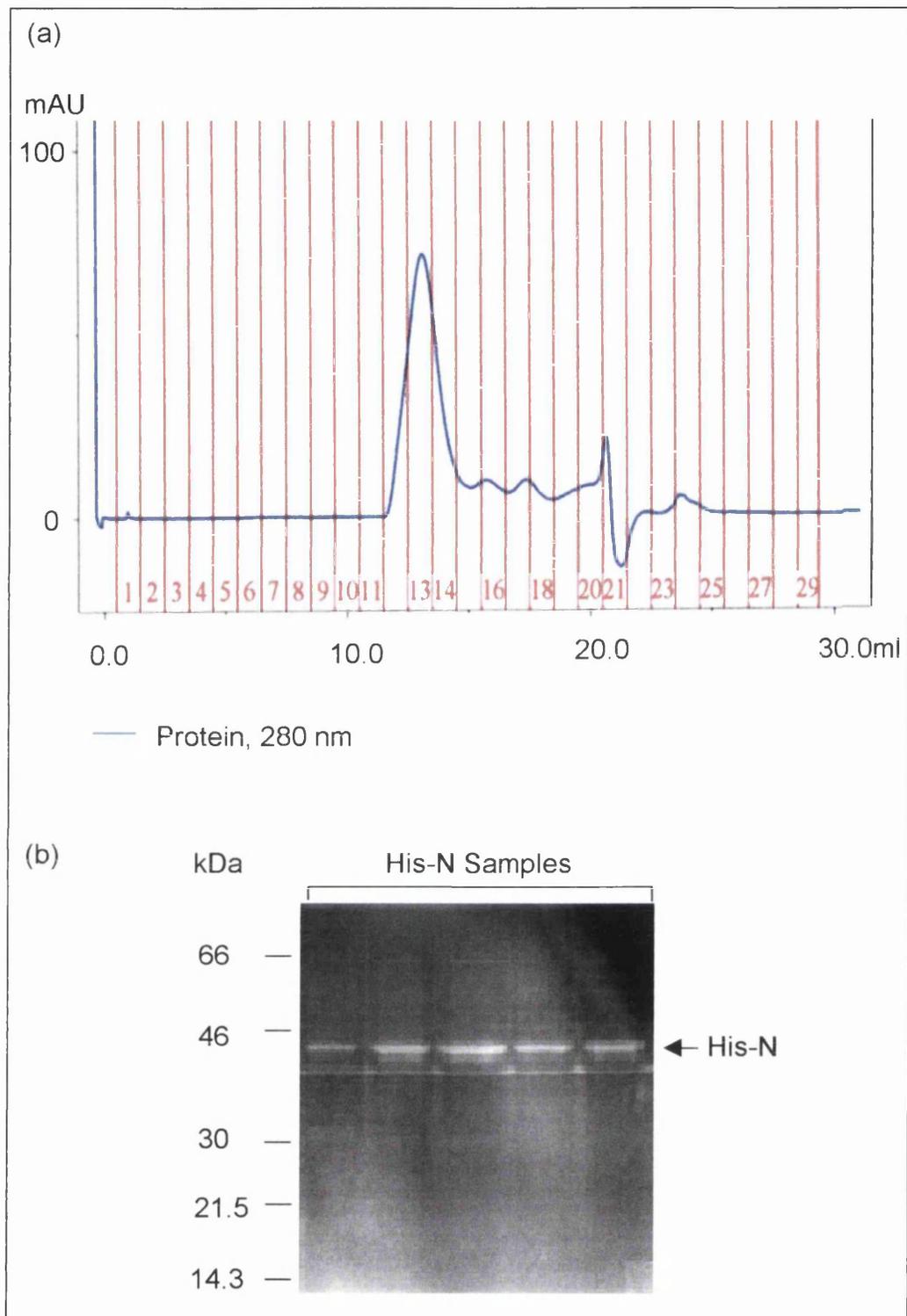
Figure 7.4: Size-exclusion FPLC of Histidine-tag N (a) and subsequent analysis of the purified protein (b).

His-N protein, purified by DEAE ion exchange chromatography, was applied to the size-exclusion column and eluted with Tris-based binding buffer (a). Fractions were separated by 10% polyacrylamide SDS-PAGE and analysed by silver-staining (b).

the protein. To confirm the identity and purity of the protein, fractions 11 to 15 were separated by SDS-PAGE and visualised by silver-stain (Figure 7.4 b). The His-N protein of approximately 46 kDa was present in each sample. At higher concentrations, there appears to be a second protein band visible immediately beneath the His-N band, although this may be an artefact of the gel or of image processing. The breakdown products of His-N protein were not detected.

The final product of affinity, ion-exchange and size-exclusion chromatography was considered of acceptable purity for use in UV crosslinking experiments and mass spectrometry.

7.4 UV Crosslinking Assay

7.4.1 Optimisation of assay conditions (i) UV crosslinking

The association between N protein and RNA was investigated using the UV crosslinking assay. The temperature, and duration, of incubation and the energy level of UV radiation were investigated to determine optimal reaction conditions. The most favourable conditions were an incubation of 20 minutes at 33°C followed by exposure to 25 000 μjoules cm^{-2} x100 UV radiation (Figure 7.5). In lane 1, 0.6 μg His-N was combined with 6500 cpm radiolabelled specific RNA, whereas 0.3 μg His-N was used in lane 2. A single band of approximately 60 kDa was present in each lane and represented crosslinked RNA and protein. Band intensity decreased as N protein concentration was reduced confirming the presence of N within the band complex.

The third sample lane in Figure 7.5 was generated by combining sodium hydroxide-treated N protein with RNA. In Chapter 5, it was determined that N protein binds RNA in a non-sequence specific manner. It is possible that the His-N may associate with bacterial RNA during expression; this has been alluded to before, as being problematic in the initial RNA-binding studies. Sodium hydroxide treatment was attempted to remove potential bacterial RNA from the protein. N protein was separated from associated RNA by SDS-PAGE and transferred to nitrocellulose membrane. The protein band was located by coomassie blue staining and eluted from the membrane by overnight incubation with 0.2 mM sodium hydroxide. RNA-free N protein was then combined with radiolabelled RNA and UV crosslinking performed. Full-length His-N was notably absent and only a single band of approximately 30 kDa was routinely
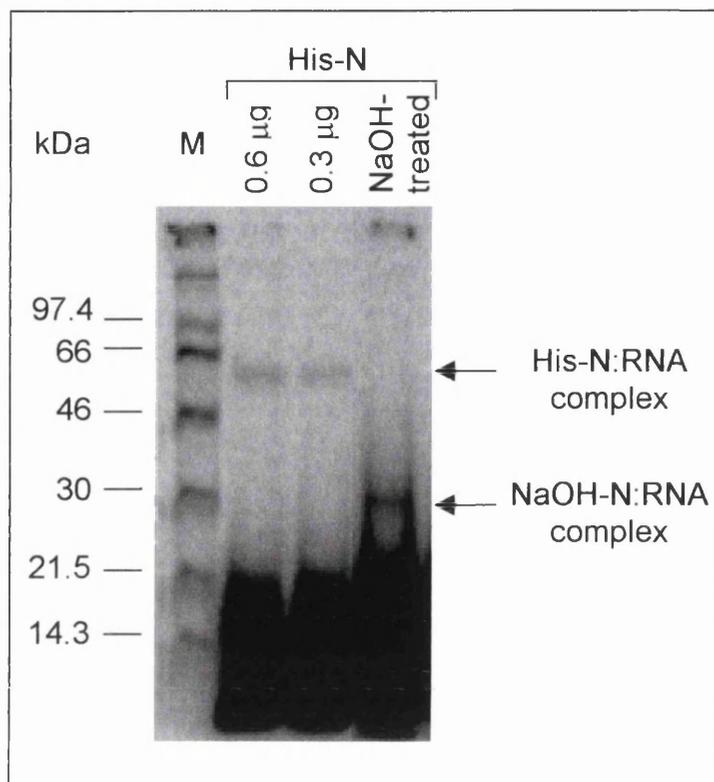
Figure 7.5: UV crosslinking assay of Histidine-tag N and radiolabelled RNA.

The UV crosslinking assay was performed with purified His-N and with sodium hydroxide (NaOH) -treated His-N protein. The latter was generated by separating purified His-N by SDS-PAGE, transferring to nitrocellulose membrane and from there releasing the protein by incubation with 0.2 mM sodium hydroxide (NaOH). This treatment was performed to disrupt any existing association between His-N and bacterial RNA, remaining from recombinant protein expression.

Protein and 10 000 cpm specific radiolabelled RNA were incubated for 20 minutes at 33°C and crosslinked by exposure to UV radiation. Protein:RNA complexes were separated by 10% polyacrylamide SDS-PAGE and visualised by phosphorimagery. (Marker, M).

obtained. This may be the result of sodium hydroxide cleavage of the N protein, creating a peptide that was able to bind RNA. The band was investigated by silver- and coomassie blue-stained protein gels and by western blot analysis with various N protein MAbs, which were previously mapped (Murray *et al.*, 2001), and the histidine-tag MAb. The band was absent from each assay and therefore remained unidentified, however due to the reactivities of the MAbs we can exclude certain regions of the N in being involved in RNA binding (see Discussion Chapter 9).

7.4.2 Optimisation of assay conditions (ii) RNase A concentration

Individual N proteins probably associate with only a few nucleotides within an RNA sequence. For example, the N proteins of rabies and Sendai virus each bind eight and six nucleotides respectively (Egelman *et al.*, 1989; Iseni *et al.*, 1998; Vulliemoz & Roux, 2001). Therefore, either the probe (47 nt) may be bound by more than one N protein or regions of the RNA probe may remain unbound by protein. In the UV crosslinking assay, RNA not protected by protein was removed by RNase A. The protein may only be visualised when it is bound to radiolabelled RNA therefore observation of an RNase A-treated protein band is dependent on the presence of labelled nucleotides. In Figure 7.6 (a), a noticeable decrease in size was observed following RNase A digestion of the His-N:RNA complex, with a reduction from 60 kDa to 46 kDa. Decrease in mass was accompanied by a reduction of the band intensity, as a result of the removal of unbound nucleotides. At the gel bottom, the intense band of free RNA probe disappeared completely upon addition of RNase A.

RNase A was titred to determine its optimal concentration. N protein and radiolabelled RNA were each used at approximately half the concentrations used in Figure 7.6 (a) resulting in lower band intensities (Figure 7.6 b). The N protein-RNA band intensity and to a limited extent its size, increased as the concentration of RNase A was decreased from 10 µg to 1 ng. Free RNA was absent at RNase A concentrations of 10 µg and 1 µg. Thus, 1 µg of RNase A was used in subsequent experiments as the N protein-RNA band was visible but minimal full-length free probe was present.

The success of RNase A treatment suggests that, unlike the viral nucleocapsid, the protein-RNA complex generated was not resistant to RNase digestion. This may

Figure 7.6: RNase A digestion of UV crosslinked Histidine-tag N and radiolabelled RNA.

0.6 μg (a) or 0.3 μg (b) purified His-N protein was incubated with 10 000 cpm specific radiolabelled RNA for 20 minutes at 33°C and crosslinked by exposure to UV radiation. RNase A was added to the crosslinked samples and incubated for 15 minutes at 33°C. Digests were separated by 10% polyacrylamide SDS-PAGE and visualised by phosphorimagery. In (b), RNase A was titrated in a range from 0.001 μg to 10 μg. (Marker, M).
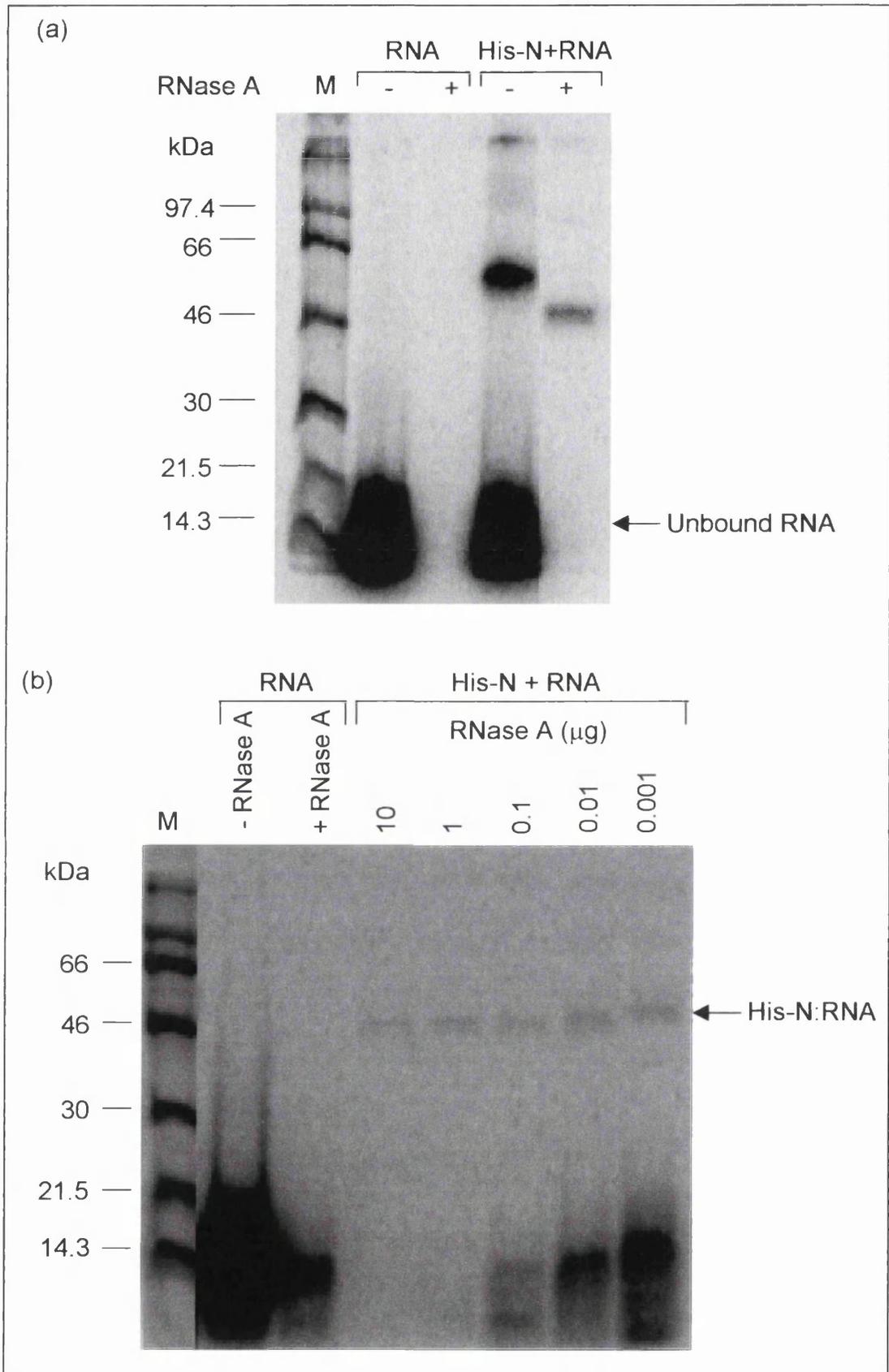
suggest that probe is not encapsidated along its entire length but instead is bound by N protein in a less structured conformation.

7.4.3 Optimisation of assay conditions (iii) Radiolabel

In initial experiments, radiolabelled CTP was used to generate the RNA probe. However, only nine CTP residues are present in the RNA sequence used in comparison with 24 ATP residues. Therefore, probes with ^{32}P-CTP and with ^{32}P-ATP were synthesised and the intensity of the resulting N protein-RNA bands was compared. A range from 2500 cpm to 20 000 cpm of labelled RNA was combined with 0.6 µg N protein, crosslinked and digested with 1 µg RNase A prior to separation by SDS-PAGE (Figure 7.7 a). RNase A treatment was omitted in Figure 7.7 (b). An His-N-RNA band at 46 kDa was present within each lane. Band intensity increased as the concentration of both CTP- and ATP-radiolabelled RNAs were increased. Bands created with the labelled-ATP probe were more intense than those with CTP-labelled probe. Band intensities appeared to approximate to the expected ratio of 8 : 3, that is, 24 ATP residues to nine CTP residues. Due to the improved band intensity, specific probe created with radiolabelled-ATP was used in subsequent experiments.

7.4.4 Optimisation of assay conditions (iv) Proteolytic digestion

To detect specific regions of N protein responsible for binding RNA, proteolytic analysis of N, bound to RNA, was performed. Initially, various means of N protein digestion were investigated in the absence of RNA. Digestion was performed with the endoproteinases V8, clostripain, LysC and trypsin (Promega). Their individual digestion products were predicted by GCG software analysis to yield from 18 (clostripain) to 48 (trypsin) peptides. Digestions were performed in UV crosslinking binding buffer and were incubated overnight at either room temperature or 37°C. The extent of N protein digestion was investigated by 10% SDS-PAGE and by silver-staining (Figure 7.8 a). Protein digestion had occurred with each enzyme, though with varied success. Full-length protein remains in each sample, though at a low concentration with trypsin. Clostripain digestion was the least successful with only two faint product bands visible, the majority of protein remained as full-length. Many digestion products were created with V8 protease but several were only a little shorter than full-length His-N protein. The products of LysC digestion were smaller but limited in number, the major bands were 27 kDa and 20 kDa. N protein digestion with trypsin was the most complete with
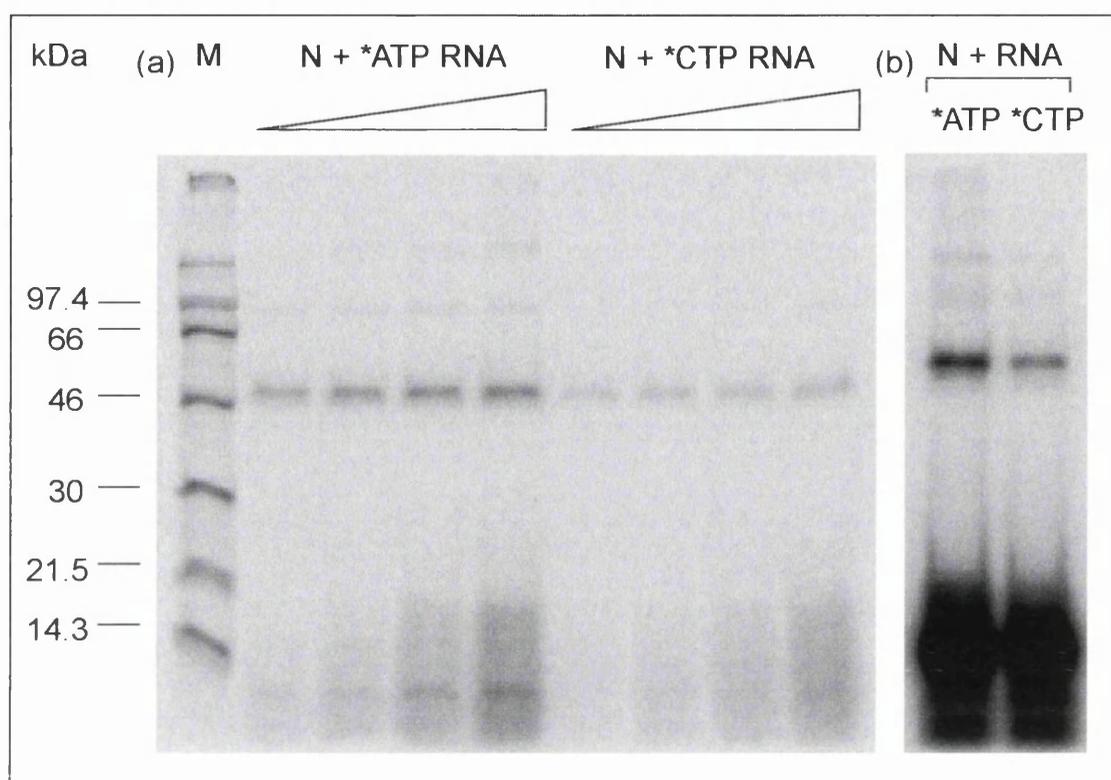
Figure 7.7: UV crosslinking assay of Histidine-tag N and RNA labelled with either ^{32}P-ATP or ^{32}P-CTP.

0.6 μg His-N protein was combined with a range from 2500 cpm to 20 000 cpm specific ^{32}P-ATP or ^{32}P-CTP radiolabelled RNA, incubated for 20 minutes at 33°C and crosslinked by exposure to UV radiation. RNA was digested with 1 μg RNase A for 15 minutes at 33°C (a). Samples were separated by 10% polyacrylamide SDS-PAGE and visualised by phosphorimagery. In (b), RNase A digestion was not performed. (Marker, M).

Figure 7.8: Protease digestion of Histidine-tag N protein.

The proteases V8, clostripain and trypsin were used in a weight ratio of 1: 20 with His-N protein. LysC and His-N were used in a weight ratio of 1: 100. Digestion was performed overnight at either room temperature or at 37°C (clostripain). Products were separated by 10% polyacrylamide SDS-PAGE and visualised by silver-stain (a). The total possible number of digestion products with each protease are as follows: 29 (V8 protease), 18 (clostripain), 31 (LysC) and 48 (trypsin). However, fewer fragments were expected as N protein conformation may make certain protease digestion sites inaccessible.

The products of His-N protein digestion with trypsin were separated by 16% polyacrylamide Tris-tricine gel electrophoresis and visualised by silver-stain (b) and by western blot, with polyclonal antibody directed against the N protein (c). (Marker, M).

(a)

kDa   His-N   V8   Clost.   LysC   Trypsin

66 —

46 —

30 —

21.5 —

14.3 —

← Full-length His-N

(b)

kDa   M   His-N   His-N + Trypsin

66 —
46 —
30 —
21.5 —
14.3 —

(c)

His-N   His-N + Trypsin   kDa

— 46

— 30
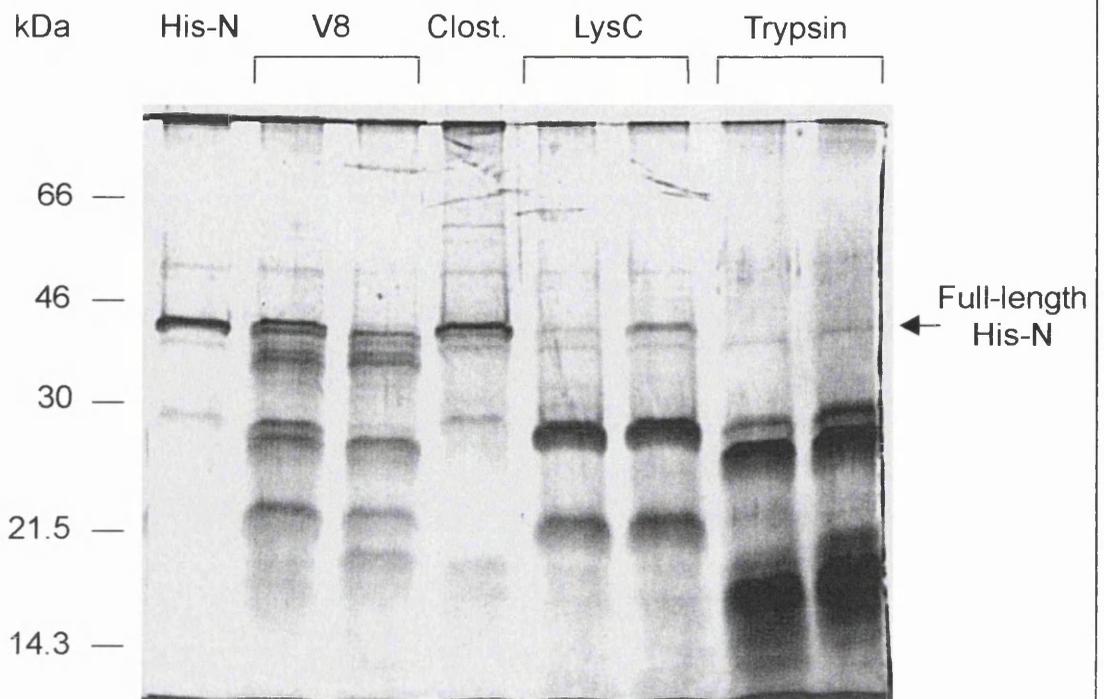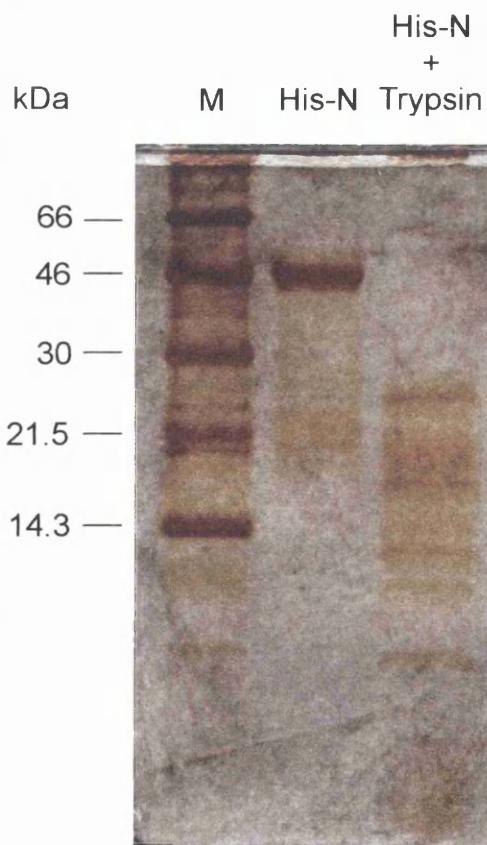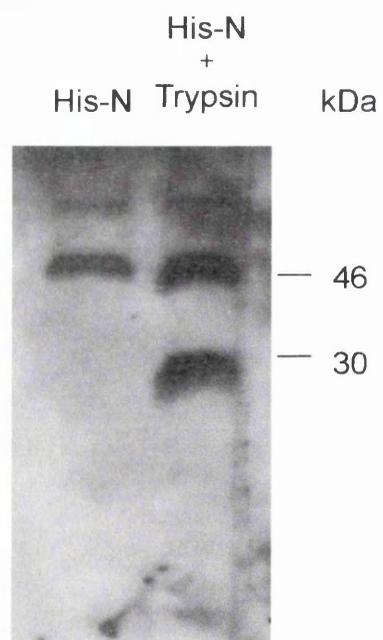
only minimal full-length protein remaining. Many peptides were generated including several at approximately 21.5 kDa and smaller. Digestion products of this size are most desirable for eventual analysis by mass spectrometry.

The products of trypsin digestion were further analysed by silver-stain and by western blot. His-N was treated with trypsin for three hours at 37°C and the digestion products separated by 16% polyacrylamide gel and visualised by silver-stain (Figure 7.8 b). Trypsin digestion removed full-length N protein and created a number of peptides smaller than 30 kDa. Eight discrete bands were clearly visible, though additional peptides may also be present. Several bands with a mass of 14.3 kDa or less may be seen. A western blot of His-N, digested with trypsin, was performed using a polyclonal antiserum specific to the N protein (Figure 7.8 c). In this example, full-length protein remained at 46 kDa. An intense band of a digestion product was visible at approximately 30 kDa. In addition, very faint bands of approximately 21.5 kDa and smaller than 14.3 kDa were present. However, many of the smaller peptides visible by silver-stain were not detected by the antibody. The peptides detected by silver-stain and by western blot are relatively large suggesting that incomplete digestion of N protein has occurred. Should digestion proceed to completion, 48 peptides would be generated, the majority of which would be undetectable by these means.

The small number of relatively large digestion products visible in the silver-stained gel suggests that many sites are not available due to N protein conformation. Protein denaturation was attempted to make more trypsin sites available and thereby generate smaller digestion products. This was achieved by either heating His-N to 95°C for 20 minutes, or chemically denaturing the protein with guanidine hydrochloride, prior to crosslinking. Denatured N protein-RNA association occurred, however, in the absence of trypsin, the complex was retained in the well and did not migrate into the gel. Upon trypsin treatment, no bands were visible, possibly as a result of complete digestion. The RNA-binding site of N may be conformational, created by folding of the protein. Therefore, use of the denatured protein to bind RNA may not be an appropriate method to investigate the RNA-binding site. Trypsin digestion of native protein generated relatively small peptides (Figure 7.8 b), this was considered sufficient for initial experiments and denaturation of the N protein was pursued no further.

7.4.5 Optimisation of assay conditions (v) UV crosslinking and protease digestion of His-N:RNA.

Trypsin digestion of N was combined with the UV crosslinking assay. Protease digestion was performed for three hours at 37°C following N protein-RNA exposure to UV radiation and RNase A digestion. The radiolabelled products were separated by 16% polyacrylamide gel and visualised by phosphorimagery (Figure 7.9). In Figure 7.9, three groups of samples [RNA only], [N protein and RNA only] and [N protein, RNA and trypsin], digested with 0 μg, 0.1 μg or 1 μg RNase A, are present. As previously observed, the N protein:RNA complex forms a band of approximately 60 kDa that decreases to approximately 46 kDa upon addition of RNase A. In samples digested with trypsin, two novel bands replace that of full-length N protein. These bands are of approximate sizes 13.5 kDa and 10 kDa (N^{*1} and N^{*2}, respectively) and persist upon the increase of RNase A concentration. These features, the intensity of the bands and their absence from other lanes, suggest that they represent N peptides associated with radiolabelled nucleotides.

7.5 Mass Spectrometry Analysis

Two N peptides bound by radiolabelled nucleotides were identified by UV crosslinking analysis of N protein and RNA, subsequent proteolytic and nucleic acid digestion and visualisation by SDS-PAGE. Mass spectrometry was attempted to identify the sequences of N peptides involved in nucleic acid binding.

Samples to be analysed by mass spectrometry were prepared with non-radiolabelled RNA. The preparation of these samples was performed in parallel with samples created with radiolabelled RNA. SDS-PAGE was performed to confirm recovery of the radiolabelled samples and was taken as an indication of the status of the unlabelled samples. Slide-A-Lyzer mini dialysis units were used to dialyse the nucleotide-bound N peptide samples against 0.1% tri-fluoroacetic acid (TFA). Samples were removed from the dialysis units and those with radiolabelled RNA were investigated by 16% polyacrylamide Tris-tricine gel electrophoresis (Figure 7.10). The expected bands of full-length N protein:RNA complex, both with and without RNase A, were visible (lanes 3 and 4). A faint band, of approximately 13.5 kDa (N^*), was visible in the lane with trypsinised sample (lane 5). This band was also present in the sample digested with RNase A alone (lane 4). However, the band was absent from lane 2 (digested RNA

Figure 7.9: UV crosslinking assay of Histidine-tag N and radiolabelled RNA with subsequent RNase A and trypsin digestion.

2 µg purified His-N protein and 15 000 cpm specific radiolabelled RNA were incubated for 20 minutes at 33°C and crosslinked by exposure to UV radiation. 0, 0.1 µg or 1 µg RNase A was added and the samples incubated for 15 minutes at 33°C. Trypsin was added to certain samples in a ratio of 1: 20. The reactions were incubated overnight at 37°C. Samples were separated by 16% polyacrylamide Tris-tricine gel electrophoresis and visualised by phosphorimagery.

A band of approximately 46 kDa is present in lanes of full-length His-N-RNA (N) treated with RNase A. In lanes of sample further treated with trypsin, bands of 13.5 kDa (N *^{1}) and 10 kDa (N *^{2}) are present. Intensity of digested RNA at the gel bottom is less in protein-RNA samples treated with both RNase A and trypsin. (Marker, M).

Lanes	1	2	3	4	5	6
N	-	-	+	+	+	+
RNA	+	+	+	+	+	+
RNase	-	+	-	+	+	+
Trypsin	-	-	-	-	+	+
Dialysis	-	-	-	-	-	+

46 —
30 —
21.5 —
14.3 —
12.5 —

← His-N

← N *1
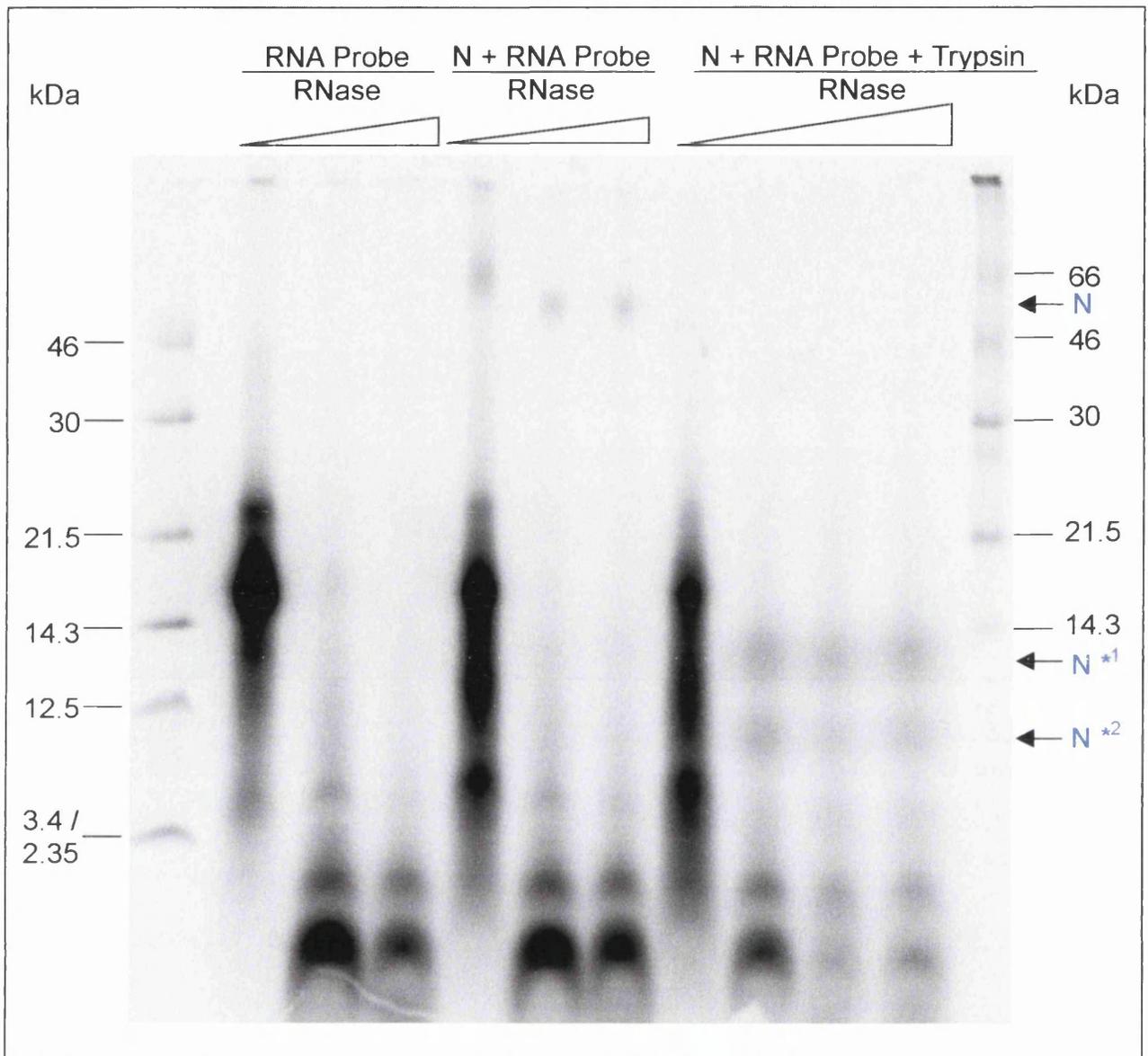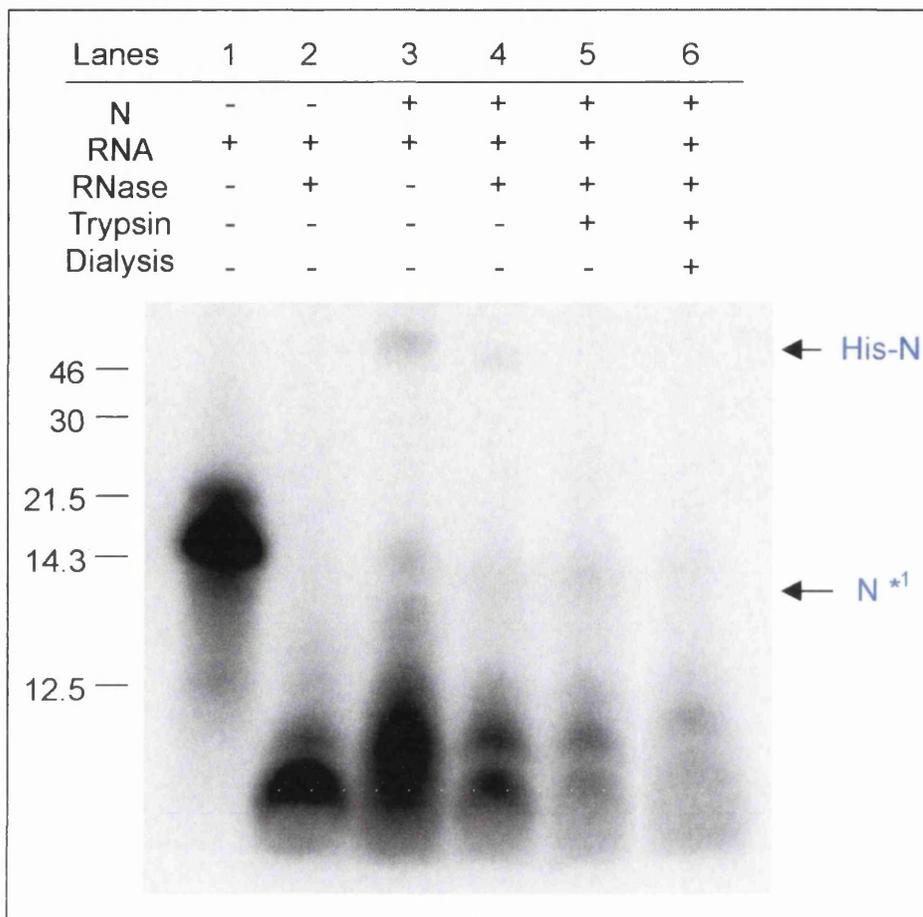
Figure 7.10: UV crosslinking assay of Histidine-tag N and radiolabelled RNA, with subsequent RNase A and trypsin digestion and dialysis.

1 μg purified His-N protein and 10 000 cpm specific radiolabelled RNA were incubated for 20 minutes at 33°C and crosslinked upon exposure to UV radiation. Samples were digested with 0.1 μg RNase A for 15 minutes at 33°C followed by digestion with 0.1 μg trypsin overnight at 37°C. Dialysis was performed in a Slide-A-Lyzer mini-dialysis unit (3000 mwco) against 0.1% TFA for 4 hours at 4°C. Samples were separated by 16% polyacrylamide Tris-tricine electrophoresis and visualised by phosphorimagery. A band of 13.5 kDa (N *1) was recovered from dialysis of the final digestion products.

only) suggesting that it was of N protein origin. More importantly, the 13.5 kDa peptide and additional smaller bands were recovered from dialysis, indicating similar success with the unlabelled samples. Therefore, the latter samples were snap frozen in a dry ice-methanol bath and freeze-dried overnight at -60°C, under vacuum. They were then sent to A. Wilson at the Department of Microbiology, University of York and analysis by mass spectrometry performed by P. Ashton.

Four samples were prepared and analysed by MALDI mass spectrometry. Prior to analysis of each sample, the machine was calibrated using close external calibration leading to a mass accuracy of 0.2 Da. Within the preliminary study, a mass spectrum of less than 3800 m z^{-1} was investigated in each sample. This would not enable detection of the labelled peptides of sizes 13.5 kDa and 10 kDa observed in Figure 7.9.

Samples of [RNase A & Trypsin] and of [RNA & RNase A & Trypsin] provided negative controls enabling peaks contributed by RNA, trypsin and RNase A to be discounted from the experimental spectra. Peptides were not detected in either control suggesting that peaks obtained in the experimental samples originated from N protein alone. Spectra were obtained of samples [N protein & RNase A & Trypsin] (sample 1) and [N protein & RNA & RNase A & Trypsin] (sample 2). They were of good quality suggesting that trypsin digestion had gone largely to completion and that no buffer salts remained. The spectra of experimental samples are present in Figure 7.11 and the peak masses displayed in Table 7.1. The peak values listed are of the mono-isotopic mass of each peptide, that is, the mass of the peptide containing only ^{12}C atoms. Peptides are usually represented by a series of peaks that differ by 1 Da or more as ^{13}C atoms become incorporated.

The mass peaks of each list were used to search the NCBI 'nr' database using MASCOT, with a tolerance of 0.5 Da. The N protein was not retrieved from either list. However, there are several masses in each spectra that correspond to predicted products of tryptic-digestion of the N protein. Peptide positions within N protein are displayed in Figure 7.12. Of these peptide sequences, most were present in either both samples or in sample 2 only. There was only one peptide identified in sample 1 and absent from sample 2 that potentially represented sequence altered in mass due to association with nucleotides. The peptide sequence is AYAEQLK (352 to 358 amino acids) and has a mass of 822.4774 Da. In the nucleocapsid protein of related viruses,

Figure 7.11: Graphs of the spectra obtained from the mass spectrometry of digested, UV crosslinked Histidine-tag N and RNA.

The control sample of His-N, RNase A and trypsin produced the mass spectrum in (a). The experimental sample of His-N, specific RNA, RNase A and trypsin produced the mass spectrum in (b). Arrows indicate masses that correlate with peptide sequences present in His-N. Black arrows indicate masses present in both samples. Red arrows identify masses unique to each spectrum. These masses correlate to the following peptide sequences:

1 AYAEQLK, 2 NQDLYDAAK, 3 MLKEMGEVAPEYR, 4 LCGMLLITEDANHK.
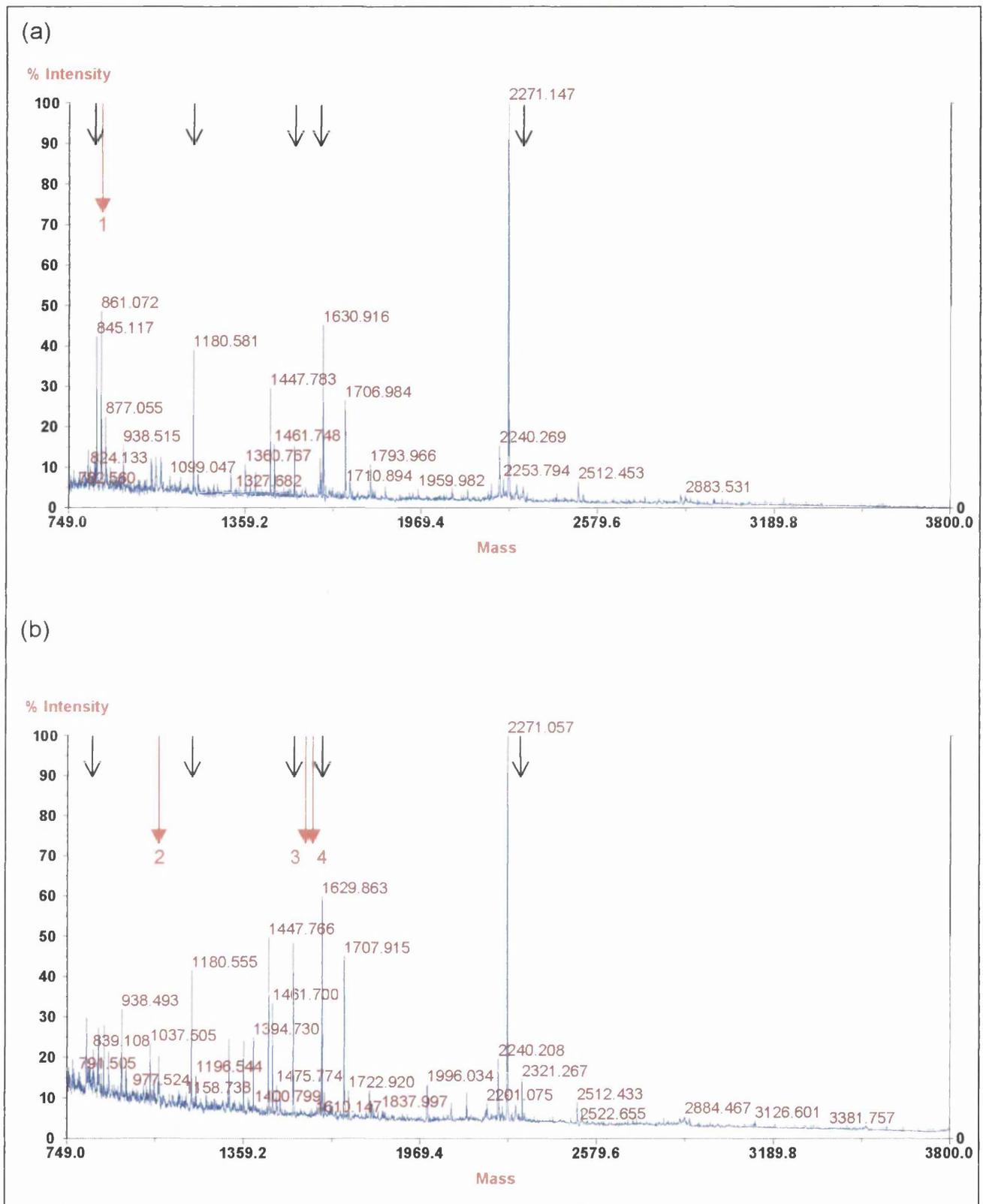
Table 7.1: Particle masses of digested, UV crosslinked Histidine-tag N and RNA obtained by mass spectrometry.

Sample 1 is the control sample [His-N protein & RNase & Trypsin]. Sample 2 is the experimental sample [His-N protein & RNA & RNase & Trypsin]. Masses in m z^{-1}.

Sample 1	Sample 2	N Peptide Sequence
	759.442	
	766.3549	
767.3994		
	787.4917	
798.5936		
	815.4605	
816.5365	**816.5087**	**SGLTAVIR**
822.4774		**AYAEQLK**
823.1333		
	827.6026	
	828.4718	
839.0978	839.1085	
842.5499		
845.1173		
	855.0874	
	858.4345	
	859.3721	
861.0729	861.5828	
877.0553	877.0491	
	883.5024	
	893.028	
918.3294		
922.6907		
938.515	938.4934	
948.5035		
	952.4926	
	954.4809	
	963.6028	
	967.7312	
	976.4101	
	1013.5844	
1034.1417		
	1036.5503	
	1037.5059	**NQDLYDAAK**
1050.1397	1050.0955	
1052.0722		
	1065.5528	
	1066.0928	
1066.6012	1066.544	
1069.0783		

Table 7.1: continued.

Sample 1	Sample 2	N Peptide Sequence
1069.607		
	1084.7039	
	1088.1625	
	1114.6798	
	1140.5016	
	1162.5725	
	1179.5887	
1180.5818	**1180.5551**	**EMGEVAPEYR**
	1184.6323	
	1196.544	
	1198.7425	
	1231.6573	
	1233.6199	
	1289.6612	
	1300.751	
	1308.6354	
1310.6927	1310.6898	
	1323.7349	
1326.7204		
	1329.7413	
	1346.8352	
	1347.5231	
	1360.7487	
	1362.7044	
1360.7679	1390.6666	
1394.7285	1394.7306	
	1431.9245	
	1442.9607	
1446.7023	1446.7691	
1447.7832	1447.7662	
1461.7486	**1461.7008**	**DIANSFYEVFEK**
	1474.7444	
	1475.7749	
	1485.7084	
	1501.7906	
1530.8932	1530.8506	
	1531.8512	
	1534.7367	**MLKEMGEVAPEYR**
	1556.7914	
	1557.8808	**LCGMLLITEDANHK**
1566.5763		
1567.9811		
	1568.8403	
	1572.6446	
1620.7776	1620.8015	
	1628.8499	

Table 7.1: continued.

Sample 1	Sample 2	N Peptide Sequence
1629.9216	**1629.8632**	LGGEAGFYHILNNPK
	1640.8078	
	1667.9294	
	1676.8248	
	1688.4504	
1706.9847	1706.9136	
	1710.7809	
	1716.8162	
1722.9355	1722.9209	
	1734.83	
	1751.4945	
1793.9668	1793.9314	
	1796.8372	
	1802.9332	
	1841.8725	
	1845.8296	
	1857.1835	
	1960.6533	
	1993.9926	
2076.0719		
	2077.1376	
	2118.2952	
	2130.0706	
	2131.0559	
	2136.7016	
	2200.0408	
	2202.1364	
2239.3157	2239.2003	
	2248.1334	
	2255.6927	
2270.1554	2270.0576	
	2273.9852	
	2299.0849	
	2300.1388	
2320.3252	**2320.3374**	ENGVINYSVLDLTAEELEAIK
	2321.267	
	2346.1578	
	2352.964	
	2435.8487	
2511.4753	2511.3874	
	2515.1021	
	2527.3576	
	2548.2899	
	2706.2879	
	2718.5462	

Table 7.1: Continued.

Sample 1	Sample 2	N Peptide Sequence
	2866.4886	
	2882.4656	
	2931.5098	
	2963.2653	

1   HHHHHHHHHHSSGHIEGRHMLEDPMALSKVKLNDTLNKDQLLSSSKYTIQ

51   RSTGDSIDTPNYDVQKHINKLCGMLLITEDANHKFTGLIGMLYAMSRLGR

101 EDTIKILRDAGYHVKANGVDVTTHRQDINGKEMKFEVLTLASLTTEIQIN

151 IEIESRKSYKKMLK.EMGEVAPEYRHDSPDCGMIILCIAALVITKLAAGDR

201 SGLTAVIRRANNVLKNEMKRYKGLLPKDIANSFYEVFEKHPHFIDVFVHF

251 GIAQSSTRGGSRVEGIFAGLFMNAYGAGQVMLRWGVLAKSVKNIMLGHAS

301 VQAEMEQVVEVYEYAQKLGGEAGFYHILNNPKASLLSLTQFPHFSSVVLG

351 NAAGLGIMGEYRGTPRNQDLYDAAK.AYAEQLK.ENGVINYSVLDLTAEELE

401 AIKHQLNPKDNDVEL

Figure 7.12: The position of peptides identified by mass spectrometry within the Histidine-tag N protein sequence.

The Histidine-tag and plasmid-derived linkage amino acids are in blue type. Sites of potential trypsin cleavage are underlined. Peptides identified by mass spectrometry are in red type. The peptide present in [N protein & RNA & RNase & Trypsin] only is in green type.

the region responsible for binding RNA occurs most frequently in the amino terminal two-thirds of the protein (Buchholz *et al*, 1993). However, this peptide sequence is present within the carboxy region of the N protein. Larger masses are present in sample 2 that approximately correspond to this peptide sequence associated with either one or two nucleotides.

The peptide identified as a candidate for modification by RNA is considerably smaller than radiolabelled peptides visualised by SDS-PAGE (Figure 7.9). These latter peptides were of approximate sizes 10 kDa and 13.5 kDa, ten-fold larger than the peptide sequences identified by mass spectrometry. Sequenced peptides were either too small to be retained by the gel or were masked by bands created by digestion of RNA probe. The larger peptides escaped detection by mass spectrometry due to the parameters of measurement being limited.

Initial investigation by the UV crosslinking assay and mass spectrometry has been successful in terms of retrieval of peptide sequences from the resulting samples. Their success in investigating the RNA-binding sequence has been more limited with the identification of only one peptide candidate. However, analysis by mass spectrometry has been performed only once due to time constraints and machine failure. The assay should be repeated to confirm results and to identify other sequences of N protein involved in binding RNA. In addition, greater concentrations of N protein could be used, additional methods of protein digestion investigated and the spectra of larger masses analysed.

7.6 Discussion

The purpose of work presented in this chapter was to investigate the domains of the N protein responsible for binding RNA. Analysis by UV crosslinking and by mass spectrometry was performed to complement that achieved with N protein mutants, presented in Chapter 6.

The His-tag N protein had previously been purified by affinity chromatography, presented in Chapter 4. However, bacterial contaminants from recombinant protein expression remained in the purified protein sample (Figure 4.9). Therefore, to enable analysis by UV crosslinking and by mass spectrometry, further purification of the N

protein was required. Ion-exchange and size-exclusion chromatography were performed to generate the final purified product, suitable for subsequent analysis. Initially, the experimental pI of His-N was obtained to enable selection of an appropriate resin-bound ion for ion-exchange chromatography. The value obtained (pH 7.35 – 7.5) was similar to the theoretical value of 7.63, obtained using IEF software available in SEQLab. This value is lower than the theoretical value obtained for N protein (pI 8.37). Despite the addition of 10 histidine residues in the protein tag, native N is more basic than the His-N protein. Determination of protein pI led to the selection of 0.2 M Tris (pH 6.5) and SP as an appropriate buffer and ion-resin, respectively. The buffer pH was selected to maintain the protein with a net positive charge. However, under the buffer conditions selected, protein purification by cation-exchange chromatography was not achieved. In place, the weak cation DEAE was considerably more successful. His-N was purified by DEAE anion-exchange FPLC, with protein elution at 0.2 M NaCl forming a single, tight peak in the resulting trace (Figure 7.3). Upon investigation by silver-stain gel and western blot, bacterial proteins were no longer detected. However, His-N breakdown products were present. These smaller protein bands were removed by size-exclusion chromatography and their absence confirmed by silver-stain (Figure 7.4).

Protein purity was considered sufficient for use in the UV crosslinking assay. Incubation at 33°C for 20 minutes followed by exposure to 25 000 µjoules cm^{-2} x100 UV radiation were determined as optimal conditions. In Figure 7.5, bands representing His-N-radiolabelled RNA association are 60 kDa in size. Also present in this figure is a band created by the crosslinking of RNA with His-N treated with 0.2 mM NaOH. The purpose of this treatment was to remove bacterial RNA from the recombinant protein. This was achieved by protein separation by SDS-PAGE, transfer to nitrocellulose and subsequent protein recovery from the membrane by NaOH treatment. The intensity of the band generated by UV crosslinking was very high, however, the protein fragment responsible could not be detected by silver-stain gel or by western blot (data not included). Presumably, the fragment was generated by NaOH chemical cleavage of His-N. The strong RNA interaction exhibited by this protein truncation makes further analysis desirable as it would aid identification of the N protein RNA-binding site.

Upon RNase A treatment of UV crosslinked His-N:RNA complex, the band decreases in both size (from 60 kDa to 46 kDa) and intensity as a result of nucleotide loss from

the complex. 1 μg RNase A was selected for use in subsequent assays following enzyme titration (Figure7.6). Unlike viral nucleocapsids, the RNA:His-N complex is sensitive to RNase A digestion suggesting that the RNA probe is not encapsidated by His-N. Protein binding may perhaps be less structured with either individual or only few protein monomers binding along the length of the oligo. Indeed, the detection of only a single band upon RNA-N binding, irrespective of RNase A treatment, would suggest the occurrence of binding between individual N proteins and RNA. Otherwise, a series of bands of increasing size would be visible, created by the step-wise addition of protein monomers to RNA probe. This limited binding may be affected by the presence of bacterial RNA, remaining from recombinant protein expression. Also, it may result from the addition of RNA probe, post-His-N translation, and therefore, after possible protein maturation and potential alteration of RNA-binding domains.

His-N was digested individually with proteases V8, clostripain, LysC and trypsin (Figure 7.8). Digestions with LysC and with trypsin were most successful with little full-length His-N remaining. LysC digestion generated only a limited number of products, including two main bands of 27 kDa and 20 kDa in size. Trypsin digestion was most successful, producing at least eight products, all of which were smaller than 30 kDa. Western blot analysis with polyclonal antibody directed against N detected an intense band of approximately 30 kDa. Faint bands of 21.5 kDa and less than 14.3 kDa were also visible.

UV crosslinking of N protein with RNA and subsequent digestion with RNase A and trypsin yielded two radiolabelled peptides N^{*1} (13.5 kDa) and N^{*2} (10 kDa) (Figure 7.9). The identity of these peptides remains unknown.

To further investigate nucleotide-bound N peptides, mass spectrometry was performed using a MALDI mass spectrometer. Spectra obtained for [N & RNase A & Trypsin] and [N & RNA & RNase A & Trypsin] were of good quality suggesting digestion had largely gone to completion and that buffer salt was absent (Figure 7.11). Masses from each sample were investigated in the NCBI 'nr' database using MASCOT. A number of N peptides generated by trypsin digestion were identified in both samples. However, only one peptide was unique to sample 1, possibly representing a peptide modified by the addition of RNA and hence, absent from sample 2. This peptide was AYAEQLK, present from amino acid 352 to 358 within the carboxy terminal of the N protein (Figure

7.12). The peptide has a mass of 822.4774 Da, considerably smaller than those observed in gel analysis of the UV crosslinking assay. The lack of detection of peptides N^{*1} and N^{*2}, identified by UV crosslinking, was due to the parameters of measurement of the mass spectrometer being unsuitable for the analysis of peptides of this size.

Masses of a size appropriate for peptide AYAEQLK bound by one or two nucleotides are present in sample 2. To investigate the presence of like sequence in these larger masses, the peptides could be fragmented and the resulting fragmentation patterns compared with that of the 822.4774 Da peak, allowing for possible mass shifts that correspond to presence of one or more nucleotide. Alternatively, the peptides may be sequenced on a Q-To-F mass spectrometer. However, peptide analysis by these means would require considerably higher concentrations of N protein. In addition, the original analysis by MALDI mass spectrometry should be repeated to identify other possible sites of RNA binding.

In conclusion, UV crosslinking analysis of N protein and RNA identified two RNA-binding peptides. An additional peptide sequence possibly involved in N protein-RNA interaction was identified by mass spectrometry.

8 A Study of Ribonucleoprotein Structure

8.1 Introduction

The nucleocapsid protein of RSV is able to bind RNA in a non-sequence specific manner. Therefore, it is possible that His-N protein associates with bacterial RNA during expression and is able to form nucleocapsid-like structures, as reported for the expression of N proteins from other *Mononegavirales* in heterologous expression systems (Iseni *et al.*, 1998; Masters & Banerjee, 1988; Meric *et al.*, 1994, Spehner *et al.*, 1991). The occurrence of potential ribonucleoproteins (RNPs) in bacteria expressing His-N or mutant derivatives was investigated by caesium chloride (CsCl) gradients and electron microscopy (EM). The presence of RNA was investigated. Therefore, this study may give more insight to the domain, or domains on N responsible for binding RNA.

8.2 CsCl Gradient Isolation of Nucleocapsid-like Structures

Two buoyant density gradient types were employed. Preformed gradients with a CsCl range from 20% (w/w) to 40% (w/w) provided a relatively quick method of separation. Though sample equilibrium was not obtained, separation was sufficient to enable potential RNPs to be investigated by EM. CsCl buoyant density equilibrium gradients provided more refined separation allowing the components to better reach their equilibrium buoyant density. The samples analysed were the clarified lysates of cells expressing the N protein of interest. All bacterial lysates were digested with DNase and RNase A to reduce sample viscosity prior to centrifugation and subsequent density gradient separation (Sections 2.12.1 & 2.12.2). Retention of RNA following RNase treatment is taken as a defining feature of *Paramyxoviridae* nucleocapsids (Fearns *et al.*, 1997).

8.2.1 Buoyant Density Equilibrium Gradients

Initial density gradients were performed to compare the behaviour of RSV N protein prepared from different sources. These included clarified lysates of bacteria expressing histidine-tag N protein, Sf21 cells infected with recombinant baculovirus (Bac-N) expressing N protein (provided by P. Yeo) and CV-1 cells infected with RSV. Equilibrium gradients were fractionated and investigated by western blot (Figure 8.1). The blot with bacterially derived His-N protein (Figure 8.1 a) was probed with antibody

Figure 8.1: A western blot of N proteins separated by equilibrium gradient.

The clarified lysates of bacteria expressing Histidine-tag N protein (a), Sf21 cells infected with Bac-N expressing baculovirus-N (b) and CV-1 cells infected with RSV (c) were separated by step gradient. The RNP band was isolated from the 30% (w/w) CsCl region of the gradient and combined with additional 30% CsCl to form an equilibrium gradient. Gradients were fractionated and investigated by western blot. Histidine-tag N protein was probed with antibody specific to the His-tag (a). Baculovirus-N (b) and native N from RSV (c) were each probed with the polyclonal anti-N antibody αN015. Protein bands were visualised by ECL and autoradiography.

(a)
Density, g / cm³   1.2110   1.2387   1.2798 1.3045 1.3240
Fraction   2   4   6   8   10   12   14   16   18   20

(b)
Density, g / cm³   1.2055   1.2387   1.2858 1.3045 1.3240
Fraction   2   4   6   8   10   12   14   16   18   20

(c)
Density, g / cm³   1.2164   1.2446   1.2858   1.3110
Fraction   2   4   6   8   10   12   14   16   18   20
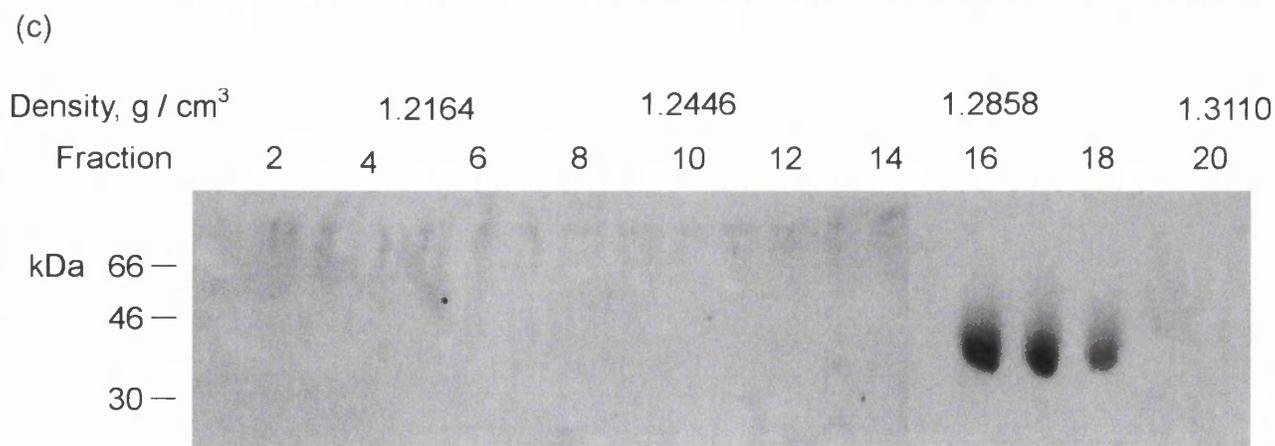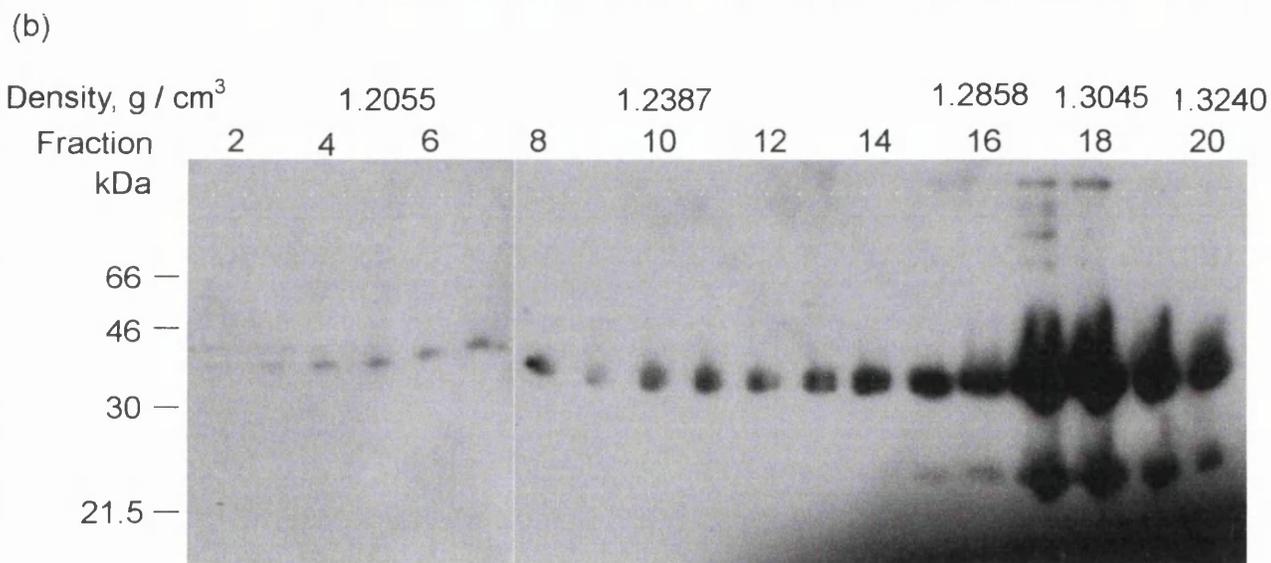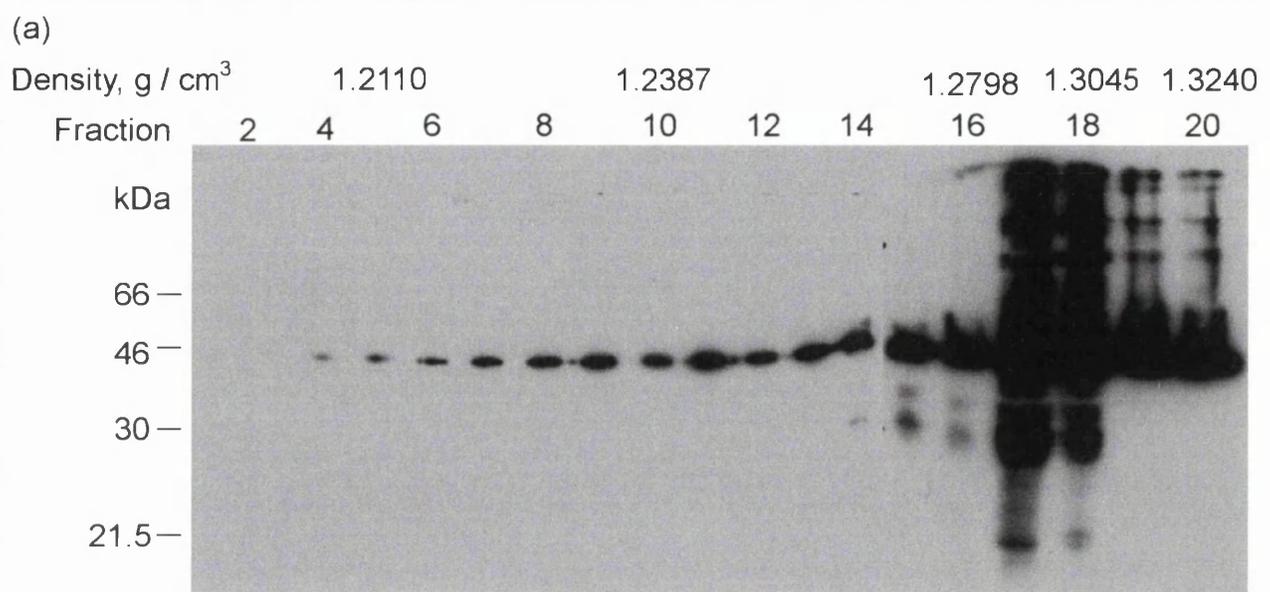
specific to the histidine-tag. Blots with baculovirus-N protein (Figure 8.1 b) and with RSV-N (Figure 8.1 c) were probed with the rabbit polyclonal antiserum, αN015, specific to the N protein. In each gradient, the density at which protein concentration peaked was approximately 1.29 g / cm^3. This is within the equilibrium density range of viral RNPs, buoyancy being a feature of RNA-bound protein (Ghosh & Ghosh, 1982; Masters & Banerjee; 1988, Myers *et al.*, 1999). Therefore, it is probable that, within these fractions, N protein was associated with RNA derived from a source other than RSV. The blots of recombinant proteins His-N and baculovirus-N were similar. Protein was at its most concentrated in fractions 17 and 18. Products of both oligomerisation and of protein breakdown were present in fractions with high protein concentration. In contrast to the recombinant proteins, RSV-N was detected in fractions 16 to 18 inclusive, and was present at considerably lower concentration as a single band of full-length protein when compared to the heterologously expressed N proteins. The absence of additional bands may be due to the lower N protein concentration applied to the gradient or perhaps greater stability of virus-derived protein. An equilibrium gradient of BSA, which does not bind RNA and therefore provides a negative control, was also performed and protein distribution analysed by SDS-PAGE stained with coomassie blue.

Densitometry analysis of the western blots in Figure 8.1 (a to c) demonstrated similar distribution of His-N and Bac-N in the equilibrium gradients, with a peak density of 1.3045 g / cm^3 common to both (Figure 8.2). The RSV-N gradient was slightly displaced from those obtained with the recombinant proteins. The RSV-N protein peak occurred at a density of approximately 1.29 g / cm^3. Despite subtle differences, the fraction density at each N protein peak was approximately 1.29 g / cm^3, again indicating that each protein was probably associated with RNA within the gradient. The slight differences may be the result of overloading and sample error.

Further equilibrium density gradients were performed to investigate the mutant N proteins described in Chapter 6. The peak protein densities of His-N and its mutant proteins occurred in a range from 1.2629 g / cm^3 (NΔ342-391) to 1.3023 g / cm^3 (His-N) (Figure 8.3). The highest densities were obtained with proteins His-N, NΔ121-160 and NΔ142-391. Peak protein fractions of gradients with NΔ1-200 and NΔ342-391 had the lowest densities, of 1.2659 g / cm^3 and 1.2629 g / cm^3, respectively. It is difficult to conclude whether these differences in peak protein fraction densities represent a
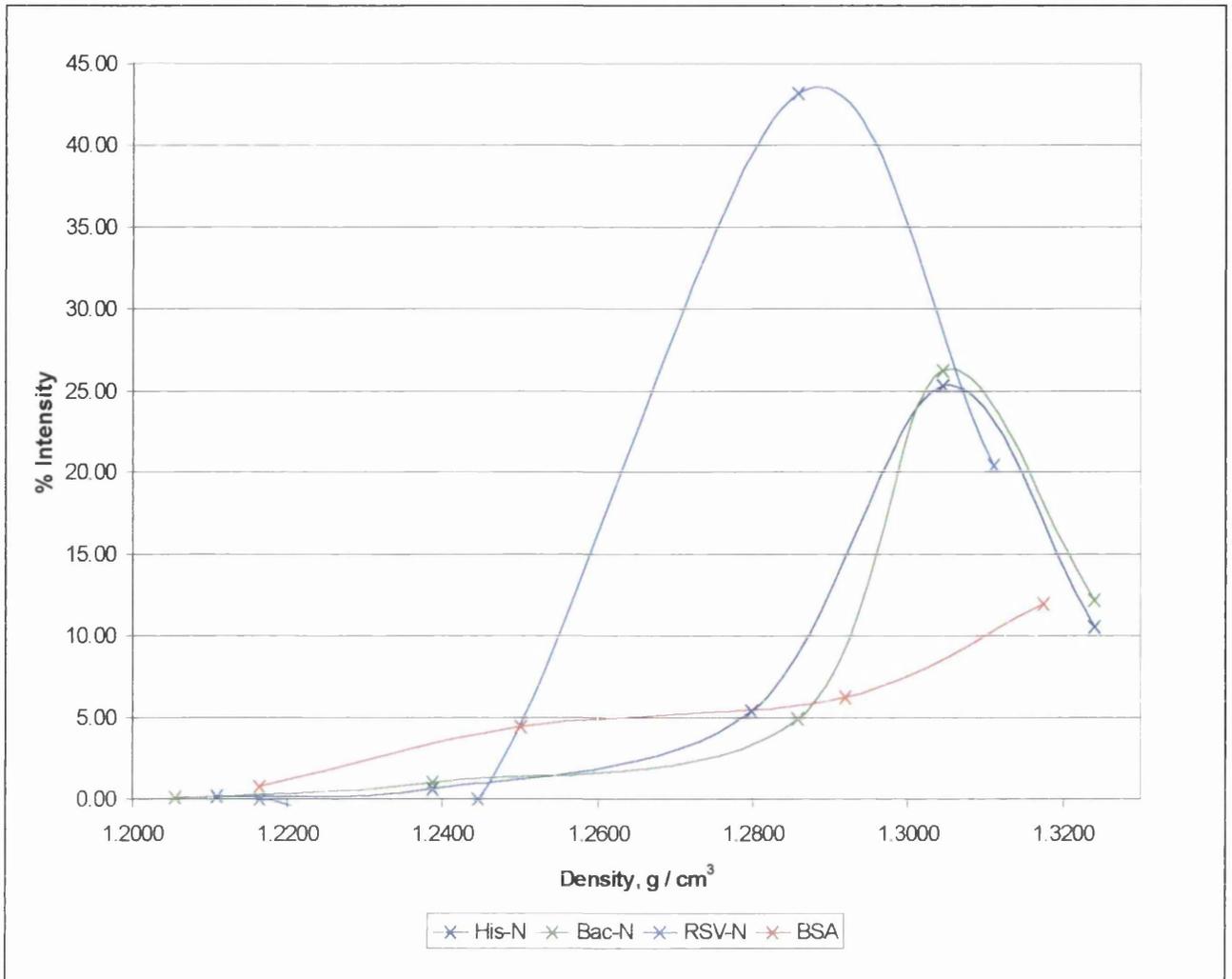
Figure 8.2: A graph showing the distribution of N proteins within buoyant density equilibrium gradients.

Traces represent native RSV-N and the recombinant proteins, histidine-tag N and baculovirus-N. The western blots of fractionated equilibrium gradients featured in Figure 8.1 were scanned and the band intensities quantified using densitometry software Quantity One (BioRad). The percentage of total protein present in individual fractions was calculated and plotted against fraction density. The traces of His-N and of baculovirus-N each display a peak in protein percentage at a density of approximately 1.3 g / cm^3. With RSV-N, this peak occurs at the lower density of approximately 1.28 g / cm^3.

Purified BSA provides a negative control.
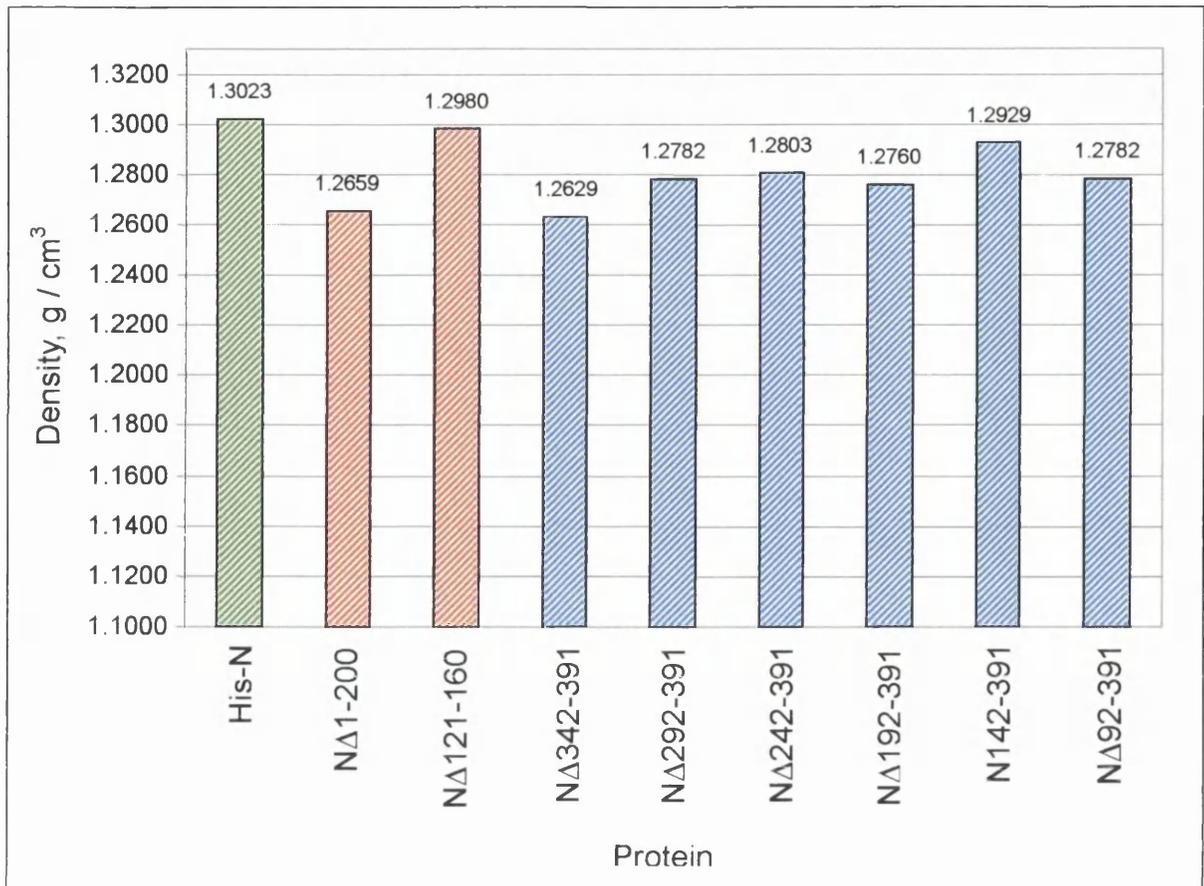
Figure 8.3: Density at which highest concentration of Histidine-tag N protein and its mutant derivatives occur in CsCl equilibrium gradients.

His-N and its mutant derivatives were separated by CsCl equilibrium gradient. Gradients were fractionated and investigated by western blot with antibody directed against the Histidine-tag. Protein bands were visualised by ECL and autoradiography. The refractive indices of gradient fractions containing the highest protein concentrations were measured and are presented in the above graph. His-N is presented in green and provides a positive control. Mutant proteins with amino-terminal deletions are presented in red. Carboxy-terminal deletion mutants are presented in blue. Peak densities range from 1.2629 g / cm^3 (NΔ342-391) to 1.3023 g / cm^3 (His-N).

difference in protein-RNA association of the mutants or result only from experimental variation. A negative control was omitted from the bar graph due to an even distribution of BSA through the gradient without a detectable peak in protein concentration (see Figure 8.2).

Differences in protein density of the His-N mutants showing significantly lower densities (NΔ1-200 and NΔ342-391) were investigated further by equilibrium gradients. The density of NΔ1-200 was compared with that of the other amino-terminus deletion mutant NΔ121-160, whose density was equivalent to His-N. Protein samples and gradients were prepared and performed concurrently. In Figure 8.4 (a), the peak in NΔ1-200 protein intensity occurred at fraction 12, considerably lower than that observed with NΔ121-160, which occurred at fraction 15. These gradients suggest that NΔ1-200 has a lower density than NΔ121-160, confirming previous observations (although in these experiments the density readings were not measured), and indicate that there may be a difference in protein-RNA binding behaviour. A similar pair-wise comparison was performed with NΔ342-391 and NΔ292-391 (Figure 8.4 b). Peaks in protein concentration occurred in the same fraction indicating that NΔ342-391 density is the same as NΔ292-391 and, presumably, as the other mutants, with the exception of NΔ1-200.

8.2.2 Preformed CsCl Gradients

Preformed density gradients (20 to 40% (w/w) CsCl) were performed with His-N and its mutant proteins. Gradients were fractionated and analysed by western blot (see Figure 8.5 for an example). In Figure 8.5, NΔ292-391 was first detected in fraction 12 and was present in all subsequent fractions. Protein concentration increased gradually until fraction 33 after which there was a dramatic increase in concentration until fraction 36. Protein concentration then decreased until the final fraction 40.

The density at which peak protein concentration occurred within preformed gradients was measured for each mutant (Figure 8.6). Variation in density occurred, with a range from 1.2561 g / cm^3 (NΔ121-160) to 1.2919 g / cm^3 (NΔ192-391). The fraction density obtained for His-N was 1.2668 g / cm^3, considerably lower than the value of 1.3023 g / cm^3 presented in Figure 8.3. There was notable variation between mutant densities within gradients. An explanation may be that the gradients were performed at different

Figure 8.4: Pair-wise comparisons of His-N mutant protein distribution within equilibrium density gradients.

His-N mutant proteins were separated by CsCl equilibrium gradient. Gradients were fractionated and protein distribution analysed by western blot with antibody directed against the histidine-tag. Protein band intensities were quantified using densitometry software Quantity One (BioRad). The percentage of total protein present in individual fractions was calculated and plotted against fraction number. Proteins NΔ121-160 and NΔ1-200 were analysed in (a). Proteins NΔ342-391 and NΔ292-391 were compared in (b). In (a), NΔ121-160 forms a tight protein peak in fraction 15. The NΔ1-200 protein peak occurs in fraction 12 and is relatively broad. This may suggest separation of distinct protein structures of different densities. A shoulder is present immediately before the peak in each trace, the significance of which is unknown. In (b), NΔ342-391 and NΔ292-391 share a similar distribution within each gradient suggesting a similarity in protein behaviour.

(a)
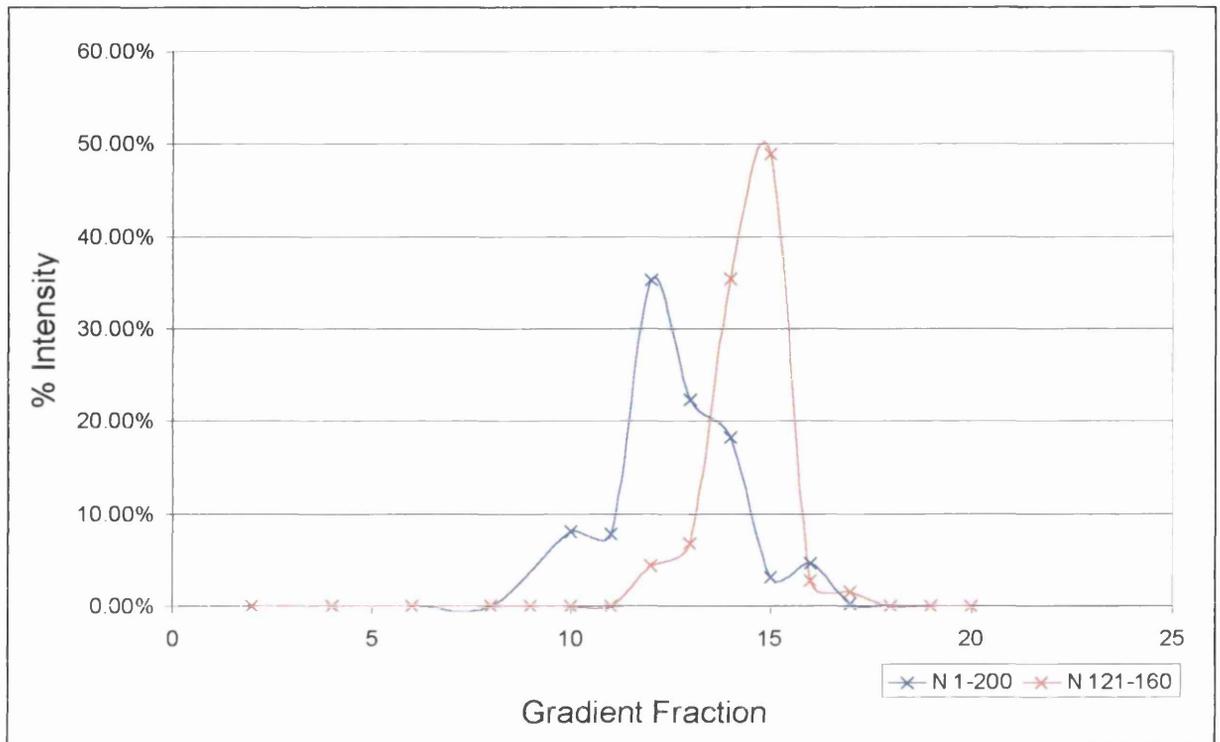
(b)
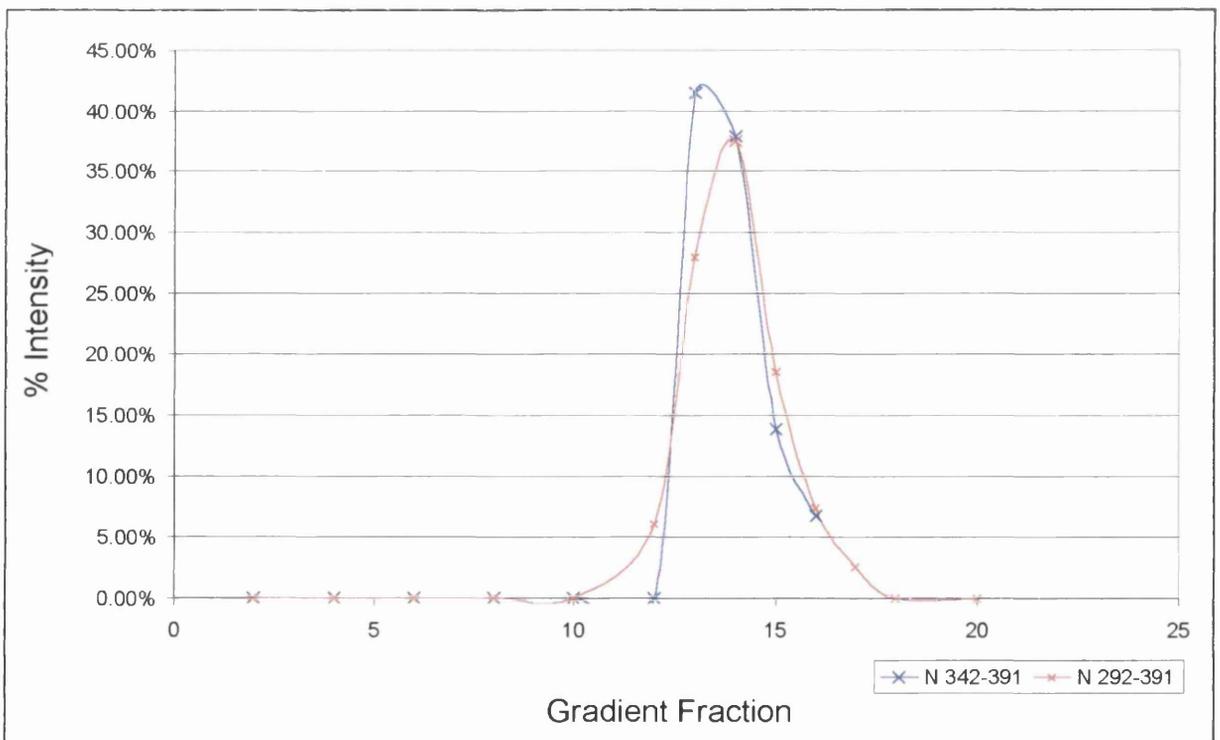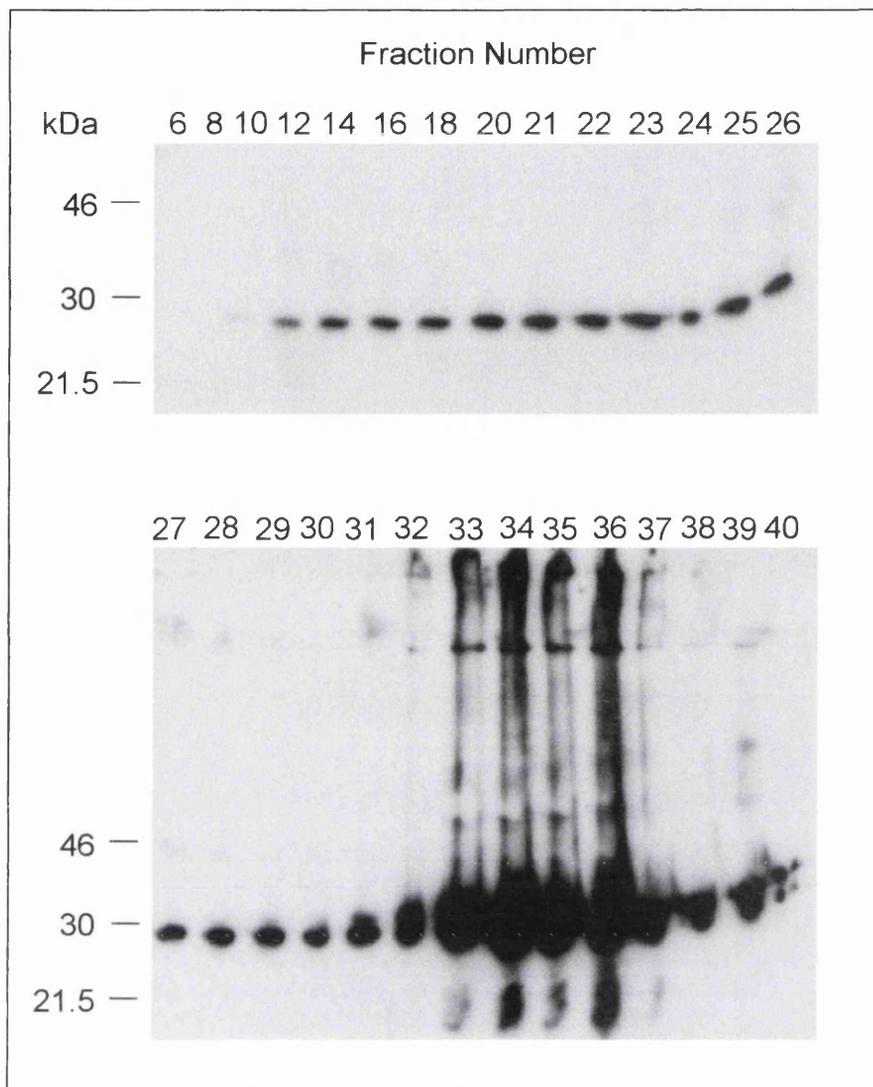
Figure 8.5: A western blot demonstrating the distribution of mutant protein NΔ292-391 within a CsCl preformed gradient.

The clarified lysate of bacteria expressing the mutant protein NΔ292-391 was applied to a 20 to 40% (w/w) preformed CsCl gradient and centrifuged. The gradient was fractionated and analysed by western blot with antibody directed against the Histidine-tag. Protein bands were visualised by ECL and autoradiography.
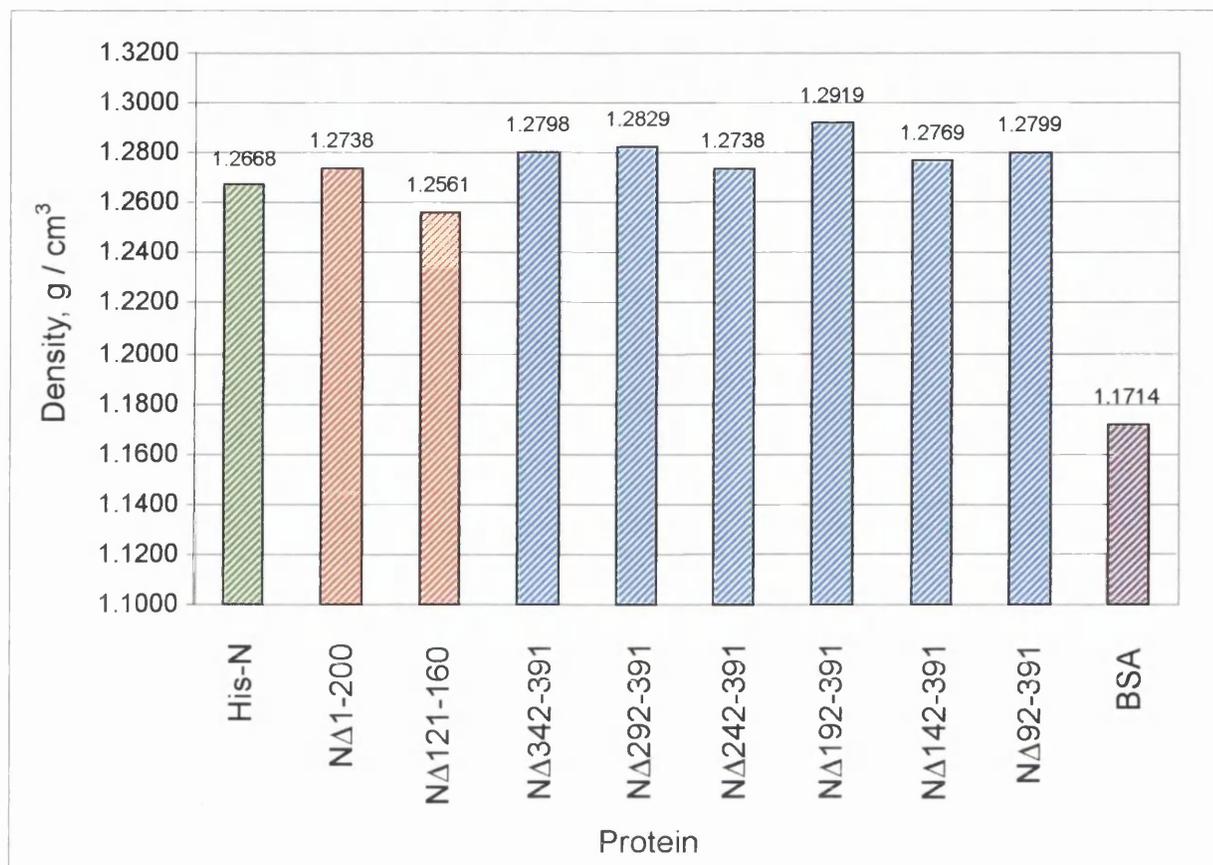
Figure 8.6: Density at which highest concentration of Histidine-tag N protein and its mutant derivatives occur in CsCl preformed gradients.

His-N and its mutant derivatives were separated by CsCl preformed gradient. Gradients were fractionated and investigated by western blot with antibody directed against the Histidine-tag. Protein bands were visualised by ECL and autoradiography. The refractive indices of gradient fractions containing the highest protein concentrations were measured and are presented in the above graph. His-N is presented in green and provides a positive control. A negative control is provided by BSA (purple). Mutant proteins with amino-terminal deletions are presented in red and carboxy-terminal deletion mutants are presented in blue. Peak densities range from 1.2561 g / cm^3 (NΔ121-160) to 1.2919 g / cm^3 (NΔ192-391).

times therefore introducing the possibility of differences in gradient preparation and conditions of centrifugation. For example, preformed gradients of NΔ1-200 and NΔ242-391 and of NΔ142-391 and NΔ92-391 were performed concurrently. Within each pair, density readings were found to be similar. In addition, centrifuge unreliability may have contributed to the variation in density, given its inclination to spontaneously stop. Upon comparison of the fraction densities at which His-N protein peaks occurred in the different gradient types, it appeared that protein analysed by preformed gradient did not reach equilibrium. However, separation by preformed gradient retained credibility upon comparison of mutant proteins with the negative control, BSA. Unlike buoyant equilibrium gradients, a peak in BSA protein concentration was observed in this gradient-type. There was a dramatic difference in fraction density between the mutants, potentially bound to RNA, and BSA that does not associate with RNA.

Experimental variation and lack of buoyant density equilibrium may provide an explanation for the peak densities obtained for mutants NΔ1-200 and NΔ121-160. By preformed gradient, NΔ121-160 density (1.2561 g / cm^3) was lower than that of NΔ1-200 (1.2738 g / cm^3), the inverse of results obtained by buoyant equilibrium gradient. Both values would be considered acceptable for RNA-bound protein when compared with the negative control. By equilibrium gradient, the density at which NΔ1-200 concentration was at its peak suggested a difference in protein behaviour compared with the other mutants (Figures 8.3 and 8.4), possibly an alteration of RNA-binding ability. However, this supposition becomes less clear upon consideration of the preformed gradients. However, for a number of factors the use of preformed gradients, at least in our hands, should not be used as a reliable indicator of density, as previously commented on by Masters & Banerjee, (1988). However, the preformed gradients provided a more convenient method for the isolation of potential RNPs and their analysis by EM.

8.3   Electron Microscopy of N Protein and its Mutants

N proteins from different origins (see Section 8.2.1) and His-N protein mutants were investigated by electron microscopy (EM). Aliquots of CsCl gradient fractions with densities of between 1.29 to 1.31 g / cm^3 were adsorbed on carbon-coated copper grids, stained with 1% (w/w) phosphotungstic acid and analysed by EM (Figure 8.7). Nucleocapsid structures were visible in grids created with RSV-derived N. Structures

Figure 8.7: Examples of structures observed by EM after isolation of N protein on CsCl gradients.

Histidine-tag N and its mutant derivatives were expressed in bacteria. Bac-N was expressed in Sf21 cells infected with recombinant baculovirus. Clarified cell lysates were separated by CsCl density gradient. The fractions containing high concentrations of N protein, identified by western blot, were investigated by EM.

A nucleocapsid obtained from RSV-infected CV-1 cells is present in the insert in Bac-N, demonstrating similarity in native and baculovirus-derived nucleocapsid form. N_{wt} refers to an untagged form of N, expressed and purified from bacteria (construction details not included). The mutant NΔ242-391 was difficult to purify and frequently contained a high amount of contaminating bacterial debris. A picture of NΔ242-391 nucleocapsid-like structures was therefore unavailable.

Bac-N

Nwt

His-N

β-Gal
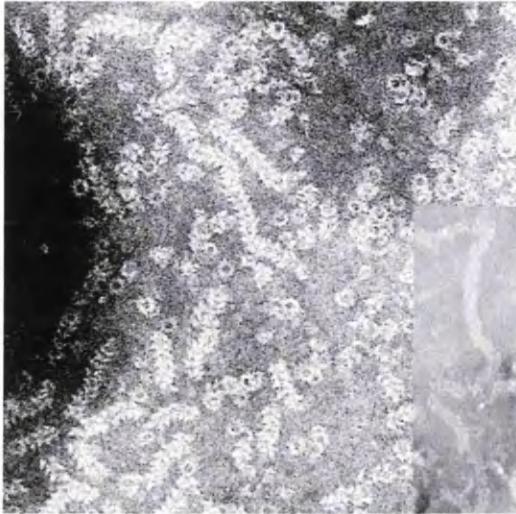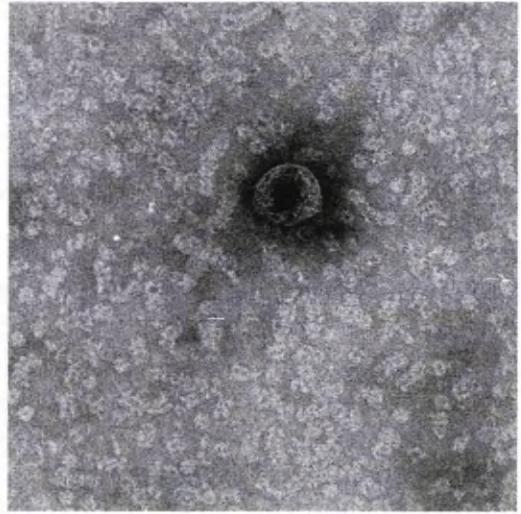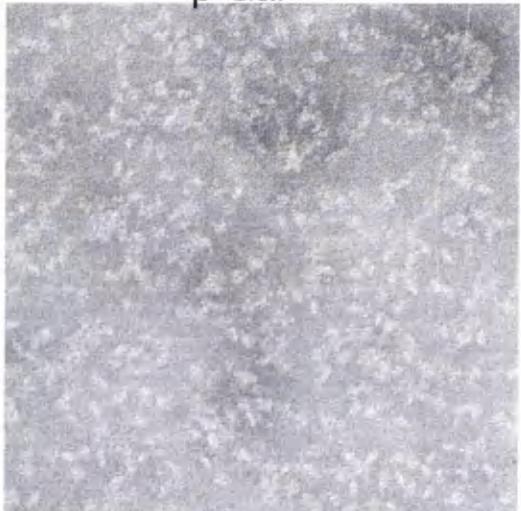
NΔ 342-391
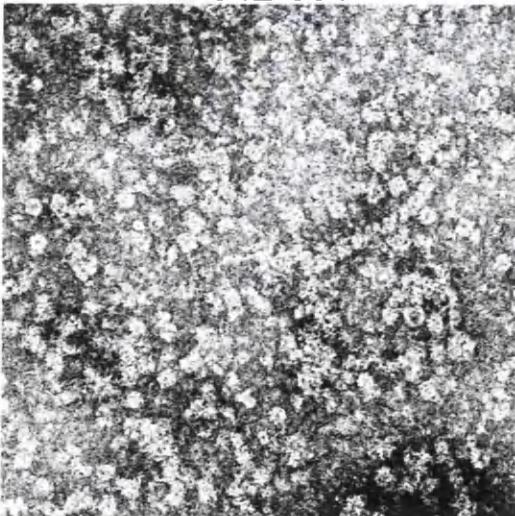
NΔ 292-391
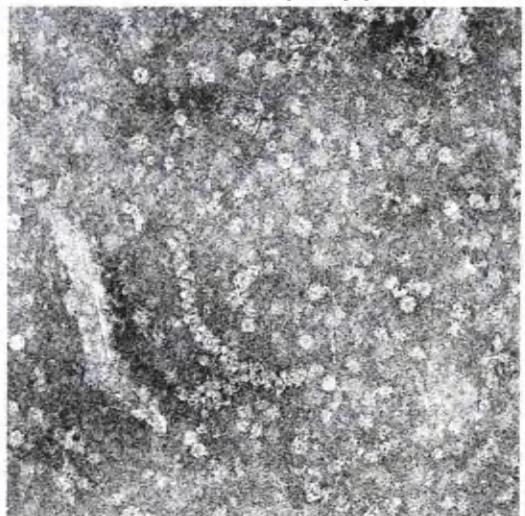
NΔ 192-391
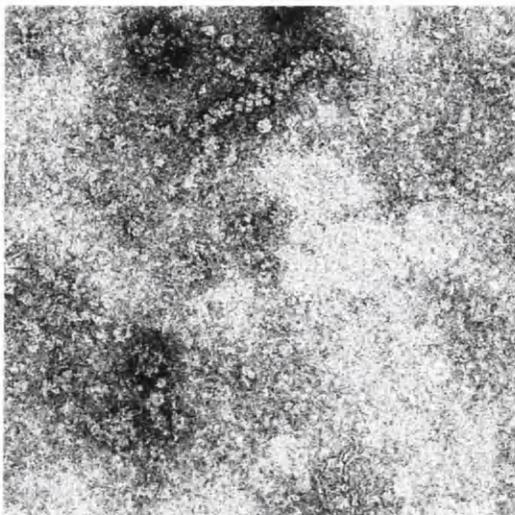
NΔ 142-391
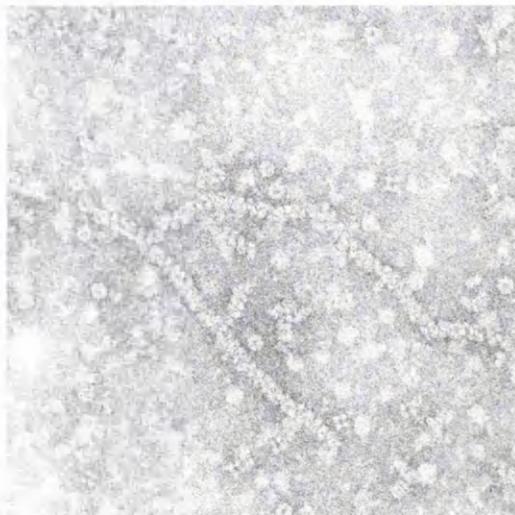
NΔ 92-391
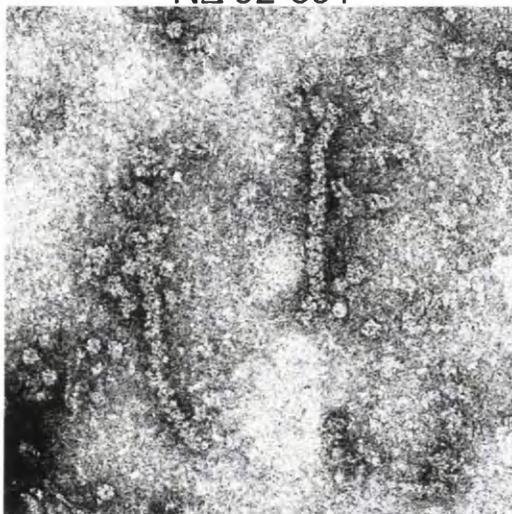
NΔ121-160
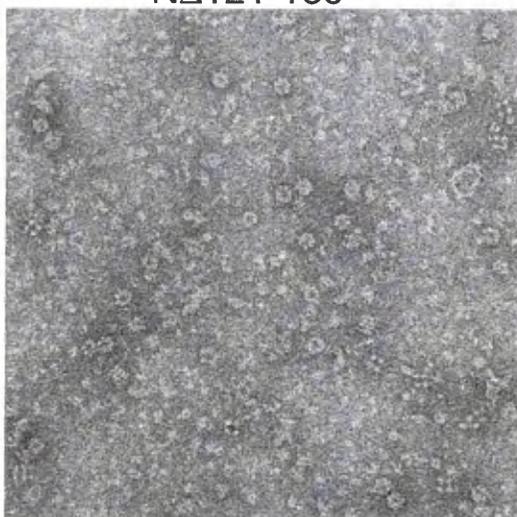
NΔ 1-200
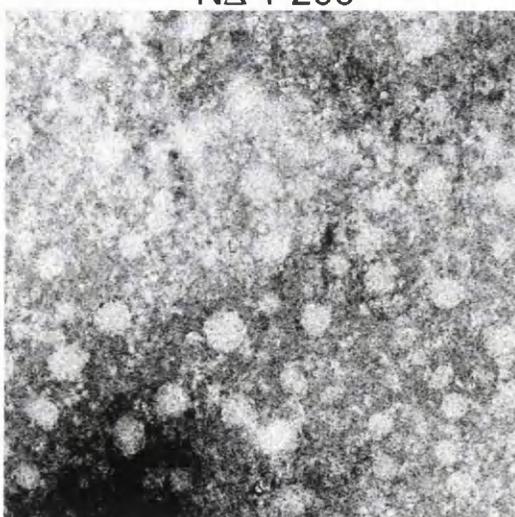
were long, slender and had a herringbone-like appearance, which suggested polarity of a supposedly helical structure. Nucleocapsid-like structures of similar appearance were observed after expression of RSV N by baculovirus, a phenomenon previously observed by Meric *et al.* (1994). In addition, ring structures of a similar width to the nucleocapsids were visible. Nucleocapsid-like structures were visible with His-N expressed in *E. coli*. However, the nucleocapsid-like structures formed were generally shorter than those obtained from RSV-infected cells or from recombinant baculovirus expression. They had the appearance of stacks of discs and lacked the polarity observed in non-bacterial systems. Ring structures were also present with the bacterially expressed protein.

A lysate of bacteria expressing His-tag β-galactosidase was used as negative control. The lysate was separated by buoyant density gradient and equivalent fractions at which His-N would be most concentrated in similar gradients were analysed by EM. Neither ring nor nucleocapsid-like structures were observed. This confirms that these structures were a feature of N protein expression and were neither native to bacteria nor a feature of histidine-tag proteins.

The N mutants were also investigated by EM. Ring and nucleocapsid-like structures were observed with each of the carboxy terminus deletion mutants (Figure 8.7 and Table 8.1). Ring abundance varied between mutant preparations. However, nucleocapsid-like structures created with each mutant protein had the shared appearance of stacks of discs lacking polarity, similar to those created with full-length His-N protein. The mutant NΔ242-391 was difficult to purify and frequently contained a high amount of contaminating bacterial debris. A picture of NΔ242-391 nucleocapsid-like structures was therefore unavailable. As previously observed (Figure 8.3), the carboxy deletion mutants migrate to a buoyant density of approximately 1.29 g / cm^3 upon equilibrium gradient centrifugation, indicating that they may be associated with RNA.

With NΔ121-160 only rings were apparent, no nucleocapsid-like structures were observed. EM investigation of NΔ1-200 preparations revealed an absence of ring and nucleocapsid-like structures. Only large protein aggregates were observed (Figure 8.7 and Table 8.1).

Protein	Rings	Nucleocapsid-like Structures
His-N	+	+
NΔ342-391	+	+
NΔ292-391	+	+
NΔ242-391	+	+
NΔ192-391	+	+
NΔ142-391	+	+
NΔ92-391	+	+
NΔ1-200	-	-
NΔ121-160	+	-

Table 8.1: Structures present in His-N mutant gradient samples, observed by EM. '+' indicates the presence and '-' indicates the absence of structure.

8.4 RNA Extraction from N Protein Buoyant Density Gradients

The protein densities and EM data, suggest the possible association of RNA with N proteins. N-derived proteins were isolated by immuno-precipitation from the appropriate fractions, using an antibody to the histidine-tag. Protein recovery was confirmed by western blot. RNA was extracted by phenol/chloroform, ethanol precipitation and subsequently end-labelled with ^{32}P; the products were analysed by denaturing polyacrylamide gel electrophoresis and visualised by phosphorimagery.

RNA was detected in association with His-N but not with NΔ1-200 (Figure 8.8 b). Each protein had been successfully immuno-precipitated as demonstrated by western blot (Figure 8.8 a). Discreet RNA bands of different sizes were present in the His-N protein lane. The different lengths may represent the ring and nucleocapsid-like structures of varying sizes observed by EM.

Immuno-precipitation and RNA extraction from the remaining N protein mutants was performed (Figure 8.9). By western blot, bands representing His-N, NΔ121-160, NΔ242-391, NΔ192-391, NΔ142-391 and NΔ92-391 were present (Figure 8.9 a). NΔ1-

Figure 8.8: Immuno-precipitation and RNA extraction assay performed on CsCl gradient fractions (i).

The clarified lysates of bacteria expressing His-N protein and NΔ1-200 were separated by CsCl density gradient. The gradients were fractionated and investigated by western blot with antibody directed against the Histidine-tag. Immunoprecipitation of His-N and NΔ1-200 was performed in gradient fractions containing highest protein concentrations. Precipitation was confirmed by western blot with antibody directed against the Histidine-tag (a). RNA was phenol: chloroform extracted from the precipitated protein, end-labelled with ^{32}P and investigated by 6% polyacrylamide gel electrophoresis (b). Radiolabelled bands were visualised by phosphorimagery. Negative controls include the immuno-precipitate from gradient fractions of untransformed bacterial lysate, (a) and (b), and protein A alone (b). Both His-N and NΔ1-200 proteins have been successfully immuno-precipitated. In (a), a band larger than 46 kDa is visible in the control and NΔ1-200 lanes, possibly created by capture of the detection antibody by Protein A, remaining from immuno-precipitation.
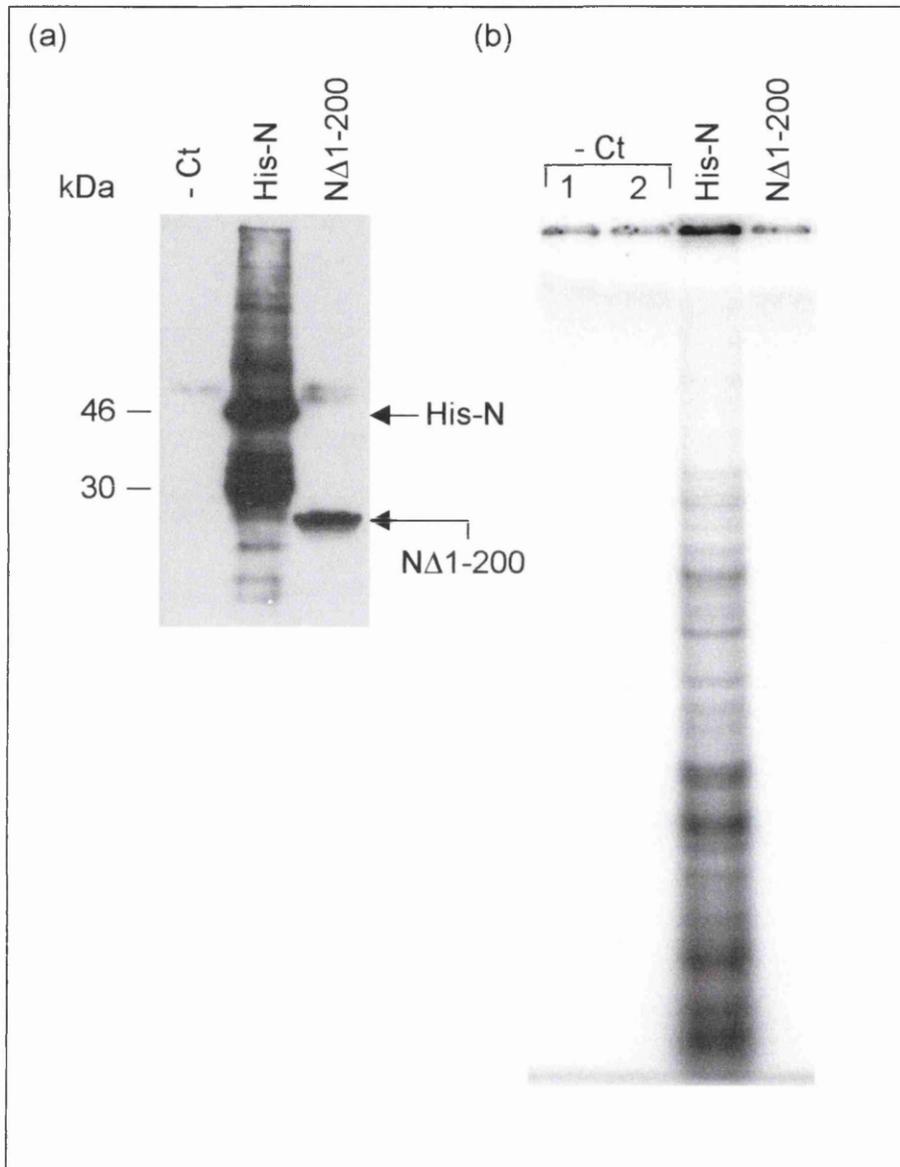
Figure 8.9: Immuno-precipitation and RNA extraction assay performed on CsCl gradient fractions (ii).

The clarified lysates of bacteria expressing His-N protein and its mutant derivatives were separated by CsCl density gradient. The gradients were fractionated and investigated by western blot with antibody directed against the Histidine-tag. Immunoprecipitation of His-N protein and its mutant derivatives was performed in gradient fractions containing highest protein concentrations. Precipitation was confirmed by western blot with antibody directed against the Histidine-tag (a). RNA was phenol: chloroform extracted from the precipitated protein, end-labelled with ^{32}P and investigated by 12% polyacrylamide gel electrophoresis (b). Radiolabelled bands were visualised by phosphorimagery. The lysate of bacteria expressing histidine-tag β-galactosidase separated by CsCl gradient provided a control in both (a) and (b). Lanes 1 to 8 present mutant protein samples as follows: 1 (NΔ1-200), 2 (NΔ121-160), 3 (NΔ342-391), 4 (NΔ292-391), 5 (NΔ242-391), 6 (NΔ192-391), 7 (NΔ142-391) and 8 (NΔ92-391).

(a)

kDa   β-Gal N   1   2   3   4   5   6   7   8

97.4 —
66 —
46 —
30 —
21.5 —
14.3 —

(b)
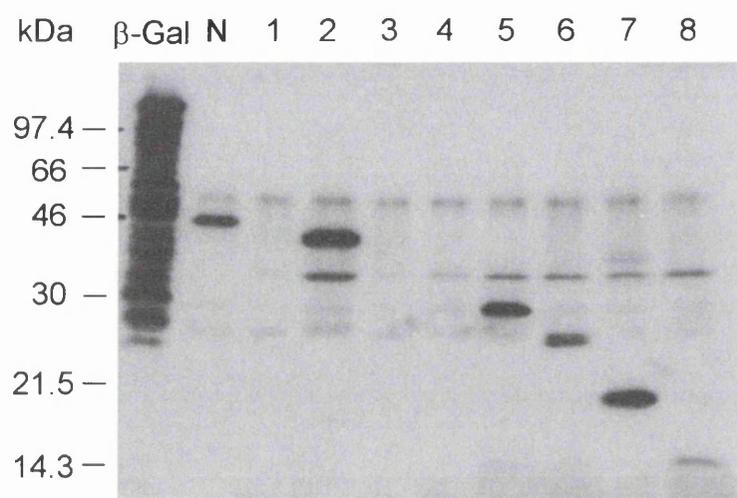
β-Gal N   1   2   3   4   5   6   7   8
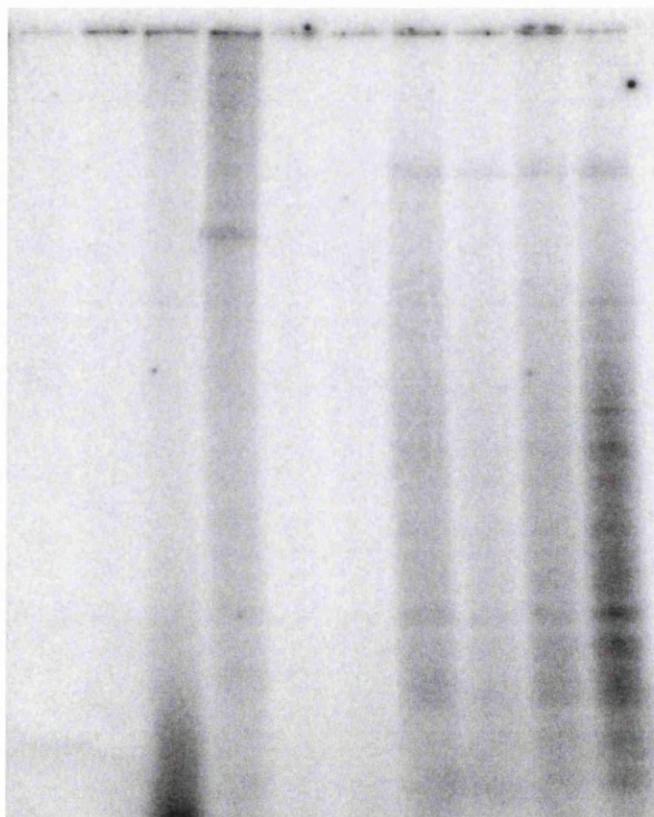
200, NΔ342-391 and NΔ292-391 were absent suggesting a failure of immuno-precipitation. A second smaller band of unknown origin is present in some lanes. However, the resulting RNA profiles of the different proteins were sufficiently dissimilar to exclude the possible influence of this unknown protein. His-tag β-galactosidase was used as negative control in this experiment. Upon examination of the phosphorimage (Figure 8.9 b), RNA was found to have been associated with NΔ242-391, NΔ192-391, NΔ142-391 and NΔ92-391. As observed in Figure 8.8 (b), there was variety in RNA length and band intensity. NΔ121-160 in Figure 8.8 (b) possessed a single RNA band present within the lane. This may represent the predominance of a single structure within NΔ121-160. Unfortunately, RNA was absent from the N protein lane suggesting failure of either RNA extraction or radiolabelling.

To clarify this data, the experiment was repeated (Figure 8.10). Upon this occasion, two samples each of His-N, NΔ1-200, NΔ121-160, NΔ342-391 and NΔ292-391 were investigated. Each protein was recovered by immunoprecipitation, with the exception of NΔ342-391 samples and one sample of NΔ292-391 had a low concentration of recovered protein (Figure 8.10 a). Radiolabelled RNA was found with His-N, NΔ121-160 and NΔ292-391, though band intensity was much reduced in the sample extracted from low protein concentration (Figure 8.10 b). The pattern of RNA bands was similar in each case suggesting that each protein associated with RNA in such a way as to provide greater stability to specific lengths. Therefore, the resulting protein-RNA structure may contain areas of weakness leading to the creation of varied lengths of RNA upon disruption. The intensity of RNA bands appears to reflect the concentration of protein recovered by immuno-precipitation. Several bands were observed with NΔ121-160 in contrast with the single band noted in Figure 8.9 (b). RNA was not isolated from lanes of NΔ1-200 and NΔ342-391. The latter protein was not recovered by immuno-precipitation therefore the absence of RNA was not unexpected. However, NΔ1-200 protein was precipitated therefore the lack of RNA confirmed the inability of NΔ1-200 to bind RNA.

8.5 Discussion

The analysis of full-length N proteins from different origins by equilibrium density gradient indicated that recombinant N behaved in a similar manner to that derived from RSV (Figure 8.1 and 8.2). Each protein migrated to a density similar to that of viral

Figure 8.10: Immuno-precipitation and RNA extraction assay performed on CsCl gradient fractions (iii).

The clarified lysates of bacteria expressing His-N protein, NΔ1-200, NΔ121-160, NΔ342-391 and NΔ292-391 were separated by CsCl density gradient. The gradients were fractionated and investigated by western blot with antibody directed against the Histidine-tag. Immunoprecipitation of His-N protein and its mutant derivatives was performed in gradient fractions containing highest protein concentrations. Precipitation was confirmed by western blot with antibody directed against the Histidine-tag (a). RNA was phenol: chloroform extracted from the precipitated protein, end-labelled with ^{32}P and investigated by 12% polyacrylamide gel electrophoresis (b). Radiolabelled bands were visualised by phosphorimagery.
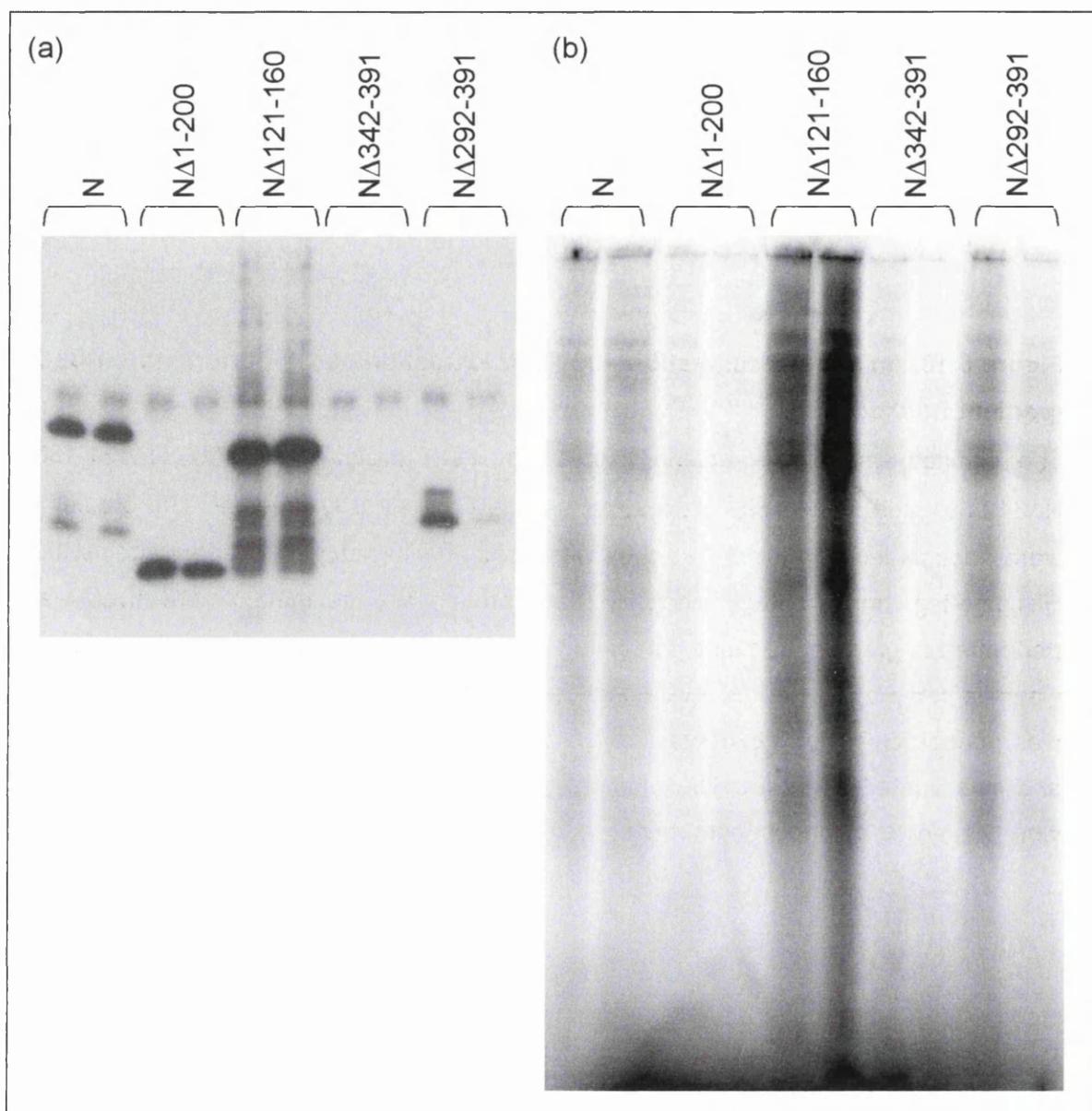
ribonucleoproteins (RNPs) (1.29 to 1.31 g / cm^3). The recombinant proteins, baculovirus-N and His-N, shared a common peak density of ca. 1.30 g / cm^3. RSV-derived N was present at a density of 1.29 g / cm^3. The slight differences in protein densities may arise from over-expression of recombinant protein and subsequent over-loading of the gradient in comparison with that derived from RSV. Sampling error may account for the differences observed. However, it is clear that N protein, expressed from heterologous sources, behaved in a similar manner to viral protein.

Mutant N proteins were investigated by equilibrium density gradient (Figure 8.3). Peak densities occurred in the range of 1.2629 g / cm^3 (NΔ342-391) to 1.3023 g / cm^3 (His-N). The range of densities may reflect subtle differences in protein-RNA association or structure, causing mutant proteins to peak at slightly different densities. Alternatively, experimental variation may offer an explanation. The lowest peak densities were obtained with proteins NΔ1-200 and NΔ342-391, which led to their further investigation (Figure 8.4). It was subsequently shown that NΔ342-391 density was identical to that of NΔ292-391, indicating a similarity in protein behaviour. The amino-terminal mutant proteins NΔ1-200 and NΔ121-160 were studied in parallel. Analysis confirmed that NΔ1-200 density was lower than that of NΔ121-160, and hence other mutants. This would suggest that NΔ1-200 does not form RNPs, that is, if we take density to reflect the association of protein with RNA.

Full-length and mutant N proteins were also investigated on preformed CsCl gradients. The mutant protein peak densities lay in a range from 1.2561 g / cm^3 (NΔ121-160) to 1.2919 g / cm^3 (NΔ192-391) (Figure 8.6). In general, values were lower than those obtained by equilibrium gradient. This may result from failure of preformed gradients to reach buoyant density equilibrium as mentioned previously.

Potential nucleocapsid-like structures were observed under the EM (Figure 8.7). Isolates with RSV-N and with baculovirus-N contained long, slender RNPs with a herringbone-like appearance. In addition, ring structures were obtained following baculovirus-N expression. It is postulated that the majority of the rings are created upon shearing of the larger nucleocapsid-like structures (P. Yeo, personal communication). Therefore, their presence in recombinant protein and absence from the RSV samples may suggest that structures created with the former are less stable. EM investigation of recombinant His-tag N protein samples identified the presence of

both nucleocapsid-like and ring structures. The nucleocapsid-like structures were shorter and lacked the polarity observed with N protein expressed in eukaryotic cells, having the appearance of a stack of discs. Structural appearance may reflect differences in conformation of the recombinant proteins expressed in bacteria compared with that expressed in eukaryotic cells.

Ring and nucleocapsid-like structures were observed for each of the carboxy-deletion N mutants. Structure abundance and nucleocapsid length varied between samples. However, their appearance was comparable to that of structures formed with full length His-N.

The lines of evidence above, that is, the similar densities to viral nucleocapsids and the formation of nucleocapsid-like structures, suggest that RNA was encapsidated by the recombinant proteins. RNA was recovered from His-N, NΔ121-160 and the carboxy terminal deletion mutants following protein isolation on CsCl gradients. RNA was not detected with NΔ1-200 and NΔ342-391, although for the latter, lack of protein expression or isolation means that no conclusive arguments can be made. Subsequent attempts have been made to analysis this construct, but expression has been problematical (C. Loney, personal communication). Thus, like other Mononegavirales, the RNA binding domain may be found in the amino-terminus of N. A significant proportion of the RNA binding domain exists within the first 100 amino acids of the N protein.

Differences were observed with the amino-terminal N mutants. Rings, but no nucleocapsid-like structures, were obtained with NΔ121-160. This perhaps suggests that the internal deletion affects either the ability of His-N to form more than one turn of the helix or that helices, if able to form, were unstable, resulting in the production of rings. RNA extracted from NΔ121-160 formed a single discreet band (Figure 8.9). It is possible that the single RNA species was obtained from the circular rings. However, RNAs of different lengths were isolated from a second gradient performed with NΔ121-160 (Figure 8.10). In the associated western blot, products of protein breakdown are visible. Therefore, it is possible that the additional RNA bands may result from RNA-binding by these protein fragments. NΔ1-200 samples lacked ordered structures, but contained numerous large, pleomorphic aggregates of an undefined nature. The unstructured appearance of NΔ1-200 and the absence of recovered RNA suggest that

the carboxyl-terminus of N has little, if any, involvement in the binding of RNA and the assembly of nucleocapsid-like structures.

The work presented in this chapter represents only preliminary data and requires repetition to confirm and further elucidate these observations. However, the use of a different protein expression system may be advisable. Samples prepared for EM were considered to be "dirty" due to contaminating bacterial proteins and probably not suitable for high resolution structural determination. This was also apparent when noted that the nucleocapsids obtained from bacterial cells have a different appearance to those generated in eukaryotic cells. N protein sequences are identical; the differences probably arise from a change in protein conformation that may occur during nucleocapsid assembly. Therefore, though bacteria provided a good host for preliminary studies, it may be advisable to graduate to a more advanced expression system for further analysis, such as the baculovirus system.

9 General Discussion

At the beginning of this study, RSV research had only recently been re-established within the Institute of Virology. Therefore, a considerable portion of the project was dedicated to generating reagents of use within the laboratory and developing suitable protocols for their investigation. Work presented in this thesis was performed concurrently with investigations into other aspects of RSV N protein biology such as a study on the association between N and cellular proteins. In addition, the interaction between viral N and P proteins and the domains responsible therein were investigated by P. Yeo, C. Loney and J. Murray. Investigation of the N protein-RNA interaction, and of the resulting macromolecular structures presented herein, complemented the ongoing work on the RSV nucleocapsid within the laboratory.

The main objectives achieved during the project may be summarised as follows:

- A system for the expression and purification of recombinant N protein (His-N) was developed, providing the core reagent for this study. The protein was subsequently used, within the laboratory, in experiments to map potential P-binding sites, generating data that has been published (Murray *et al.*, 2001).

- The RNA sequence specificity of N was investigated, employing a variety of *in vitro* techniques. The N demonstrated a lack of sequence specificity in its interaction with RNA.

- The location of the RNA-binding domain within N was investigated. Considering data from Chapters 6, 7 and 8, it was deduced that regions that contribute significantly to the binding and encapsidation of RNA are situated within the amino-terminal portion of the N protein.

- EM studies demonstrated the formation of ordered structures upon His-N protein expression in bacteria. Data presented herein, in conjunction with laboratory-acquired knowledge of the P binding sites and the panel of N-specific Mabs, may present a powerful tool in the further study of the nucleocapsid.

- The protocol for His-N purification has been refined to the extent that it may, with further development, enable full-length N protein to be investigated by high resolution structural studies such as X-ray crystallography.

9.1 Preliminary Studies and Expression of Recombinant N protein.

Before the aims of the project were addressed, virus growth and protein behaviour were studied by western blot, RIPA and IFA (Chapter 3). The expression of viral proteins of expected size was confirmed by western blot, employing a panel of antibodies directed against specific viral components. RIPA investigation of protein interactions occurring during RSV infection confirmed the interaction between N and P, an established association previously observed in human RSV (Garcia-Barreno et al., 1996; Slack & Easton, 1998) and in bovine RSV (Krishnamurthy & Samal, 1998). Possible interactions between N, M and M2-1 were observed, though the precise nature of the association requires further elucidation. The interesting nature of the association between N protein and MAb α-N15 was observed in the RIPA. In contrast with other α-N MAbs, α-N15 precipitated only a small amount of N and no P protein. Peptide mapping studies identified the antagonistic nature of the binding of MAb α-N15 and P to the N protein (Murray et al., 2001). It is possible that MAb α-N15 detects an epitope that is found only on nucleocapsids, further studies are in progress. Similar MAbs have been described for RV (Kawai et al., 1999).

The N protein ORF was successfully cloned into pET16b (Novagen) and expressed in E. coli strain BL21 (DE3) pLysS. Sufficient quantities of soluble protein were obtained to allow purification by nickel-affinity chromatography. Western blot analysis detected considerable amounts (> 70% of the total expressed) of recombinant protein within the insoluble fraction. Subsequent experiments suggest that a proportion of this may represent His-N assembled with bacterial RNA into nucleocapsid-like structures, a feature common to N proteins (Iseni et al., 1998; Spehner et al., 1991). Structures were pelleted with cell debris upon ultra-centrifugation (> 50,000 xg) of the lysate prior to isolation of the protein. Therefore, only soluble His-N, mostly in the form of a monomer, is purified. This initial level of His-N purity was sufficient to enable analysis of the sequence specificity of His-N:RNA binding. Further purification steps were necessary to prepare protein suitable for UV crosslinking and analysis by mass spectrometry. The level of His-N purity eventually obtained may conceivably enable higher resolution

structural analysis such as X-ray crystallography. However, problems presented by the inherent stickiness of the protein would need to be addressed to enable sample concentration to a level appropriate for such studies. One approach would be proteolytic digestion of His-N to obtain more soluble fragments. Thus, the data in Figure 7.8 may be helpful in elucidating which proteases to employ.

9.2 N-RNA Interactions

In vitro analysis demonstrated the ability of His-N to bind RNA in a non-sequence specific manner, a feature common to the nucleocapsid proteins of other Mononegaviruses, for example MV, VSV and SeV (Curran *et al.*, 1995; Masters & Banerjee, 1988; Spehner *et al.*, 1991). However, within the infected cell, only assembly upon viral genomic/antigenomic RNA is observed. Masters & Banerjee (1988) demonstrated that the specificity of interaction observed between VSV N and viral RNA was conferred by the P protein. P apparently maintains N in soluble form, preventing illegitimate assembly on non-viral RNA. It is believed that nucleocapsid assembly is permitted upon the recognition by the N:P complex of an encapsidation signal within the viral leader sequence. Similar P-directed prevention of non-specific association between N and cellular RNAs has been demonstrated for SeV and MV (Curran, 1996; Curran *et al.*, 1993; Spehner *et al.*, 1997). In SV5, this role may be provided by the V protein (Precious *et al.*, 1995). Yang *et al.* (1998) reported that the rabies virus N had a greater affinity for viral leader RNA than for that of non-viral origin. However, non-specific encapsidation was still observed.

The location of the RSV N protein domain responsible for binding RNA formed the core objective of this study. Though sequence homology among the paramyxovirus N proteins is relatively low, comparison of the amino-terminal 80% of each suggests reasonably conserved sequences; greater variation occurred in the carboxy-terminal region (Curran *et al.*, 1993). The amino-terminus of paramyxovirus N proteins are relatively basic and reasonably hydrophobic in nature. Combined, these characteristics suggest that the RNA-binding site may be present within the amino-terminal domain. The protein may be visualised as a globular amino-terminal core with an exposed carboxy-terminal tail, supported by the bilobed appearance of VSV N protein under the EM (Masters & Banerjee, 1988).

The N protein RNA-binding domain may be conformational in nature, that is, non-adjacent amino acids, rather than a single primary sequence, may contribute to the interaction. This may make identification of the region(s) responsible more complicated. From data presented in Chapters 5 through to 8, a potential RNA-binding region may be deduced. By north-western analysis, a band of approximately 30 kDa was detected in samples of purified His-N able to bind RNA (Figure 5.9). An amino-terminus His-N breakdown product of the same mass was frequently detected upon western blot analysis with antibody directed against the histidine-tag (for example, in Figure 4.6). Therefore, it is proposed that this places the RNA-binding domain in the amino-terminal two-thirds of N. Lack of involvement of the carboxy-tail has been observed for SeV. Mountcastle *et al.* (1974) demonstrated that trypsin treatment of SeV nucleocapsids removed the carboxy- tail of the N protein; nucleocapsid structure was left largely intact though the structure increased in rigidity. Despite their normal appearance and provision of resistance to RNase digestion, loss of the SeV N carboxy tail prevented the nucleocapsid from acting as a template for RNA synthesis. Subsequently, using the mini-genome system, Khattar *et al.* (2000) confirmed that the ability of bovine RSV N to encapsidate RNA and for the nucleocapsid to function as a template, are separate; the division of the amino- and carboxy-termini of the RSV N apparently mimic the SeV scenario.

It is proposed that the RNA-binding domain of RSV N is not present at the extreme amino terminus. This assumption is based on the NaOH protein fragment investigated by UV crosslinking and observed to bind RNA (Figure 7.5). Analysis of the fragment by western blot was performed with antibodies specific to the N protein and the histidine-tag. However, the fragment was unreactive with any of the MAbs tested. MAb α-N009 binds to an epitope within the region 16 to 35 aa, i.e. near the extreme end of the amino terminus of N, which is also the position of the histidine-tag. Therefore, the lack of detection with either antibody suggests that this region of the amino-terminus is absent from the NaOH fragment. Thus, the RNA-binding site may be located in the region from ca. 35 to 260 aa, as deduced from the above and from the estimated size of the amino-terminal 30 kDa fragment.

The locale of the N RNA-binding domain may be further refined by data obtained from the structural studies of His-N and its mutant derivatives. The largest RSV N deletion mutant, NΔ92-391, was still found in association with RNA. Therefore, the location of

the RNA-binding site may be found between amino acids 35 and 92. Deletions of amino acids 1 and 399 of the SeV N protein (524 aa) either abolished nucleocapsid assembly and presumably RNA-binding or caused reduction in protein function (Buchholz *et al.*, 1993). Deletions in the carboxy-terminal of the SeV N had no effect on nucleocapsid formation, a situation similar to that reported in this thesis. The ability of NΔ92-391 to bind RNA has now been confirmed on a number of occasions (C. Loney and K. MacLellan, personal communication).

The ability of His-N to bind RNA is unaffected by the histidine-tag, as demonstrated by UV crosslinking, north-western analysis and structural studies. If the hypothesised RNA-binding domain is correct, the site at the extreme amino-terminal of the protein, occupied by the His-tag, is not involved in the association with RNA and nucleocapsid assembly, a result supported by conclusions from Khattar *et al.*, (2000).

The presence of the RNA-binding site in the amino terminus of N protein was further supported by the failure to extract RNA from NΔ1-200 protein. This mutant migrated to a slightly lower peak density in CsCl equilibrium gradient compared with that of full-length and remaining His-N mutants. The difference in protein behaviour was mirrored by the results of RNA extraction. Examination by EM revealed the absence of ordered structures in NΔ1-200 samples and the presence of large protein aggregates. CsCl gradient analysis of SeV N and mutant derivatives indicated that proteins with a density lower than 1.27 g / cm^3 were devoid of RNA (Myers *et al.*, 1999). A similar low density was observed with the RSV mutant NΔ1-200. In addition, Myers *et al.* (1999) suggest that RNA-binding by the SeV N is achieved via hydrophobic interactions, rather than by acid − base association. In this present study, the isoelectric point of N was reduced in His-N but was still able to bind RNA.

Preliminary investigation, by mass spectrometry, identified a peptide within the carboxy terminal of His-N that was potentially modified by association with RNA. This peptide may represent another region of N involved in, or near to, the RNA-binding domain. If involved, it is only minimally responsible for binding RNA as absence from the carboxy deletion mutants apparently does not hinder association with RNA. Kouznetzoff *et al.* (1998) predicted an involvement of amino acids 298 to 352 in the RNA-binding region of rabies virus N (450 aa), i.e. three-quarters along the length of the protein. Conceivably, the folding of the N protein may bring regions of the amino- and carboxy-

termini into proximity which could explain the data. Recent work within the laboratory has produced data suggesting that a potential P binding site in the amino-terminus of the RSV N protein interacts with the carboxy-terminus region (C. Loney, unpublished results). Krishnamurthy & Samal (1998) confirmed the conformational nature of bovine RSV N by an investigation of the N-N interaction using the yeast two-hybrid system. Internal deletions were not tolerated from the protein resulting in the failure of self-association. As the deletions were generally non-overlapping, the major conclusion could be that N has a high degree of conformation which may be easily disrupted. Fortunately, this was not observed within this study. A change in the conformation of N may also occur upon the binding of RNA. Analysis by mass spectrometry detected novel peptides present after the binding of RNA to His-N (Table 7.1). Their occurrence may result from a change in protein conformation upon association with RNA and subsequent exposure of trypsin cleavage sites, hitherto protected. Conformational change has been demonstrated in the maturation of the nucleocapsids of MV and RV (Gombart *et al.*, 1993; Kawai *et al.*, 1999). It has been noted that the RSV N acquires a resistance to proteolytic digestion during the infection process, suggesting the occurrence of a maturation step (Cash *et al.*, 1979). Whether this can be related to the conformational change observed upon His-N:RNA interaction is intriguing.

9.3 Nucleocapsid Structure

The presence of RNA complexed to His-N within the CsCl gradients suggested the possibility that nucleocapsid-like structures may form within the bacteria. Meric *et al.* (1994) previously reported that, when expressed from recombinant baculovirus, RSV N spontaneously formed nucleocapsid-like structures, although whether they contained RNA was not investigated. Similarly, in the absence of other viral components, the N protein of MV has the ability to self-assemble into nucleocapsid-like particles, possibly by associating with cellular RNAs (Spehner *et al.*, 1997; Warnes *et al.*, 1995). Similar behaviour had been observed with SeV N, producing nucleocapsids identical in density and morphology to the authentic particles only smaller, possibly due to the assembly with shorter RNAs (Myers *et al.*, 1997). Recently N:RNA complexes have been reported for VSV and rabies virus after heterologous expression.

It was determined that all of the carboxy-terminal mutants and His-N formed nucleocapsid-like structures. The data within this thesis indicates that they are

associated with RNA. A second species of His-N:RNA complex was also observed, that is the occurrence of 'rings' which appear to represent one turn of the nucleocapsid. The nucleocapsid-like structures formed with His-N have a different appearance to those generated in eukaryotic cells, as observed by EM (Figure 8.7). His-N forms structures lacking polarity with the appearance of a stack of discs compared with the longer nucleocapsid of herringbone appearance observed with viral-N and Bac-N. This difference suggests that the bacterial cell environment in some way affects recombinant N protein. It perhaps causes altered conformation and thereby affects the protein-protein interaction potentially responsible for the change in conformation from a series of discs to structures with polarity. Thus, it looks like the final stage of nucleocapsid formation cannot occur in bacterial cells, whether this can be related to the discussion above about changes in N conformation, and nucleocapsid maturation, remains an enigma.

The amino terminus deletion mutant NΔ121-160, which bound RNA (Figure 8.9), produced only ring structures. Therefore, the absence of longer nucleocapsid-like structures is probably not a feature of disrupted RNA association but is more likely a feature of altered protein-protein binding. Interaction between N proteins on adjacent turns of the nucleocapsid helix may be disrupted in this mutant, increasing the fragility of the nucleocapsid and its propensity to shear into component rings (P. Yeo, personal communication).

Ring structures generated with His-N were isolated from CsCl gradients and investigated by computer image analysis (work performed by D. Bhella) (Figure 9.1). Rings were observed to contain 10 potential subunits, the same number as that observed in the rings isolated from nucleocapsid preparations of the rhabdoviruses VSV and rabies virus (Green *et al.*, 2000; Iseni *et al.*, 1998). In contrast, 13 subunits have been observed in the ring structures of SeV (Egelman *et al.*, 1989). Comparison of ring structure and N protein mass suggests that RSV is more similar to rabies and the rhabdoviruses than to the paramyxoviruses. Comparison with other members of the *Paramyxoviridae* has shown that RSV, and probably the subfamily *Pneumovirinae*, are structurally distinct from the *Paramyxovirinae* (P.Yeo, D. Bhella, personal communication). This represents an active area of research for the group. The NΔ121-160 construct may provide a source of ring structures, free of contaminating
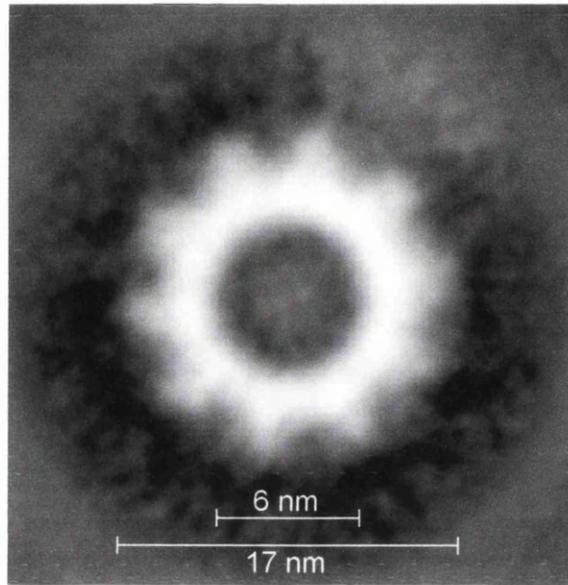
Figure 9.1: A ring structure isolated from fractions of a His-N CsCl gradient and visualised by computer image analysis of EM pictographs.

The ring is composed of 10 His-N monomers encapsulating RNA in an RNase A-resistant structure. The internal pore is 6 nm in diameter and the external ring structure is 17 nm in diameter.

nucleocapsid-like structures, for high level structural resolution which would enhance our knowledge of the nucleocapsid.

9.4   A Potential Map of the RSV N Protein

Data presented in this thesis, and generated in the laboratory by colleagues, have contributed to a proposed model of domain structure for N protein (Figure 9.2). As previously described, the deduced RNA-binding site has been located to the region 35 to 92 aa. The region 121 to 160 aa appears important in helical stability of the resulting nucleocapsid. Therefore, it is involved in an aspect of N-N interaction. P protein binding sites have been located to positions 46 to 65 and 301 to 335. (Murray *et al.*, 2001). An additional putative binding site for P is located from 211 to 230 (C. Loney, personal communication). Also found within the carboxy-terminus (amino acids 192-391) is a site for binding M (data not included, P. Yeo personal communication). There are obvious gaps in the map, for example amino acids 161 to 210, whose function has not been determined. These may simply represent hinge regions linking domains together or acting as regions of flexibility. Future work will try to identify other interacting regions, i.e. for M2-1, both in this linear model and in a three-dimensional spatial model based on reconstruction of the RSV nucleocapsid.

9.5   Conclusion

Work presented in this thesis has generated reagents, in the form of histidine-tagged N protein and deletion mutant derivatives, for use in the laboratory. Protocols for the purification of full-length recombinant protein have been established. In addition, protocols for UV crosslinking and mass spectrometry have been developed and optimised in anticipation of their further use. Future work will involve the transference of the mutants into the baculovirus system as one problematical aspect of this study has been the contamination of protein preparations with bacterial proteins. The baculovirus produces cleaner preparations (and shows that hindsight is a great advantage). Thus, the transfer of mutants to recombinant baculovirus and their subsequent expression may generate proteins, and hence nucleocapsids, of more realistic conformation. Mutants may then be used to further map the N protein domains. Purification of baculovirus-expressed N proteins may be performed to enable subsequent high-resolution analysis by NMR or by X-ray crystallography. The work presented in this thesis has made a number of observations that should form the basis of future work

Figure 9.2: Potential domains of the N protein of RSV.

The deduced RNA-binding site is present between amino acids 35 and 92. The N protein may be divided more broadly into an amino terminus domain (1 to 92 aa) able to bind RNA and assemble into nucleocapsid-like structures and a carboxy terminus (93 to 391 aa), dispensable for nucleocapsid assembly *in vitro*. The region 121 to 160 aa (purple) is possibly required for helical stability of the nucleocapsid. Therefore, it is potentially required for N-N interactions. Two P-binding sites (46 to 65 aa and 301 to 335 aa) are presented in blue. A putative binding site for P (211 to 230 aa) is also presented in blue and marked '*'. Antigenic sites to which MAbs α-N3, α-N9 and α-N15 bind are presented in yellow.
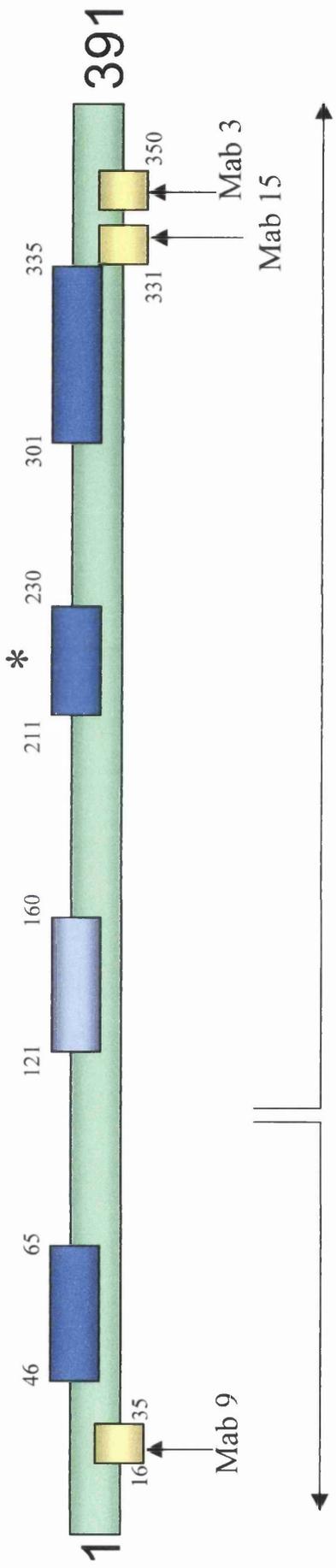
within the RSV group. It has initiated work on structural aspects of RSV that were not envisaged at the beginning of the study and hopefully will contribute to the understanding of RSV.

10 References

Ahmadian, G., Randhawa, J. S. & Easton, A. J. (2000). Expression of the ORF-2 protein of the human respiratory syncytial virus M2 gene is initiated by a ribosomal termination-dependent reinitiation mechanism. *Embo J* **19**, 2681-2689.

Alwan, W. H., Record, F. M. & Openshaw, P. J. (1992). CD4+ T cells clear virus but augment disease in mice infected with respiratory syncytial virus. Comparison with the effects of CD8+ T cells. *Clin Exp Immunol* **88**, 527-536.

Anderson, L. J., Hierholzer, J. C., Tsou, C., Hendry, R. M., Fernie, B. F., Stone, Y. & McIntosh, K. (1985). Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Infect Dis* **151**, 626-633.

Atreya, P. L. & Kulkarni, S. (1999). Respiratory syncytial virus strain A2 is resistant to the antiviral effects of type I interferons and human MxA. *Virology* **261**, 227-241.

Atreya, P. L., Peeples, M. E. & Collins, P. L. (1998). The NS1 protein of human respiratory syncytial virus is a potent inhibitor of minigenome transcription and RNA replication. *J Virol* **72**, 1452-1461.

Bachi, T. (1988). Direct observation of the budding and fusion of an enveloped virus by video microscopy of viable cells. *J Cell Biol* **107**, 1689-1695.

Bachi, T. & Howe, C. (1973). Morphogenesis and ultrastructure of respiratory syncytial virus. *J Virol* **12**, 1173-1180.

Behrendt, C. E., Decker, M. D., Burch, D. J. & Watson, P. H. (1998). International variation in the management of infants hospitalized with respiratory syncytial virus. International RSV Study Group. *Eur J Pediatr* **157**, 215-220.

Belshe, R. B., Van Voris, L. P. & Mufson, M. A. (1982). Parenteral administration of live respiratory syncytial virus vaccine: results of a field trial. *J Infect Dis* **145**, 311-319.

Bermingham, A. & Collins, P. L. (1999). The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc Natl Acad Sci U S A* **96**, 11259-11264.

Blumberg, B. M., Leppert, M. & Kolakofsky, D. (1981). Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* **23**, 837-845.

Bont, L., Heijnen, C. J., Kavelaars, A., van Aalderen, W. M., Brus, F., Draaisma, J. T., Geelen, S. M., van Vught, H. J. & Kimpen, J. L. (1999). Peripheral blood cytokine responses and disease severity in respiratory syncytial virus bronchiolitis. *Eur Respir J* **14**, 144-149.

Brandenburg, A. H., Neijens, H. J. & Osterhaus, A. D. (2001). Pathogenesis of RSV lower respiratory tract infection: implications for vaccine development. *Vaccine* **19**, 2769-2782.

Brandt, C., Power, U. F., Plotnicky-Gilquin, H., Huss, T., Nguyen, T., Lambert, P. H., Binz, H. & Siegrist, C. A. (1997). Protective immunity against respiratory syncytial virus in early life after murine maternal or neonatal vaccination with the recombinant G fusion protein BBG2Na. *J Infect Dis* **176**, 884-891.

Buchholz, C. J., Spehner, D., Drillien, R., Neubert, W. J. & Homann, H. E. (1993). The conserved N-terminal region of Sendai virus nucleocapsid protein NP is required for nucleocapsid assembly. *J Virol* **67**, 5803-5812.

Calain, P. & Roux, L. (1993). The rule of six, a basic feature for efficient replication of Sendai Virus defective interfering RNA. *J Virol* **67**, 4822-4830

Campbell, N. A. (1993). Biology, 3rd edn, pp. 850-872: The Benjamin/Cummings Publishing Company, Inc.

Cane, P. A., Matthews, D. A. & Pringle, C. R. (1991). Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *J Gen Virol* **72**, 2091-2096.

Cannon, M. J., Openshaw, P. J. & Askonas, B. A. (1988). Cytotoxic T cells clear virus but augment lung pathology in mice infected with respiratory syncytial virus. *J Exp Med* **168**, 1163-1168.

Caravokyri, C. & Pringle, C. R. (1992). Effect of changes in the nucleotide sequence of the P gene of respiratory syncytial virus on the electrophoretic mobility of the P protein. *Virus Genes* **6**, 53-62.

Carroll, A. R. & Wagner, R. R. (1979). Role of the membrane (M) protein in endogenous inhibition of in vitro transcription by vesicular stomatitis virus. *J Virol* **29**, 134-142.

Cash, P., Pringle, C. R. & Preston, C. M. (1979). The polypeptides of human respiratory syncytial virus: products of cell-free protein synthesis and post-translational modifications. *Virology* **92**, 375-384.

Chanock, R. M. & Finberg, L. (1957). Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiological aspects of infection in infants and young children. *American Journal of Hygiene* **66**, 291-300.

Chanock, R. M., Roizman, B. & Myers, R. (1957). Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent. I. Isolation, properties and characterisation. *American Journal of Hygiene* **66**, 281-290.

Cherrie, A. H., Anderson, K., Wertz, G. W. & Openshaw, P. J. (1992). Human cytotoxic T cells stimulated by antigen on dendritic cells recognize the N, SH, F, M, 22K, and 1b proteins of respiratory syncytial virus. *J Virol* **66**, 2102-2110.

Chin, J., Magoffin, R. L., Shearer, L. A., Schieble, J. H. & Lennette, E. H. (1969). Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. *Am J Epidemiol* **89**, 449-463.

Collins, P. L., Anderson, K., Langer, S. J. & Wertz, G. W. (1985). Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus. *Virology* **146**, 69-77.

Collins, P. L., Dickens, L. E., Buckler-White, A., Olmsted, R. A., Spriggs, M. K., Camargo, E. & Coelingh, K. V. (1986). Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order. *Proc Natl Acad Sci U S A* **83**, 4594-4598.

Collins, P. L., Hill, M. G., Camargo, E., Grosfeld, H., Chanock, R. M. & Murphy, B. R. (1995). Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development. *Proc Natl Acad Sci U S A* **92**, 11563-11567.

Collins, P. L., Hill, M. G., Cristina, J. & Grosfeld, H. (1996a). Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A* **93**, 81-85.

Collins, P. L., Huang, Y. T. & Wertz, G. W. (1984). Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus. *Proc Natl Acad Sci U S A* **81**, 7683-7687.

Collins, P. L., McIntosh, K. & Chanock, R. M. (1996b). Respiratory Syncytial Virus. In *Fields Virology*, 3rd edn, pp. 1313-1352. Edited by B. N. Fields, D. M. Knipe, P. M. Howley, R. M. Chanock, J. L. Melnick, T. P. Monath, B. Roizman & S. E. Straus. Philadelphia, Pa: Lippincott-Raven.

Collins, P. L., Mink, M. A., Hill, M. G., 3rd, Camargo, E., Grosfeld, H. & Stec, D. S. (1993). Rescue of a 7502-nucleotide (49.3% of full-length) synthetic analog of respiratory syncytial virus genomic RNA. *Virology* **195**, 252-256.

Collins, P. L., Mink, M. A. & Stec, D. S. (1991). Rescue of synthetic analogs of respiratory syncytial virus genomic RNA and effect of truncations and mutations on the expression of a foreign reporter gene. *Proc Natl Acad Sci U S A* **88**, 9663-9667.

Collins, P. L. & Mottet, G. (1993). Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *J Gen Virol* **74**, 1445-1450.

Collins, P. L., Olmsted, R. A., Spriggs, M. K., Johnson, P. R. & Buckler-White, A. J. (1987). Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc Natl Acad Sci U S A* **84**, 5134-5138.

Compans, R. W., Mountcastle, W. E. & Choppin, P. W. (1972). The sense of the helix of paramyxovirus nucleocapsids. *J Mol Biol* **65**, 167-169.

Connors, M., Collins, P. L., Firestone, C. Y. & Murphy, B. R. (1991). Respiratory syncytial virus (RSV) F, G, M2 (22K), and N proteins each induce resistance to RSV challenge, but resistance induced by M2 and N proteins is relatively short-lived. *J Virol* **65**, 1634-1637.

Connors, M., Collins, P. L., Firestone, C. Y., Sotnikov, A. V., Waitze, A., Davis, A. R., Hung, P. P., Chanock, R. M. & Murphy, B. R. (1992). Cotton rats previously immunized with a chimeric RSV FG glycoprotein develop enhanced pulmonary pathology when infected with RSV, a phenomenon not encountered following immunization with vaccinia–RSV recombinants or RSV. *Vaccine* **10**, 475-484.

Coronel, E. C., Murti, K. G., Takimoto, T. & Portner, A. (1999). Human parainfluenza virus type 1 matrix and nucleoprotein genes transiently expressed in mammalian cells induce the release of virus- like particles containing nucleocapsid-like structures. *J Virol* **73**, 7035-7038.

Coronel, E. C., Takimoto, T., Murti, K. G., Varich, N. & Portner, A. (2001). Nucleocapsid incorporation into parainfluenza virus is regulated by specific interaction with matrix protein. *J Virol* **75**, 1117-1123.

Crowe, J. E., Jr., Bui, P. T., Davis, A. R., Chanock, R. M. & Murphy, B. R. (1994a). A further attenuated derivative of a cold-passaged temperature- sensitive mutant of human respiratory syncytial virus retains immunogenicity and protective efficacy against wild-type challenge in seronegative chimpanzees. *Vaccine* **12**, 783-790.

Crowe, J. E., Jr., Bui, P. T., London, W. T., Davis, A. R., Hung, P. P., Chanock, R. M. & Murphy, B. R. (1994b). Satisfactorily attenuated and protective mutants derived from a partially attenuated cold-passaged respiratory syncytial virus mutant by introduction of additional attenuating mutations during chemical mutagenesis. *Vaccine* **12**, 691-699.

Curran, J. (1996). Reexamination of the Sendai virus P protein domains required for RNA synthesis: a possible supplemental role for the P protein. *Virology* **221**, 130-140.

Curran, J., Homann, H., Buchholz, C., Rochat, S., Neubert, W. & Kolakofsky, D. (1993). The hypervariable C-terminal tail of the Sendai paramyxovirus nucleocapsid protein is required for template function but not for RNA encapsidation. *J Virol* **67**, 4358-4364.

Curran, J. & Kolakofsky, D. (1988a). Ribosomal initiation from an ACG codon in the Sendai virus P/C mRNA. *Embo J* **7**, 245-251.

Curran, J. & Kolakofsky, D. (1988b). Scanning independent ribosomal initiation of the Sendai virus X protein. *Embo J* **7**, 2869-2874.

Curran, J., Marq, J. B. & Kolakofsky, D. (1995). An N-terminal domain of the Sendai paramyxovirus P protein acts as a chaperone for the NP protein during the nascent chain assembly step of genome replication. *J Virol* **69**, 849-855.

Curran, J. A., Richardson, C. & Kolakofsky, D. (1986). Ribosomal initiation at alternate AUGs on the Sendai virus P/C mRNA. *J Virol* **57**, 684-687.

Egelman, E. H., Wu, S. S., Amrein, M., Portner, A. & Murti, G. (1989). The Sendai virus nucleocapsid exists in at least four different helical states. *J Virol* **63**, 2233-2243.

Evans, J. E., Cane, P. A. & Pringle, C. R. (1996). Expression and characterisation of the NS1 and NS2 proteins of respiratory syncytial virus. *Virus Res* **43**, 155-161.

Falsey, A. R., Cunningham, C. K., Barker, W. H., Kouides, R. W., Yuen, J. B., Menegus, M., Weiner, L. B., Bonville, C. A. & Betts, R. F. (1995). Respiratory syncytial virus and influenza A infections in the hospitalized elderly. *J Infect Dis* **172**, 389-394.

Fearns, R. & Collins, P. L. (1999a). Model for polymerase access to the overlapped L gene of respiratory syncytial virus. *J Virol* **73**, 388-397.

Fearns, R. & Collins, P. L. (1999b). Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol* **73**, 5852-5864.

Fearns, R., Collins, P. L. & Peeples, M. E. (2000). Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus. *J Virol* **74**, 6006-6014.

Fearns, R., Peeples, M. E. & Collins, P. L. (1997). Increased expression of the N protein of respiratory syncytial virus stimulates minigenome replication but does not alter the balance between the synthesis of mRNA and antigenome. *Virology* **236**, 188-201.

Fernie, B. F. & Gerin, J. L. (1980). The stabilization and purification of respiratory syncytial virus using MgSO4. *Virology* **106**, 141-144.

Friedewald, W. T., Forsyth, B. R., Smith, C. B., Gharpure, M. A. & Chanock, R. M. (1968). Low-temperature-grown RS virus in adult volunteers. *Jama* **203**, 690-694.

Fulginiti, V. A., Eller, J. J., Sieber, O. F., Joyner, J. W., Minamitani, M. & Meiklejohn, G. (1969). Respiratory virus immunization. I. A field trial of two inactivated respiratory virus vaccines; an aqueous trivalent parainfluenza virus vaccine and an alum-precipitated respiratory syncytial virus vaccine. *Am J Epidemiol* **89**, 435-448.

Garcia-Barreno, B., Delgado, T. & Melero, J. A. (1996). Identification of protein regions involved in the interaction of human respiratory syncytial virus phosphoprotein and nucleoprotein: significance for nucleocapsid assembly and formation of cytoplasmic inclusions. *J Virol* **70**, 801-808.

Garcia-Barreno, B., Portela, A., Delgado, T., Lopez, J. A. & Melero, J. A. (1990). Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *Embo J* **9**, 4181-4187.

Garenne, M., Ronsmans, C. & Campbell, H. (1992). The magnitude of mortality from acute respiratory infections in children under 5 years in developing countries. *World Health Stat Q* **45**, 180-191.

Gharpure, M. A., Wright, P. F. & Chanock, R. M. (1969). Temperature-sensitive mutants of respiratory syncytial virus. *J Virol* **3**, 414-421.

Ghosh, K. & Ghosh, H. P. (1982). Synthesis in vitro of full length genomic RNA and assembly of the nucleocapsid of vesicular stomatitis virus in a coupled transcription-translation system. *Nucleic Acids Res* **10**, 6341-6351.

Gimenez, H. B., Hardman, N., Keir, H. M. & Cash, P. (1986). Antigenic variation between human respiratory syncytial virus isolates. *J Gen Virol* **67**, 863-870.

Glezen, W. P., Paredes, A., Allison, J. E., Taber, L. H. & Frank, A. L. (1981). Risk of respiratory syncytial virus infection for infants from low- income families in relationship to age, sex, ethnic group, and maternal antibody level. *J Pediatr* **98**, 708-715.

Glezen, W. P., Taber, L. H., Frank, A. L. & Kasel, J. A. (1986). Risk of primary infection and reinfection with respiratory syncytial virus. *Am J Dis Child* **140**, 543-546.

Gombart, A. F., Hirano, A. & Wong, T. C. (1993). Conformational maturation of measles virus nucleocapsid protein. *J Virol* **67**, 4133-4141.

Gonzalez-Reyes, L., Ruiz-Arguello, M. B., Garcia-Barreno, B., Calder, L., Lopez, J. A., Albar, J. P., Skehel, J. J., Wiley, D. C. & Melero, J. A. (2001). Cleavage of the human respiratory syncytial virus fusion protein at two distinct sites is required for activation of membrane fusion. *Proc Natl Acad Sci U S A* **98**, 9859-9864.

Graham, B. S., Henderson, G. S., Tang, Y. W., Lu, X., Neuzil, K. M. & Colley, D. G. (1993). Priming immunization determines T helper cytokine mRNA expression patterns in lungs of mice challenged with respiratory syncytial virus. *J Immunol* **151**, 2032-2040.

Green, T. J., Macpherson, S., Qiu, S., Lebowitz, J., Wertz, G. W. & Luo, M. (2000). Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. *J Virol* **74**, 9515-9524.

Grinnell, B. W. & Wagner, R. R. (1984). Nucleotide sequence and secondary structure of VSV leader RNA and homologous DNA involved in inhibition of DNA-dependent transcription. *Cell* **36**, 533-543.

Guerrero-Plata, A., Ortega, E. & Gomez, B. (2001). Persistence of respiratory syncytial virus in macrophages alters phagocytosis and pro-inflammatory cytokine production. *Viral Immunol* **14**, 19-30.

Hall, C. B. (2000). Nosocomial respiratory syncytial virus infections: the "Cold War" has not ended. *Clin Infect Dis* **31**, 590-596.

Hall, C. B. & Douglas, R. G., Jr. (1981). Modes of transmission of respiratory syncytial virus. *J Pediatr* **99**, 100-103.

Hall, C. B., Geiman, J. M., Douglas, R. G., Jr. & Meagher, M. P. (1978). Control of nosocomial respiratory syncytial viral infections. *Pediatrics* **62**, 728-732.

Hall, C. B., McBride, J. T. & Walsh, E. E. (1983). Aerosolised ribavirin treatment of infants with respiratory syncytial virus infection. *New England Journal of Medicine* **308**, 1442-1447.

Heggeness, M. H., Scheid, A. & Choppin, P. W. (1980). Conformation of the helical nucleocapsids of paramyxoviruses and vesicular stomatitis virus: reversible coiling and uncoiling induced by changes in salt concentration. *Proc Natl Acad Sci U S A* **77**, 2631-2635.

Heggeness, M. H., Scheid, A. & Choppin, P. W. (1981). the relationship of conformational changes in the Sendai virus nucleocapsid to proteolytic cleavage of the NP polypeptide. *Virology* **114**, 555-562.

Hendricks, D. A., McIntosh, K. & Patterson, J. L. (1988). Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. *J Virol* **62**, 2228-2233.

Hendry, R. M., Burns, J. C., Walsh, E. E., Graham, B. S., Wright, P. F., Hemming, V. G., Rodriguez, W. J., Kim, H. W., Prince, G. A., McIntosh, K. & et al. (1988). Strain-specific serum antibody responses in infants undergoing primary infection with respiratory syncytial virus. *J Infect Dis* **157**, 640-647.

Hengst, U. & Kiefer, P. (2000). Domains of human respiratory syncytial virus P protein essential for homodimerization and for binding to N and NS1 protein. *Virus Genes* **20**, 221-225.

Hoffman, M. A. & Banerjee, A. K. (2000). Analysis of RNA secondary structure in replication of human parainfluenza virus type 3. *Virology* **272**, 151-158.

Hruska, J. F., Morrow, P. E., Suffin, S. C. & Douglas, R. G., Jr. (1982). In vivo inhibition of respiratory syncytial virus by ribavirin. *Antimicrob Agents Chemother* **21**, 125-130.

Huang, Y. T., Collins, P. L. & Wertz, G. W. (1985). Characterization of the 10 proteins of human respiratory syncytial virus: identification of a fourth envelope-associated protein. *Virus Res* **2**, 157-173.

Hussell, T., Georgiou, A., Sparer, T. E., Matthews, S., Pala, P. & Openshaw, P. J. (1998). Host genetic determinants of vaccine-induced eosinophilia during respiratory syncytial virus infection. *J Immunol* **161**, 6215-6222.

Isaacs, D. (1991). Reducing hospital respiratory infections. *Nurs Times* **87**, 36.

Iseni, F., Barge, A., Baudin, F., Blondel, D. & Ruigrok, R. W. (1998). Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J Gen Virol* **79**, 2909-2919.

Johnson, K. M., Chanock, R. M., Rifkind, D., Kravetz, H. M. & Knight, V. (1961). Respiratory syncytial virus. IV. Correlation on virus shedding, serologic response and illness in adult volunteers. *JAMA* **176**, 663-667.

Johnson, P. R. & Collins, P. L. (1988). The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B: sequence conservation provides a structural basis for antigenic relatedness. *J Gen Virol* **69**, 2623-2628.

Johnson, P. R. & Collins, P. L. (1989). The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (RSV) of antigenic subgroups A and B: sequence conservation and divergence within RSV genomic RNA. *J Gen Virol* **70**, 1539-1547.

Johnson, P. R., Jr., Olmsted, R. A., Prince, G. A., Murphy, B. R., Alling, D. W., Walsh, E. E. & Collins, P. L. (1987a). Antigenic relatedness between glycoproteins of human respiratory syncytial virus subgroups A and B: evaluation of the contributions of F and G glycoproteins to immunity. *J Virol* **61**, 3163-3166.

Johnson, P. R., Spriggs, M. K., Olmsted, R. A. & Collins, P. L. (1987b). The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Natl Acad Sci U S A* **84**, 5625-5629.

Johnson, S., Oliver, C., Prince, G. A., Hemming, V. G., Pfarr, D. S., Wang, S. C., Dormitzer, M., O'Grady, J., Koenig, S., Tamura, J. K., Woods, R., Bansal, G., Couchenour, D., Tsao, E., Hall, W. C. & Young, J. F. (1997). Development of a humanized monoclonal antibody (MEDI-493) with potent in vitro and in vivo activity against respiratory syncytial virus. *J Infect Dis* **176**, 1215-1224.

Kakuk, T. J., Soike, K., Brideau, R. J., Zaya, R. M., Cole, S. L., Zhang, J. Y., Roberts, E. D., Wells, P. A. & Wathen, M. W. (1993). A human respiratory syncytial virus (RSV) primate model of enhanced pulmonary pathology induced with a formalin-inactivated RSV vaccine but not a recombinant FG subunit vaccine. *J Infect Dis* **167**, 553-561.

Kawai, A., Toriumi, H., Tochikura, T. S., Takahashi, T., Honda, Y. & Morimoto, K. (1999). Nucleocapsid formation and/or subsequent conformational change of rabies virus nucleoprotein (N) is a prerequisite step for acquiring the phosphatase-sensitive epitope of monoclonal antibody 5-2-26. *Virology* **263**, 395-407.

Kellner, J. D., Ohlsson, A., Gadomski, A. M. & Wang, E. L. (1996). Efficacy of bronchodilator therapy in bronchiolitis: a meta-analysis. *Archives of Paediatric and Adolescent Medicine* **150**, 1166-1172.

Khattar, S. K., Yunus, A. S., Collins, P. L. & Samal, S. K. (2000). Mutational analysis of the bovine respiratory syncytial virus nucleocapsid protein using a minigenome system: mutations that affect encapsidation, RNA synthesis, and interaction with the phosphoprotein. *Virology* **270**, 215-228.

Khattar, S. K., Yunus, A. S. & Samal, S. K. (2001). Mapping the domains on the phosphoprotein of bovine respiratory syncytial virus required for N-P and P-L interactions using a minigenome system. *J Gen Virol* **82**, 775-779.

Kim, H. W., Canchola, J. G., Brandt, C. D., Pyles, G., Chanock, R. M., Jensen, K. & Parrott, R. H. (1969). Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *Am J Epidemiol* **89**, 422-434.

Kim, H. W., Leikin, S. L., Arrobio, J. O., Brandt, C. D., Chanock, R. M. & Parrott, R. H. (1976). Cell-mediated immunity to respiratory syncytial virus induced by inactivated vaccine or by infection. *Paediatrics Research* **10**, 75-78.

Kimman, T. G. & Westenbrink, F. (1990). Immunity to human and bovine respiratory syncytial virus. *Arch Virol* **112**, 1-25.

Kouznetzoff, A., Buckle, M. & Tordo, N. (1998). Identification of a region of the rabies virus N protein involved in direct binding to the viral RNA. *J Gen Virol* **79**, 1005-1013.

Krishnamurthy, S. & Samal, S. K. (1998). Identification of regions of bovine respiratory syncytial virus N protein required for binding to P protein and self-assembly. *J Gen Virol* **79**, 1399-1403.

Krusat, T. & Streckert, H. J. (1997). Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch Virol* **142**, 1247-1254.

Kuo, L., Fearns, R. & Collins, P. L. (1996a). The structurally diverse intergenic regions of respiratory syncytial virus do not modulate sequential transcription by a dicistronic minigenome. *J Virol* **70**, 6143-6150.

Kuo, L., Grosfeld, H., Cristina, J., Hill, M. G. & Collins, P. L. (1996b). Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol* **70**, 6892-6901.

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

Lamb, R. A. & Kolakofsky, D. (1996). Paramyxoviridae: The viruses and their replication. In *Fields Virology*, 3rd edn, pp. 1177-1204. Edited by B. N. Fields, D. M. Knipe, P. M. Howley, R. M. Chanock, J. L. Melnick, T. P. Monath, B. Roizman & S. E. Straus. Philadelphia, Pa: Lippincott-Raven.

Law, B. J., Wang, E. E., MacDonald, N., McDonald, J., Dobson, S., Boucher, F., Langley, J., Robinson, J., Mitchell, I. & Stephens, D. (1997). Does ribavirin impact on the hospital course of children with respiratory syncytial virus (RSV) infection? An analysis using the pediatric investigators collaborative network on infections in Canada (PICNIC) RSV database. *Pediatrics* **99**, E7.

Lehmkuhl, H. D., Smith, M. H. & Cutlip, R. C. (1980). Morphogenesis and structure of caprine respiratory syncytial virus. *Arch Virol* **65**, 269-276.

Levine, S., Klaiber-Franco, R. & Paradiso, P. R. (1987). Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J Gen Virol* **68**, 2521-2524.

Masters, P. S. & Banerjee, A. K. (1988). Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. *J Virol* **62**, 2658-2664.

Mattaj, I. W. (1993). RNA recognition: a family matter? *Cell* **73**, 837-840.

McKay, E., Higgins, P., Tyrrell, D. & Pringle, C. (1988). Immunogenicity and pathogenicity of temperature-sensitive modified respiratory syncytial virus in adult volunteers. *J Med Virol* **25**, 411-421.

Melero, J. A., Garcia-Barreno, B., Martinez, I., Pringle, C. R. & Cane, P. A. (1997). Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol* **78**, 2411-2418.

Menon, K., Sutcliffe, T. & Klassen, T. P. (1995). A randomised trial comparing the efficacy of epinephrine with salbutamol in the treatment of acute bronchiolitis. *Journal of Paediatrics* **126**, 1004-1007.

Meric, C., Spehner, D. & Mazarin, V. (1994). Respiratory syncytial virus nucleocapsid protein (N) expressed in insect cells forms nucleocapsid-like structures. *Virus Res* **31**, 187-201.

Mink, M. A., Stec, D. S. & Collins, P. L. (1991). Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* **185**, 615-624.

Morgan, E. M. (1991). The Paramyxoviruses: Evoulutionary relationships of Paramyxovirus Nucleocapsid-associated proteins. In *The Viruses*, pp. 163-176. Edited by D. W. Kingsbury. New York: Plenum Press.

Morris, J. A., Blunt, R., E. & Savage, R., E. (1956). Recovery of cytopathogenic agent from chimpanzees with coryza. *Proceedings of the Society of Experimental Biology and Medicine.* **92**, 544-550.

Mountcastle, W. E., Compans, R. W., Lackland, H. & Choppin, P. W. (1974). Proteolytic cleavage of subunits of the nucleocapsid of the paramyxovirus simian virus 5. *J Virol* **14**, 1253-1261.

Mufson, M. A., Belshe, R. B., Orvell, C. & Norrby, E. (1987). Subgroup characteristics of respiratory syncytial virus strains recovered from children with two consecutive infections. *J Clin Microbiol* **25**, 1535-1539.

Mufson, M. A., Orvell, C., Rafnar, B. & Norrby, E. (1985). Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* **66**, 2111-2124.

Murphy, S. K. & Parks, G. D. (1997). Genome nucleotide lengths that are divisible by six are not essential but enhance replication of defective interfering RNAs of the paramyxovirus simian virus 5. *Virology* **232**, 145-157.

Murray, J., Loney, C., Murphy, L. B., Graham, S. & Yeo, R. P. (2001). Characterization of Monoclonal Antibodies Raised against Recombinant Respiratory Syncytial Virus Nucleocapsid (N) Protein: Identification of a Region in the Carboxy Terminus of N Involved in the Interaction with P Protein. *Virology* **289**, 252-261.

Myers, T. M., Pieters, A. & Moyer, S. A. (1997). A highly conserved region of the Sendai virus nucleocapsid protein contributes to the NP-NP binding domain. *Virology* **229**, 322-335.

Myers, T. M., Smallwood, S. & Moyer, S. A. (1999). Identification of nucleocapsid protein residues required for Sendai virus nucleocapsid formation and genome replication. *J Gen Virol* **80**, 1383-1391.

Ogra, P. L., Ogra, S. S. & Coppola, P. R. (1975). Secretory component and sudden infant-death syndrome. *Lancet* **2**, 387-390.

Olmsted, R. A., Elango, N., Prince, G. A., Murphy, B. R., Johnson, P. R., Moss, B., Chanock, R. M. & Collins, P. L. (1986). Expression of the F glycoprotein of respiratory syncytial virus by a recombinant vaccinia virus: comparison of the individual contributions of the F and G glycoproteins to host immunity. *Proc Natl Acad Sci U S A* **83**, 7462-7466.

Openshaw, P. J., Culley, F. J. & Olszewska, W. (2001). Immunopathogenesis of vaccine-enhanced RSV disease. *Vaccine* **20 Suppl 1**, S27-31.

Pastey, M. K. & Samal, S. K. (1997). Analysis of bovine respiratory syncytial virus envelope glycoproteins in cell fusion. *J Gen Virol* **78**, 1885-1889.

Peeples, M. E. & Collins, P. L. (2000). Mutations in the 5' trailer region of a respiratory syncytial virus minigenome which limit RNA replication to one step. *J Virol* **74**, 146-155.

Piazza, F. M., Johnson, S. A., Darnell, M. E., Porter, D. D., Hemming, V. G. & Prince, G. A. (1993). Bovine respiratory syncytial virus protects cotton rats against human respiratory syncytial virus infection. *J Virol* **67**, 1503-1510.

Portner, A. & Murti, K. G. (1986). Localization of P, NP, and M proteins on Sendai virus nucleocapsid using immunogold labeling. *Virology* **150**, 469-478.

Portner, A., Murti, K. G., Morgan, E. M. & Kingsbury, D. W. (1988). Antibodies against Sendai virus L protein: distribution of the protein in nucleocapsids revealed by immunoelectron microscopy. *Virology* **163**, 236-239.

Precious, B., Young, D. F., Bermingham, A., Fearns, R., Ryan, M. & Randall, R. E. (1995). Inducible expression of the P, V, and NP genes of the paramyxovirus simian virus 5 in cell lines and an examination of NP-P and NP-V interactions. *J Virol* **69**, 8001-8010.

Prince, G. A., Hemming, V. G., Horswood, R. L., Baron, P. A. & Chanock, R. M. (1987). Effectiveness of topically administered neutralizing antibodies in experimental immunotherapy of respiratory syncytial virus infection in cotton rats. *J Virol* **61**, 1851-1854.

Pringle, C. (1991). The Paramyxoviruses: The genetics of Paramyxoviruses. In *The Viruses*, pp. 1-32. Edited by D. W. Kingsbury. New York: Plenum Press.

Pringle, C. R. (1997). The order Mononegavirales–current status. *Arch Virol* **142**, 2321-2326.

Pringle, C. R., Shirodaria, P. V., Gimenez, H. B. & Levine, S. (1981). Antigen and polypeptide synthesis by temperature-sensitive mutants of respiratory syncytial virus. *J Gen Virol* **54**, 173-183.

Radecke, F., Spielhofer, P., Schneider, H., Kaelin, K., Huber, M., Dotsch, C., Christiansen, G. & Billeter, M. A. (1995). Rescue of measles viruses from cloned DNA. *Embo J* **14**, 5773-5784.

Randhawa, J. S., Marriott, A. C., Pringle, C. R. & Easton, A. J. (1997). Rescue of synthetic minireplicons establishes the absence of the NS1 and NS2 genes from avian pneumovirus. *J Virol* **71**, 9849-9854.

Re, G. G. (1991). The Paramyxoviruses: Deletion mutants of Paramyxoviruses. In *The Viruses*, pp. 275-298. Edited by D. W. Kingsbury. New York: Plenum Press.

Roberts, S. R., Compans, R. W. & Wertz, G. W. (1995). Respiratory syncytial virus matures at the apical surfaces of polarized epithelial cells. *J Virol* **69**, 2667-2673.

Rodriguez, W. J., Gruber, W. C., Groothuis, J. R., Simoes, E. A., Rosas, A. J., Lepow, M., Kramer, A. & Hemming, V. (1997). Respiratory syncytial virus immune globulin treatment of RSV lower respiratory tract infection in previously healthy children. *Pediatrics* **100**, 937-942.

Samal, S. K. & Collins, P. L. (1996). RNA replication by a respiratory syncytial virus RNA analog does not obey the rule of six and retains a nonviral trinucleotide extension at the leader end. *J Virol* **70**, 5075-5082.

Samal, S. K., Pastey, M. K., McPhillips, T. H. & Mohanty, S. B. (1993). Bovine respiratory syncytial virus nucleocapsid protein expressed in insect cells specifically interacts with the phosphoprotein and the M2 protein. *Virology* **193**, 470-473.

Samal, S. K., Zamora, M., McPhillips, T. H. & Mohanty, S. B. (1991). Molecular cloning and sequence analysis of bovine respiratory syncytial virus mRNA encoding the major nucleocapsid protein. *Virology* **180**, 453-456.

Sanderson, C. M., McQueen, N. L. & Nayak, D. P. (1993). Sendai virus assembly: M protein binds to viral glycoproteins in transit through the secretory pathway. *J Virol* **67**, 651-663.

Schlender, J., Bossert, B., Buchholz, U. & Conzelmann, K. K. (2000). Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize alpha/beta interferon-induced antiviral response. *J Virol* **74**, 8234-8242.

SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. & Wickens, M. (1996). A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci U S A* **93**, 8496-8501.

Severson, W., Partin, L., Schmaljohn, C. S. & Jonsson, C. B. (1999). Characterization of the Hantaan nucleocapsid protein-ribonucleic acid interaction. *J Biol Chem* **274**, 33732-33739.

Sigurs, N. (2001). Epidemiologic and clinical evidence of a respiratory syncytial virus-reactive airway disease link. *Am J Respir Crit Care Med* **163**, S2-6.

Simoes, E. A. (1999). Respiratory syncytial virus infection. *Lancet* **354**, 847-852.

Simoes, E. A. (2001). Treatment and prevention of respiratory syncytial virus lower respiratory tract infection. Long-term effects on respiratory outcomes. *Am J Respir Crit Care Med* **163**, S14-17.

Slack, M. S. & Easton, A. J. (1998). Characterization of the interaction of the human respiratory syncytial virus phosphoprotein and nucleocapsid protein using the two-hybrid system. *Virus Res* **55**, 167-176.

Spehner, D., Drillien, R. & Howley, P. M. (1997). The assembly of the measles virus nucleoprotein into nucleocapsid-like particles is modulated by the phosphoprotein. *Virology* **232**, 260-268.

Spehner, D., Kirn, A. & Drillien, R. (1991). Assembly of nucleocapsidlike structures in animal cells infected with a vaccinia virus recombinant encoding the measles virus nucleoprotein. *J Virol* **65**, 6296-6300.

Spriggs, M. K., Olmsted, R. A., Venkatesan, S., Coligan, J. E. & Collins, P. L. (1986). Fusion glycoprotein of human parainfluenza virus type 3: nucleotide sequence of the gene, direct identification of the cleavage-activation site, and comparison with other paramyxoviruses [published erratum appears in Virology 1987 May;158(1):263]. *Virology* **152**, 241-251.

Stec, D. S., Hill, M. G. d. & Collins, P. L. (1991). Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* **183**, 273-287.

Sugrue, R. J., Brown, C., Brown, G., Aitken, J. & Mc, L. R. H. W. (2001). Furin cleavage of the respiratory syncytial virus fusion protein is not a requirement for its transport to the surface of virus-infected cells. *J Gen Virol* **82**, 1375-1386.

Teng, M. N. & Collins, P. L. (1998). Identification of the respiratory syncytial virus proteins required for formation and passage of helper-dependent infectious particles. *J Virol* **72**, 5707-5716.

Teng, M. N., Whitehead, S. S., Bermingham, A., St Claire, M., Elkins, W. R., Murphy, B. R. & Collins, P. L. (2000). Recombinant respiratory syncytial virus that does not express the NS1 or M2-2 protein is highly attenuated and immunogenic in chimpanzees. *J Virol* **74**, 9317-9321.

Treuhaft, M. W. & Beem, M. O. (1982). Defective interfering particles of respiratory syncytial virus. *Infect Immun* **37**, 439-444.

van den Hoogen, B. G., de Jong, J. C., Goen, J., Kuiken, T., de Groot, R., Fouchier, R. A. M. & Osterhaus, A. D. M. E. (2001). A novel Pneumovirus isolated from children with respiratory tract disease. *Nature Medicine* **6**, 719-724.

van Schaik, S. M., Obot, N., Enhorning, G., Hintz, K., Gross, K., Hancock, G. E., Stack, A. M. & Welliver, R. C. (2000). Role of interferon gamma in the pathogenesis of primary respiratory syncytial virus infection in BALB/c mice. *J Med Virol* **62**, 257-266.

Vidal, S., Curran, J. & Kolakofsky, D. (1990). Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *J Virol* **64**, 239-246.

Vulliemoz, D. & Roux, L. (2001). "Rule of six": how does the Sendai virus RNA polymerase keep count? *J Virol* **75**, 4506-4518.

Walsh, E. E., Hall, C. B., Briselli, M., Brandriss, M. W. & Schlesinger, J. J. (1987). Immunization with glycoprotein subunits of respiratory syncytial virus to protect cotton rats against viral infection. *J Infect Dis* **155**, 1198-1204.

Warnes, A., Fooks, A. R., Dowsett, A. B., Wilkinson, G. W. & Stephenson, J. R. (1995). Expression of the measles virus nucleoprotein gene in Escherichia coli and assembly of nucleocapsid-like structures. *Gene* **160**, 173-178.

Wathen, M. W., Kakuk, T. J., Brideau, R. J., Hausknecht, E. C., Cole, S. L. & Zaya, R. M. (1991). Vaccination of cotton rats with a chimeric FG glycoprotein of human respiratory syncytial virus induces minimal pulmonary pathology on challenge. *J Infect Dis* **163**, 477-482.

Wertz, G. W., Collins, P. L., Huang, Y., Gruber, C., Levine, S. & Ball, L. A. (1985). Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc Natl Acad Sci U S A* **82**, 4075-4079.

Whitehead, S. S., Bukreyev, A., Teng, M. N., Firestone, C. Y., St Claire, M., Elkins, W. R., Collins, P. L. & Murphy, B. R. (1999). Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in chimpanzees. *J Virol* **73**, 3438-3442.

Wong, T. C. & Hirano, A. (1987). Structure and function of bicistronic RNA encoding the phosphoprotein and matrix protein of measles virus. *J Virol* **61**, 584-589.

Yang, J., Hooper, D. C., Wunner, W. H., Koprowski, H., Dietzschold, B. & Fu, Z. F. (1998). The specificity of rabies virus RNA encapsidation by nucleoprotein. *Virology* **242**, 107-117.

Yao, Q. & Compans, R. W. (2000). Filamentous particle formation by human parainfluenza virus type 2. *J Gen Virol* **81 Pt 5**, 1305-1312.

Young, D. F., Didcock, L., Goodbourn, S. & Randall, R. E. (2000). Paramyxoviridae use distinct virus-specific mechanisms to circumvent the interferon response. *Virology* **269**, 383-390.

Zimmer, G., Budz, L. & Herrler, G. (2001). Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. *J Biol Chem* **276**, 31642-31650.

11 Appendices

11.1 Buffer Reagents

11.1.1 Tissue Culture

PBS

0.17 M	Sodium Chloride
3.35 mM	Potassium Chloride
1.84 mM	Potassium Dihydrogen Phosphate
10 mM	Sodium Dihydrogen Phosphate

Adjusted to pH 7.4

11.1.2 Bacterial Culture

2YT Growth Medium

5 g	Sodium Chloride
16 g	Bactotryptone
10 g	Yeast Extract

1litre dH$_2$O, sterilised by autoclaving

SOC Growth Medium

1 mM	Sodium Chloride
0.25 mM	Potassium Chloride
2 mM	Magnesium Chloride
2 mM	Glucose

1 litre dH$_2$O, sterilised by autoclaving

11.1.3 Yeast Culture

Lithium Acetate/TE Buffer

10 mM	Tris Base
1 mM	EDTA
0.1 M	Lithium Chloride

Adjusted to pH 7.6

Z-Buffer

60 mM	Disodium Hydrophosphate
46 mM	Sodium Dihydrogen Phosphate
10 mM	Potassium Chloride
2 mM	Magnesium Sulphate

Z-Buffer/X-Gal Solution

100 ml	Z-Buffer
0.27 ml	β-Mercaptoethanol
1.67 ml	X-Gal Solution (40 mg ml^{-1} Dimethyl Formamide)

11.1.4 DNA Manipulation

1x TAE Buffer

40 mM	Tris-Base
2 mM	EDTA
0.1% (v/v)	Glacial Acetic Acid

6x DNA Loading Buffer

0.25% (w/v)	Bromophenol Blue
0.25% (v/v)	Xylene Cyanol FF
30% (v/v)	Glycerol

In 1x TAE buffer

10x Ligation Buffer

250 mM	Tris-HCl pH 7.5
50 mM	Magnesium Chloride
25% (w/v)	PEG 8000
5 mM	DTT
4 mM	ATP

11.1.5 Protein Manipulation

10 mM Imidazole Phosphate Buffer

10 mM	Imidazole
10 mM	Disodium Hydrophosphate
10 mM	Sodium Dihydrogen Phosphate
500 mM	Sodium Chloride

Adjusted to pH 7.4, sterilised by filtration

Tris-Glycine Acrylamide Gel

10% Acrylamide Separating Gel		4% Acrylamide Stacking Gel	
3.3 ml	30% Acrylamide *	1.3 ml	30% Acrylamide *
2.6 ml	ProtoGel Running Buffer (1.5 M Tris-HCl, 0.384% SDS, pH 8.8)	2.5 ml	ProtoGel Stacking Buffer (0.5 M Tris-HCl, 0.4% SDS, pH 6.8)
4.1 ml	dH$_2$O	6.2 ml	dH$_2$O
200 µl	10% (w/v) APS	200 µl	10% (w/v) APS
15 µl	TEMED	15 µl	TEMED

* UltraPure ProtoGel: 30% (w/v) Acrylamide: 0.8% (w/v) Bis-Acrylamide Stock Solution (37.5:1) (National Diagnostics)

3x Sample Loading Buffer

10 ml	ProtoGel Stacking Buffer
8 ml	25% (w/v) SDS
5 ml	β-Mercaptoethanol
10 ml	Glycerol
0.25% (w/v)	Bromophenol Blue

Tank Buffer

50 mM	Tris Base
50 mM	Glycine
3.5 mM	SDS

Tris-Tricine Acrylamide Gel

16% Acrylamide Separating Gel		4% Acrylamide Stacking Gel	
16 ml	30% Acrylamide	1.62 ml	30% Acrylamide
10 ml	Tris-HCl/SDS, pH 8.45	3.10 ml	Tris-HCl/SDS, pH 8.45
	(3 M Tris Base; 10 mM SDS)		(3 M Tris Base; 10 mM SDS)
0.83 ml	dH$_2$O	7.78 ml	dH$_2$O
3.17 ml	Glycerol	200 µl	10% (w/v) APS
200 µl	10% (w/v) APS	15 µl	TEMED
15 µl	TEMED		

3x Tricine Sample Buffer

2 ml	4x Tris-HCl/SDS, pH 6.8
	(12 M Tris-Base; 40 mM SDS; pH adjusted with HCl)
2.4 ml	Glycerol
0.8 g	SDS
0.31 g	DTT
2 mg	Coomassie Blue G-250

Final volume was increased to 10 ml with dH$_2$O

Anode Running Buffer

0.2 M	Tris-Base

Adjusted to pH 8.9 with HCl

Cathode Running Buffer

0.1 M	Tris-Base
0.1 M	Tricine
0.1% (w/v)	SDS

Coomassie Blue Stain

7% (v/v)	Acetic Acid
5% (v/v)	Methanol
Coomassie Blue	

De-stain Solution

7% (v/v)	Acetic Acid
5% (v/v)	Methanol

Silver Stain Buffers (Pharmacia Biotech)

Fixative Buffer

40% (v/v)	Ethanol
10% (v/v)	Glacial Acetic Acid

Sensitising Buffer

30% (v/v)	Ethanol
0.125% (w/v)	Gluteraldehyde
0.2% (w/v)	Sodium Thiosulphate
6.8% (w/v)	Sodium Acetate

Silver Solution

0.25% (w/v)	Silver Nitrate
0.015% (w/v)	Formaldehyde

Developing Buffer

2.5% (w/v)	Sodium Carbonate
0.0075% (w/v)	Formaldehyde

Stop Buffer

1.46% (w/v)	EDTA-Na$_2$.2H$_2$O

Western Blot Transfer Buffer

25 mM	Tris Base
200 mM	Glycine
20% (v/v)	Methanol

Adjusted to pH 8.3 with HCl

Wash Buffer (Tween 20-PBS)

0.1% (v/v)	Tween 20
0.17 M	Sodium Chloride
3.35 mM	Potassium Chloride
1.84 mM	Potassium Dihydrophosphate
10 mM	Sodium Dihydrogen Phosphate

Adjusted to pH 7.4

ECL Reagents

ECL Solution 1	ECL Solution 2
2.5 mM Luminol 0.4 mM Coumaric Acid 0.1 M Tris-Cl pH 8.5	0.02% (v/v) Hydrogen Peroxide 0.1 M Tris-Cl pH 8.5

11.1.6 Immunoprecipitation

RIPA Buffer x1

150 mM	Sodium Chloride
10 mM	Tris-HCl pH 8
1% (w/v)	Sodium Deoxycholate
1% (v/v)	Triton X100
1% (w/v)	SDS

1x Machamer Buffer

150 mM	Sodium Chloride
50 mM Tris-Cl, pH 8	
5 mM	EDTA
1% (v/v)	Triton X100
Proteinase Inhibitor	

Lithium Chloride Wash Buffer

0.6 M	Lithium Chloride
0.1 M	Tris-HCl pH 8.0
1% (v/v)	β-Mercaptoethanol

11.1.7 RNA – Protein Association

TE Buffer

10 mM	Tris Base
1 mM	EDTA

Adjusted to pH 7.6

10x TBE Buffer

1 M	Tris Base
0.8 M	Boric Acid
10 mM	EDTA

Adjusted to pH 8.2-8.4

5x RNA Binding Buffer

50 mM	Tris-HCl, pH 7.2
0.5 M	Potassium Chloride
15 mM	Magnesium Acetate
5 mM	EDTA

10x HBB

250 mM	Hepes, pH 7.5
250 mM	Sodium Chloride
50 mM	Magnesium Chloride
10 mM	DTT

10x HYB100

0.2 M	Hepes, pH 7.5
1 M	Potassium Chloride
25 mM	Magnesium Chloride
1 mM	EDTA
10 mM	DTT
0.5% (v/v)	Nonidet P-40

11.1.8 Nucleocapsid Preparations

1x TNE Buffer

150 mM	Sodium Chloride
10 mM	Tris Base
1 mM	EDTA

Solutions	Weight (g) placed in 100 ml TNE Buffer	Refractive Index	Density (g / cm³)
5% (w/v) Sucrose	5	-	-
25% (w/v) Caesium Chloride	30.69	1.3550	1.2275
30% (w/v) Caesium Chloride	38.57	1.3607	1.2858
40% (w/v) Caesium Chloride	56.78	1.3735	1.4196

5x Polynucleotide Exchange Reaction Buffer

350 µM	ADP
250 mM	Imidazole (pH 6.8)
60 mM	Magnesium Chloride
5 mM	β-Mercaptoethanol

11.2 Primer Sequences

11.2.1 Primers used in the Cloning of RSV N ORF

N (Forward)

5' TGGGGCAAAGGATCCGATGGCTCTTAGCAAAGTCAAGTTGA 3'

N (Reverse)

5' TTCAGGAGCGGATCCTTCCATGATGATTTATTTGCCCCATT 3'

GGATCC represents the *Bam* HI restriction site introduced into the N ORF by cloning primers; ATG is the start codon.

11.2.2 Site-Directed Mutagenesis Primers

NΔ1-200

5' GCAGCGGCCATATCGAAGGTCGTCATATGCTCGAGGATCCGATGCCC
AAGGACATAGCCAACAGCTTCTATGAAGTGTTTGAAAAAC 3'

NΔ121-160

5' CATTGGCAAGCTTAACAACTTGTATAGCAGCATTAGTAATAA 3'

11.2.3 PCR Deletion Primers

N Forward

5' CATATGCTCGAGGATCCGATGGCTCTTAGCAAAGTC 3'

NΔ342-391

5' CTTTGCTGCATCATATGGATCCTGATTTCACGGTGTACCTCTGTACTCTC 3'

NΔ292-391

5' GGTAGAATCCTGCTTCAGGATCCAATTTTCAGGCATATTCATAAACC 3'

NΔ242-391

5' CATAGGCATTCATAAAGGATCCTGCAAATCACCCTTCAACTCTACTG 3'

NΔ192-391

5' GTAGTAAGCCTTTGTAA<u>GGATCC</u>ATTTC<u>TCA</u>TTTTAGGACATTATTAG 3'

NΔ142-391

5' CATGCCTGTATTCTGGA<u>GGATCC</u>TCTCC<u>TCA</u>TTCTTTTAGCAT TTTTTTG 3'

NΔ92-391

5' GATGTGTTGTTACATC<u>GGATCC</u>ATT<u>TCA</u>TTTTACATGATATCCC 3'

NΔ42-391

5' ATAACATGCCACATAAC<u>GGATCC</u>ATGTG<u>TCA</u>CTGCACATCATAATTAG 3'

GGATCC represents the *Bam* HI restriction site and TCA, the stop codon, introduced into the N ORF by deletion primers.

11.2.4 <u>Sequencing Primers</u>

T7 Promoter

5' ATTAATACGACTCACTATAG 3'

T7 Terminator

5' GCTAGTTATTGCTCAGCGGT 3'

<u>N SEQ 2</u> (internal)

5' GGGCAAGTGATGTTACGG 3'

<u>N SEQ 3</u> (internal)

5' ACTGATTTTGCTAAGACT 3'

<u>N REV</u> (internal)

5' ATACCTATTAACCCAGTG 3'

11.3 DNA Templates for RNA Probes

11.3.1 Specific Probe (Synthesised Oligo)

5'   AAT TTA ATA CGA CTC ACT ATA GGG ACG CGA AAA AAT GCG TAC

3'   TTA AAT TAT GCT GAG TGA TAT CCC TGC GCT TTT TTA CGC ATG

AAC AAA CTT GCA TAA ACC AAA AAA AT   3'

TTG TTT GAA CGT ATT TGG TTT TTT TA   5'

The T7 promoter is presented in red. Transcription begins immediately after the TATA sequence, generating a transcript of 47 nucleotides. RSV leader sequence is presented in blue.

11.3.2 Non-Specific Probe (Sal I-cut Bluescript plasmid)

5'   TAA TAC GAC TCA CTA TAG GGC GAA TTG GGT ACC GGG CCC

3'   ATT ATG CTG AGT GAT ATC CCG CTT AAC CCA TGG CCC GGG

CCC CTC GAG G   3'

GGG GAG CTC C   5'

The T7 promoter is presented in red. The transcribed sequence is presented in blue (32 nucleotides).

11.4 Histidine-Tagged Nucleocapsid Protein sequence

1   HHHHHHHHHHSSGHIEGRHMLEDPMALSKVKLNDTLNKDQLLSSSKYTIQ

51   RSTGDSIDTPNYDVQKHINKLCGMLLITEDANHKFTGLIGMLYAMSRLGR

101   EDTIKILRDAGYHVK ANGVDVTTHRQDINGKEMKFEVLTLASLTT EIQIN

$N_{\Delta 92\text{-}391}$ $N_{\Delta 121\text{-}160}$

151   IEIESRKSYKKMLKE MGEVAPEYRHDSPDCGMIIL CIAALVITKLAAGDR

$N_{\Delta 142\text{-}391}$ $N_{\Delta 121\text{-}160}$

201   SGLTAVIRRANNVLK NEMKRYKGLL PKDIANSFYEVFEKHPHFIDVFVHF

$N_{\Delta 192\text{-}391}$ $N_{\Delta 1\text{-}200}$

251   GIAQSSTRGGSRVEG IFAGLFMNAYGAGQVMLRWGVLAKSVKNIMLGHAS

$N_{\Delta 242\text{-}391}$

301   VQAEMEQVVEVYEYA QKLGGEAGFYHILNNPKASLLSLTQFPHFSSVVLG

$N_{\Delta 292\text{-}391}$

351   NAAGLGIMGEYRGTP RNQDLYDAAKAYAEQLKENGVINYSVLDLTAEELE

$N_{\Delta 342\text{-}391}$

401   AIKHQLNPKDNDVEL *

Reverse

Mutants

The histidine-tag and inter-linking amino acids are presented in blue.