uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

u Ottawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Valerie Suzanne Greco Stewart
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Biochemistry)
GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Characterization of Interaction of Human DNA-Dependent RNA Polymerases and Transcription Factors with RNA from the Hepatitis *delta* Virus: Novel Perspectives in HDV Biology
TITRE DE LA THÈSE / TITLE OF THESIS

Martin Pelchat
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Jocelyn Côté

Jean-François Couture

Alain Stintzi

John Taylor (University of Pennsylvania)

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Characterization of the interaction of human DNA-dependent RNA polymerases and transcription factors with RNA from the hepatitis *delta* virus: novel perspectives in HDV biology

Valerie Suzanne Greco-Stewart

A thesis submitted to the

School of Graduate Studies and Research

University of Ottawa

In partial fulfillment of the requirements for the

Biochemistry PhD program

Department of Biochemistry, Microbiology, and Immunology

Faculty of Medicine

Your file Votre référence
ISBN: 978-0-494-61386-3
Our file Notre référence
ISBN: 978-0-494-61386-3

ABSTRACT

Hepatitis *delta* virus (HDV) is a unique pathogen comprised solely of a single-stranded RNA molecule of ~1,700 nucleotides. It contains only one open-reading frame and lacks the polymerase necessary for its replication. RNAP II has been implicated in HDV RNA synthesis, although the involvement of additional host polymerases has also been suggested. The present study establishes the interaction of HDV RNA with RNAP II both in cells replicating HDV and in nuclear extracts. Binding of RNAP II to HDV RNA occurs at both terminal stem-loop domains of the HDV genome and antigenome. Mutations that eliminate the secondary structure or alter the primary sequence of these domains result in abrogation of the RNA-RNAP II interaction. Sequence analyses of these regions from 81 HDV isolates show maintenance of the rod structure, conservation of a CUG triplet, and a strong purine/pyrimidine polarization, indicating that these features are necessary for RNAP II recognition. Similar experiments demonstrate that RNAP I and RNAP III binding also occurs at these terminal stem-loop domains. RNAP I, RNAP II, and RNAP III each form independent complexes on HDV RNA that contain TBP, a subunit common to all polymerases. TBP also interacts with HDV RNA in infected cells, suggesting that RNAP binding to HDV RNA occurs through this shared subunit. To complement these studies, UV-crosslinking of nuclear extract containing HDV RNA was performed and mass spectrometry of the resultant RNPs led to the identification of the multifunctional protein PSF as an HDV-interacting protein. This interaction was confirmed by EMSA and co-immunoprecipitation experiments and binding was localized to both terminal domains of both polarities of HDV RNA. RNA affinity chromatography demonstrated that PSF and RNAP II are present on HDV RNA simultaneously and that HDV-RNAP II interaction is abolished in the presence of PSF-binding competitive RNAs, suggesting that PSF acts as a transcription factor that

facilitates the recruitment of RNAP II to the HDV template. These data provide insight into both HDV biology and RNA promoter recognition by RNAPs, establishing a solid foundation for future investigation into mechanisms of HDV replication and alternate template utilization by DNA-dependent RNA polymerases.

ACKNOWLEDGEMENTS

I would sincerely like to thank everyone who facilitated my doctoral experience and made it both fulfilling and successful. I would like to take this opportunity extend my gratitude to my family, friends, colleagues, and supervisors who made this experience possible; although by no means a comprehensive list, I would like everyone who supported and encouraged me to know that they have my sincerest appreciation and respect, and that without them acquisition of this degree would not have been possible.

My parents have continuously supported my academic endeavours and provided unlimited emotional, financial, and spiritual support. My father, Michael Greco, has always believed in me and encouraged me to use my abilities to the fullest extent of their potential. My mother Patricia has also provided me with the strength to pursue this degree, and thesis composition would not have been possible without her assistance with my son, Alexander. He is also a continuous source of inspiration, with his young, unbridled enthusiasm for life; his joy and his happiness have been motivational and have helped me aspire to be the best that I can be. My husband Wilfred has also supported me unquestioningly; he has been an incredible asset to me and to the accomplishment of this degree by helping me through the difficult times and supporting and loving me throughout the course of our relationship and marriage. He has been an incredible friend and partner, and he has my greatest appreciation. My sister Carey has also been invaluable to me, both in life and during my academic career. Our relationship is multifaceted; she is not only a sister, but a colleague and a friend, and I have relied on her in all of these capacities. I would also like to acknowledge the late Howard Easter, my grandfather, who epitomized the meaning of living a great and fulfilling life. I dedicate my thesis to these people who have so enriched my life and provided me with

the support, passion, and inspiration to fulfill my dreams and become the best person that I can become.

I would also like to acknowledge my peers, colleagues, and supervisors at the University of Ottawa who have provided me with such a rich and fulfilling experience. My supervisor, Martin Pelchat, has provided me with the opportunity to perform high-calibre research in his laboratory, and I am grateful to have had the opportunity to work with him. His brilliance with experimental design, guidance throughout my project, and support is much appreciated. I have not only acquired new skills in research in his laboratory, but learned to design experiments, think analytically, and aim high with respect to my research and future ambitions, and I am very grateful for the opportunity to have been a part of his research team. I would also like to thank my laboratory colleagues and graduate student peers, both past and present, who have worked with me to make my university experience pleasant and successful. I would like to acknowledge my Thesis Advisory Committee members Odette Laneuville, Ilona Skerjanc, and Linda Bonen and thank them for their support and guidance throughout the course of my doctoral studies. Finally, I wanted to thank all of the university professors, academic and technical assistants, and support staff who have helped me over the years and to let them know that their assistance is much appreciated. Together, you have made this experience possible and set the foundation for my life's future endeavours in the world of science.

Thank you all for providing me with this opportunity and for your support and assistance, and I promise to use the skills acquired throughout the course of my doctoral studies to the fullest extent of my abilities. I want to let you know that your investment in me is worthwhile and is much appreciated.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

aa	Amino acids
ADAR 1	Adenosine deaminase acting on ribonucleic acid (small isoform)
ARM	Arginine-rich motif
ASBVd	Avocado sunblotch viroid
ASF	Arginine/serine-rich splicing factor
bp	Base pairs
CCCVd	Coconut cadang cadang viroid
cDNA	Complimentary deoxyribonucleic acid
CEV	Citrus exocortis viroid
ChIP	Chromatin immunoprecipitation assay
CKII	Casein kinase II
CTD	Carboxy terminal domain
DIPA	*delta* interacting protein A
DNA	Deoxyribonucleic acid
DRB	5,6-dichloro-1-β-D-ribofuranosylbenzimidazole
dsDNA	Double-stranded deoxyribonucleic acid
DSIF	DRB-sensitivity inducing factor
dsRNA	Double-stranded ribonucleic acid
eEF1A1	Eukaryotic translation elongation factor 1A1
EMSA	Electrophoretic mobility shift assay
ERK1/2	Extracellular signal-related kinase 1/2
ESE	Exonic splicing enhancer
FTase	Farnesyltransferase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GTF	General transcription factor

HAT	Histone acetyltransferase
HBV	Hepatitis B virus
HDAg	Hepatitis *delta* antigen
HDAg-L	Hepatitis *delta* antigen, large isoform
HDAg-S	Hepatitis *delta* antigen, small isoform
HDV	Hepatitis *delta* virus
hnRNP	Heterogeneous nuclear ribonucleoprotein
hnRNPL	Heterogeneous nuclear ribonucleoprotein L
HSF1	Heat shock transcription factor 1
INR	Initiator element
INS	Instability element
IRES	Internal ribosome entry site
K_D	Dissociation constant
kDa	Kilodalton
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LUCA	Last universal common ancestor
MHV	Mouse hepatitis virus
miRNA	Micro ribonucleic acid
mRNA	Messenger ribonucleic acid
MSAT	Mirror-spliced antisense RNA
ncRNA	Non-coding ribonucleic acid
NE	Nuclear extract
NELF	Negative elongation factor
NEP	Nuclear-encoded polymerase
NES	Nuclear export signal
NESI	Nuclear export signal-interacting protein
NLS	Nuclear localization signal

nt	Nucleotides
NTP	Nucleoside triphosphate
ORF	Open reading frame
$p54^{nrb}$	54 kDa nuclear RNA-binding protein
PCR	Polymerase chain reaction
PEP	Plastid-encoded polymerase
PIC	Pre-initiation complex
PKC	Protein kinase C
PKR	Double-stranded RNA-activated protein kinase R
PLMVd	Peach latent mosaic viroid
PPE	Pre-messenger ribonucleic acid processing enhancer
PRMT1	Protein arginine methyltransferase 1
PSF	Polypyrimidine tract-binding protein-associated splicing factor
PSTVd	Potato spindle tuber viroid
PTB	Polypyrimidine tract-binding protein
P-TEFb	Positive transcription elongation factor B
RBD	RNA binding domain
RdRp	RNA-dependent RNA polymerase
RIPA	Ribonucleoprotein immunoprecipitation assay
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAP	DNA-dependent RNA polymerase
Rpb	RNA polymerase B (II) subunit
RRM	RNA recognition motif
rRNA	Ribosomal ribonucleic acid
RNP	Ribonucleoprotein

RT	Reverse transcriptase
RT-PCR	Reverse transcriptase-polymerase chain reaction
SELEX	Systematic evolution of ligands by exponential enrichment
SFPQ	Splicing factor proline/glutamine-rich
shRNA	Small hairpin ribonucleic acid
siRNA	Small interfering ribonucleic acid
snRNA	Small nuclear ribonucleic acid
spRNA	Small priming ribonucleic acid
ssDNA	Single-stranded deoxyribonucleic acid
ssRNA	Single-stranded ribonucleic acid
TAF	TATA-binding protein-associated factor
TBP	TATA-binding protein
TFIID	RNA polymerase II transcription factor IID
TFIIS	RNA polymerase II transcription factor IIS
tRNA	Transfer ribonucleic acid
qRT-PCR	Quantitative reverse transcriptase-polymerase chain reaction
WHV	Woodchuck hepatitis virus
YY1	Yin Yang 1

NUCLEOSIDE ABBREVIATIONS

A	Adenosine
C	Cytidine
G	Guanosine
T	Thymidine
U	Uridine

AMINO ACID ABBREVIATIONS

A	Alanine
C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
H	Histidine
I	Isoleucine
K	Lysine
L	Leucine
M	Methionine
N	Asparagine
P	Proline
Q	Glutamine
R	Arginine
S	Serine
T	Threonine
V	Valine
W	Tryptophan
Y	Tyrosine

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1:

INTRODUCTION

1.1 The hepatitis *delta* virus: an overview

The hepatitis *delta* virus (HDV) is the smallest known mammalian RNA pathogen. First identified in 1977 in patients infected with the hepatitis B virus (HBV), HDV is a subviral satellite of HBV, requiring the HBV envelope proteins for encapsidation and transmission (Rizzetto et al., 1977; Bonino et al., 1984). It is estimated that approximately 10% of patients infected with HBV are co-infected with the HDV virus (www.cdc.gov). HBV is prevalent throughout the world; it is estimated that two billion people have been exposed to HBV while about 350 million people suffer from chronic HBV infection (data acquired in 2006; www.cdc.gov). Although most prevalent in tropical and sub-tropical regions, HDV remains present in developed countries due to the high incidence of HBV infection throughout the world and is thus responsible for a significant amount of global morbidity and mortality.

HDV co-infection exacerbates HBV disease progression and increases the likelihood of developing adverse sequelae including cirrhosis of the liver, hepatocellular carcinoma, and fulminant hepatitis, a lethal form viral encephalopathy (Govindarajan et al., 1984; Jacobson et al., 1985; Kanel et al., 1984; Romeo et al., 2009). A recently-published 28-year study of patients suffering from chronic HDV infection has shown that presence of active HDV replication is strongly correlated with the development of liver failure (Romeo et al., 2009). It is estimated that 80% of individuals who are chronically infected with HDV will experience cirrhosis of the liver within five to ten years (Koytak et al., 2007). Although antiviral therapies such as treatment with α-interferon and, more recently, pegylated

interferon have demonstrated modest success in the management of HBV/HDV infection, liver transplantation has traditionally remained the only viable option for the treatment of progressive, chronic infection (Farci et al., 2007; Koytak et al., 2007). It is thus essential to establish novel methods of treatment for people infected with this pathogen - an endeavor which will require a greater understanding of the mechanism by which this virus replicates and causes disease.

HDV is unique among viruses and was determined to be similar to viroids, single-stranded RNA pathogens that infect numerous higher plant species (for current viroid reviews see Daròs et al., 2006 and Tsagris, 2008). Viroids are comprised of a single-stranded, circular RNA molecule of approximately 250 to 400 nucleotides (nt) and do not encode any proteins (Davies et al., 1974; Diener, 1974; Hall et al., 1974). Viroids are divided into two major taxa, *Avsunviroidae* and *Pospiviroidae*, based on their structure, mechanism of replication, and intracellular localization (Flores et al., 1998). Avsunviroids accumulate in the chloroplast, the photosynthetic organelle of plants, replicate by a symmetrical rolling circle mechanism, and their genomes adopt complex RNA secondary structures including many stem-loop domains and autocatalytic hammerhead motifs (reviewed by Daròs et al., 2006 and Tsagris, 2008). By contrast, pospiviroids accumulate in the nucleus of infected cells where they replicate by an asymmetrical rolling circle mechanism and their genomes acquire a rod-shaped conformation (Daròs et al., 2006; Tsagris, 2008; viroid replication is further discussed in Chapter 7.1.3). Introduced to plant tissues through mechanical insult, the mechanism of viroid pathogenesis is largely unknown, though disease caused by these pathogens accounts for a substantial amount of agricultural and economic loss. Although HDV exhibits certain properties of both viroid families, it is the only known subviral

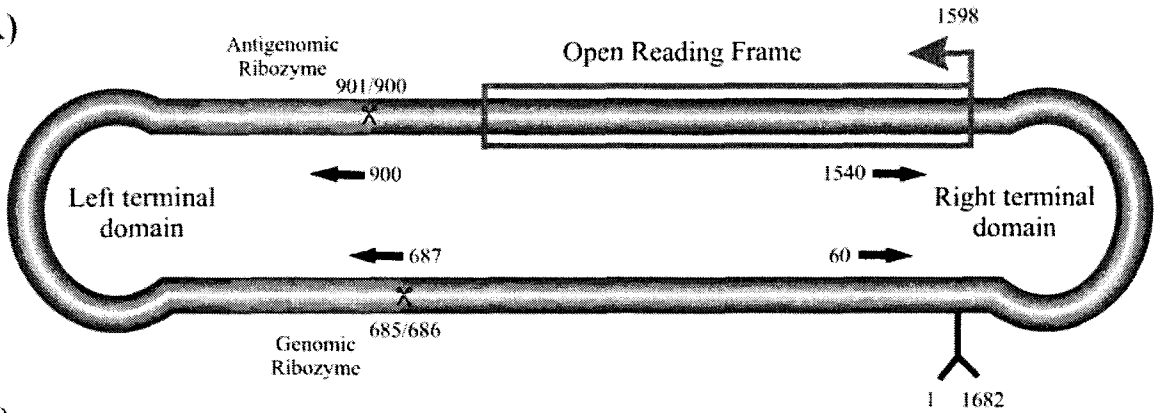
pathogen that infects mammals and subsequently demonstrates a higher level of complexity in both its genome and replication strategy.

HBV co-infection is necessary for dissemination of HDV. The small, medium, and large envelope proteins of HBV are required for encapsidation of HDV ribonucleoprotein (RNP) particles containing HDV RNA and antigens to form infectious virions (Bonino et al., 1984; reviewed by Sureau, 2006). HDV infection occurs through direct exposure to blood or body fluids containing these virions and, upon entering the bloodstream, these virions migrate to the liver since the HBV envelope proteins confer a specific hepatocyte tropism (summarized in Lazinski, 1999). Upon cellular entry, uncoating of the virion occurs, releasing HDV RNPs. These RNPs then enter the nucleus for HDV replication.

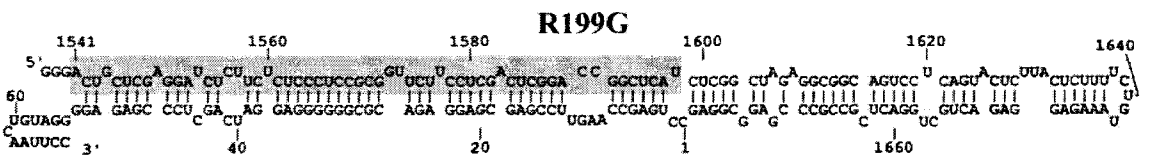
HDV was first identified by the presence of a novel antigen in the hepatocytes of HBV-infected patients and was subsequently shown to replicate in woodchucks co-infected with the woodchuck hepatitis virus (WHV), a hepadenovirus capable of packaging HDV for transmission (Ponzetto et al., 1984; Rizzetto et al., 1977). In 1988, HDV was characterized at the molecular level by the Taylor laboratory (Kuo et al., 1988a; reviewed by Lai, 2005 and Taylor, 2006). HDV is comprised of a single-stranded, G/C-rich, circular RNA molecule of approximately 1,700 nucleotides (Chen et al., 1986; Figure 1.1). It adopts a rod-shaped secondary structure that demonstrates a high proportion (~74%) of canonical intramolecular base pairing (Kuo et al., 1988a, Wang et al., 1986). Despite its extreme complimentarity, HDV is resistant to cleavage by dicer, a component of the RNA-induced silencing complex (RISC) which mediates the cellular antiviral response by degrading double-stranded, highly-complimentary RNA species (Chang et al., 2003). HDV accumulates in the nucleus of infected cells and replicates by a symmetrical, rolling circle mechanism reminiscent of that used by *Avsunviroidae* viroids (Branch and Robertson, 1984; Macnaughton and Lai, 2002b;

Figure 1.1: The hepatitis *delta* virus. A) The hepatitis *delta* genome is depicted with superimposition of both genomic and antigenomic polarities. The terminal hairpin domains, primer annealing positions for synthesis of these domains, *delta* ribozyme motifs and cleavage sites, and open reading frame for HDAg mRNA are indicated. Numbering is in accordance with Kuo et al. (1988a). B) Sequence and structure of R199G, the right terminal stem-loop domain of genomic polarity that is reported to have promoter activity. The transcribed region corresponding to HDAg mRNA is highlighted.

A)


1598
Antigenomic
Ribozyme
Open Reading Frame
901/900
900
1540
Left terminal
domain
Right terminal
domain
687
60
685/686
Genomic
Ribozyme
1
1682

B)


R199G
1541
1560
1580
1600
1620
1640
1660
40
20
1

Macnaughton et al., 2002; Figure 1.2). Replication involves transcription of the HDV genome to generate linear, concatameric species. These species are then cleaved to unit length by the endogenous *delta* ribozyme autocatalytic motif (discussed below). The monomeric HDV genomes are ligated, presumably by host factors, to generate the antigenomic species, which are subsequently transcribed as concatamers, cleaved, and ligated, regenerating the genomic species (Macnaughton et al., 2002; Reid and Lazinski, 2000). During HDV replication, two RNA species accumulate: the **genome** comprises the most abundant species present in the cell while the **antigenome** is present in an amount approximately 10-fold lower than that of its genomic counterpart (Chen et al., 1986).

One feature of HDV that facilitates its replication strategy is the presence of *delta* ribozyme motifs (Kuo et al., 1988b; Wu et al., 1989; Figure 1.3). A **ribozyme** is an RNA species that demonstrates enzymatic activity. The *delta* ribozyme possesses intrinsic endonuclease activity that causes *cis*-cleavage of HDV RNA. In addition to the *delta* ribozyme, several other ribozymes with RNA cleavage capacity exist naturally including the hammerhead, hairpin, glmS, Varkud satellite (VS), and CPEB ribozymes, the last of which is considered to be a human homologue of the *delta* ribozyme (reviewed by Cochrane and Strobel, 2008; Salehi-Ashtiani et al., 2006). Two *delta* ribozyme motifs are present on HDV RNA, one on the genome (nucleotides 659 to 772) and one on the antigenome (nucleotides 847 to 966; Chao, Y.C. et al., 1990). Cleavage by the *delta* ribozyme occurs between nucleotides 685 and 686 of the genome and nucleotides 901 and 900 of the antigenome (Kuo et al., 1988b; Sharmeen et al., 1988; Figure 1.1). The *delta* ribozyme adopts a complex, double pseudoknot structure comprised of five paired regions that facilitate autocatalytic cleavage of the genome and antigenome during replication (Ferré-D'Amaré et al., 1998; Ke et al., 2004; Figure 1.3). Cleavage by the *delta* ribozyme can occur in physiological, cellular

Figure 1.2: HDV replicates by a symmetrical, rolling circle mechanism. Transcription of linear concatamers occurs from a circular, unit-length template. The antigenomic, concatameric product is cleaved by endogenous *delta* ribozymes (black bar), ligated by host factors, and serves as a template for regeneration of the genomic species. The mRNA encoding HDAg is transcribed from a putative promoter in the right terminal stem-loop domain (black circle).

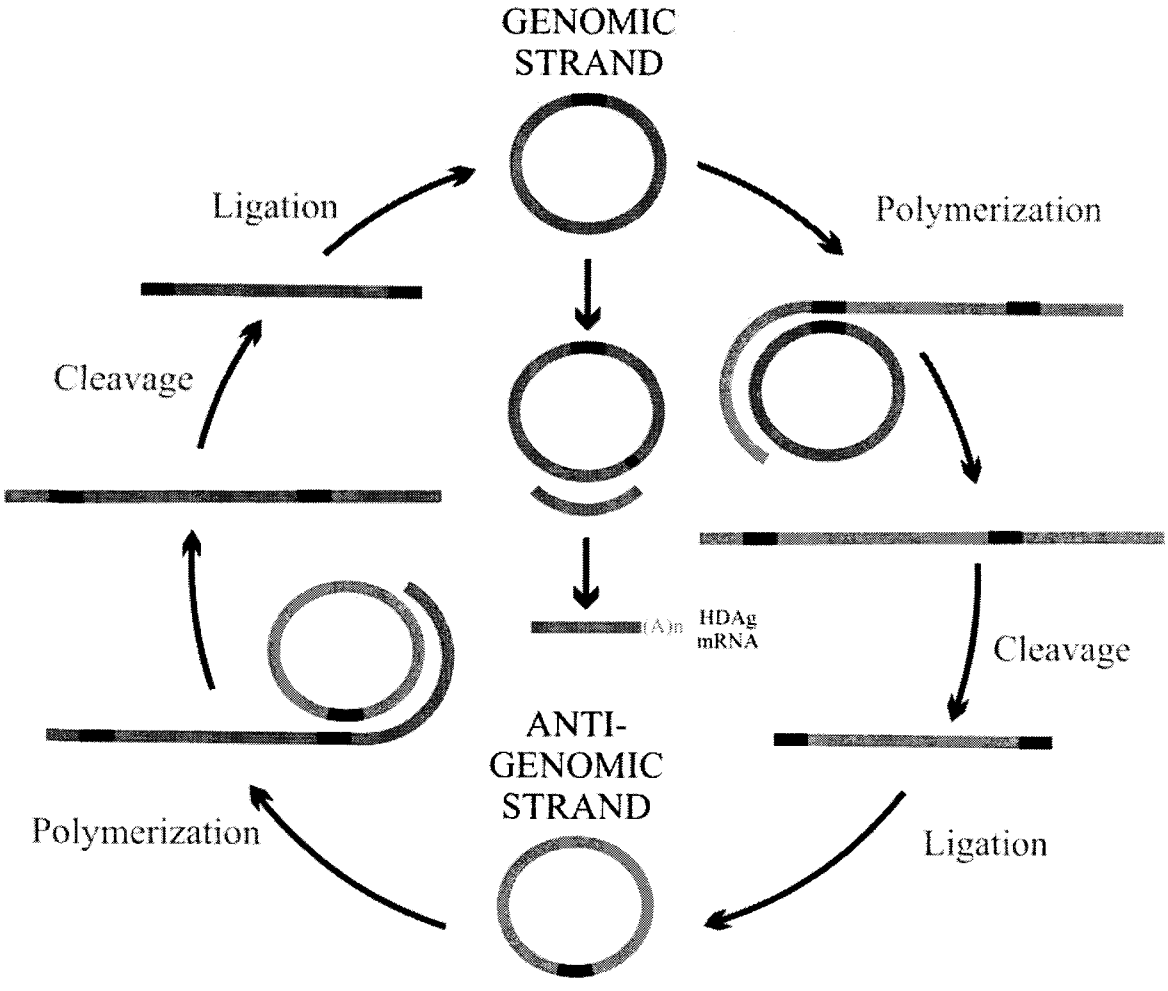
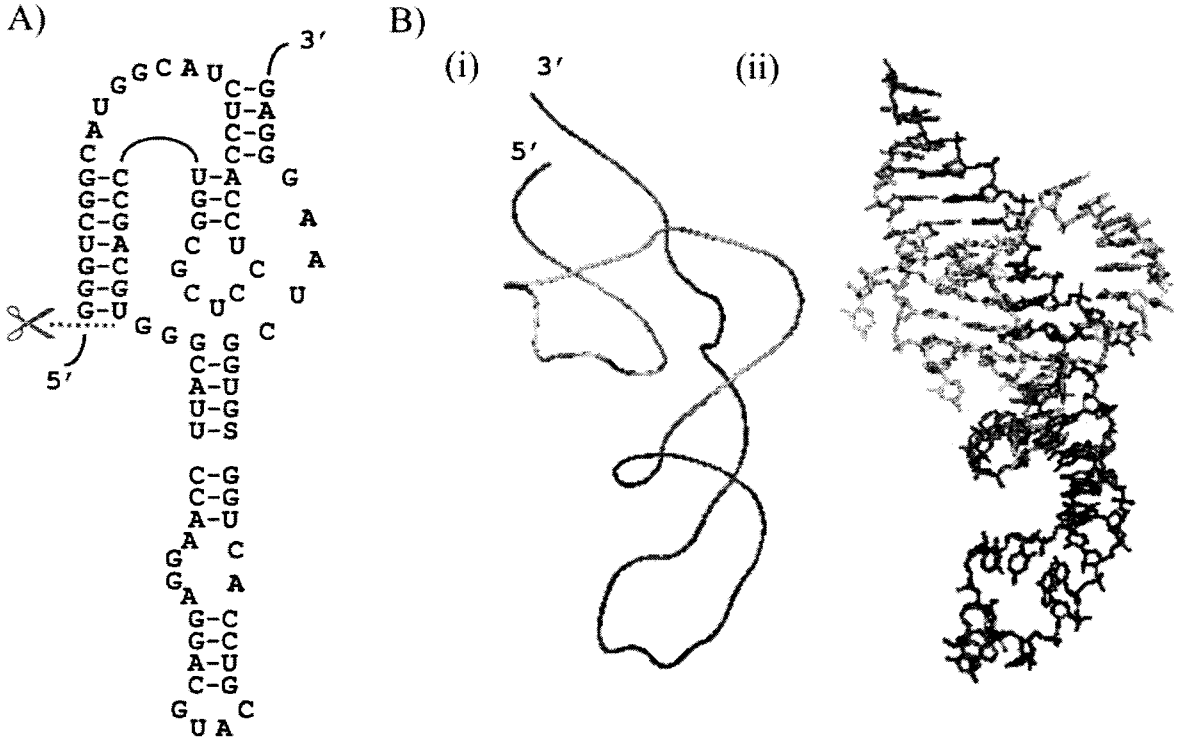

GENOMIC
STRAND
Ligation
Polymerization
Cleavage
(A)n
HDAg
mRNA
Cleavage
ANTI-
GENOMIC
STRAND
Polymerization
Ligation

Figure 1.3: The *delta* ribozyme motif. A) Nucleotide sequence of the genomic *delta* ribozyme motif. B) Crystal structure of the *delta* ribozyme motif backbone (i) and complete nucleotide structure (ii) as resolved by Ferré D'Amaré et al. (1998; PDB accession number 1CX0).

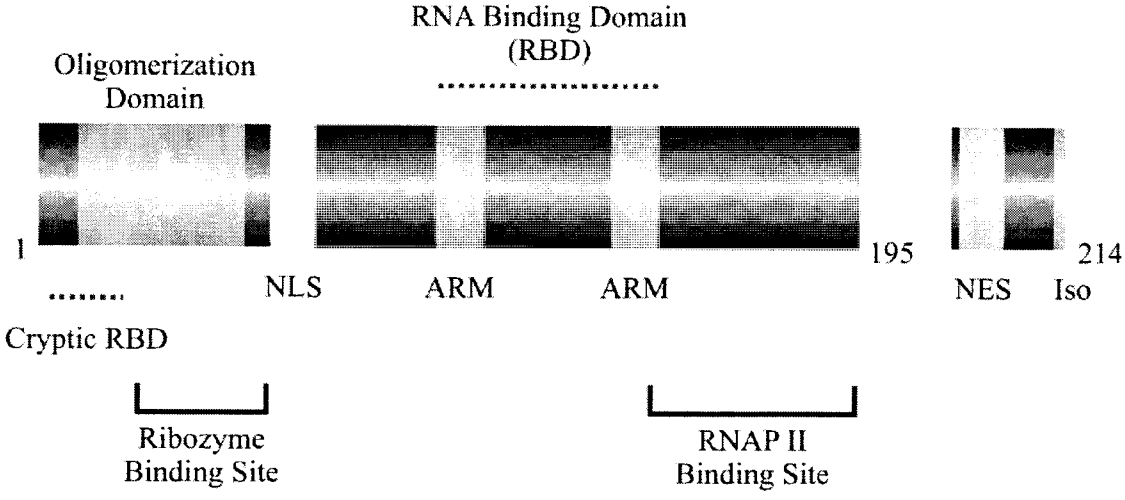

A)
B)
(i)
(ii)
3′
5′
3′
5′

conditions and requires low concentrations of divalent cation such as magnesium (0.05 mM to 0.1 mM; Rosenstein and Been, 1990). In denaturing conditions, such as in the presence of 5 M urea or 10 M formamide, cleavage is enhanced as long as the concentration of divalent cation is increased (Rosenstein and Been, 1990). Cleavage results in the generation of a terminal uridyl 2',3'-cyclic monophosphate residue and new 5' guanosyl residue with a 5'-hydroxyl group at the site of cleavage (Wu et al., 1989). Examination of dozens of HDV isolates has shown extreme conservation of the ribozymes and adjacent sequences, indicating their complete necessity for effective HDV replication and propagation (Chadalavada et al., 2007; Greco-Stewart et al., 2007).

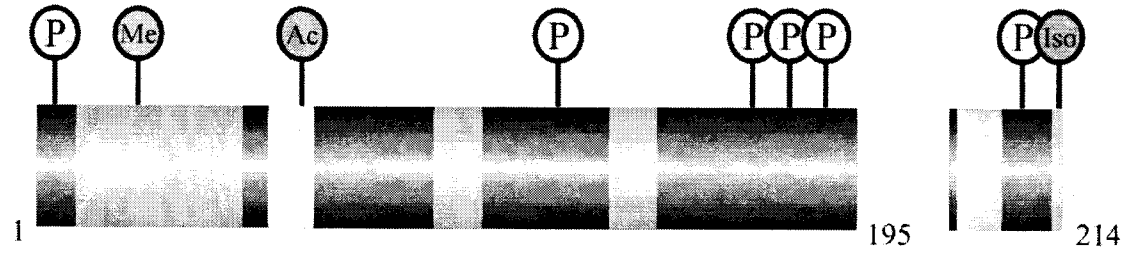
HDV contains a single open reading frame (ORF) encoding the 24 kDa small *delta* antigen [HDAg-S; 195 amino acids (aa); Weiner et al., 1988; Figure 1.4]. Post-transcriptional editing of the HDV antigenome results in loss of the HDAg-S messenger RNA (mRNA) stop codon, generating the 27 kDa large *delta* antigen (HDAg-L; 214 aa; Casey and Gerin, 1995; Weiner et al., 1988). By popular convention, the generic term "HDAg" is used in reference to both/either form of the antigen. HDAg-S is involved in HDV RNA synthesis, while HDAg-L is required for packaging of HDV RNPs into infectious particles (Chang et al., 1991; Chao, M. et al., 1990; Kuo et al., 1989; Lee et al., 1995; Ryu et al., 1992; Sureau et al., 1992). HDAg-L also acts as a dominant negative inhibitor of HDV transcription late in the HDV life cycle (Chao, M. et al., 1990; Lee et al., 1995; Macnaughton and Lai, 2002a). While HDAg-L alone is capable of working with HBV envelope proteins to achieve packaging of HDV RNA, the presence of HDAg-S enhances the process (Wang et al., 1994). Structurally, HDAg contains a single nuclear localization signal (NLS), a leucine-rich, antiparallel coiled-coil domain required for HDAg oligomerization, and two arginine-rich, highly-basic motifs necessary for RNA binding and transport (Alves et al., 2008;

Figure 1.4: Schematic representation of HDAg. A) Representation of HDAg showing significant functional regions. HDAg encompasses amino acids 1 to 195; the C-terminal expansion of HDAg-L (amino acids 196 to 214) is also indicated. NLS: nuclear localization signal (amino acids 66-75); NES: nuclear export signal (amino acids 198-210); ARM: arginine-rich motifs (amino acids 97-107 and 136-146); Iso: isoprenylation signal (amino acids 211-214). The cryptic RBD spans amino acids 2-27 while the main RBD spans both ARMs. The coiled-coil oligomerization domain spans amino acids 12-60. Ribozyme binding occurs at amino acids 24-50 and the RNAP II binding is localized to the C terminus at amino acids 150-195. B) Modification of HDAg. All possible modifications are shown concurrently, though not every modification is present simultaneously *in vivo*. Lettering for the modifications is P: phosphorylation, Me: methylation, Ac: acetylation, Iso: isoprenylation (farnesylation).

A)


RNA Binding Domain
(RBD)
Oligomerization
Domain
1
195
214
NLS
ARM
ARM
NES
Iso
Cryptic RBD
Ribozyme
Binding Site
RNAP II
Binding Site

B)


P
Me
Ac
P
P
P
P
P
Iso
1
195
214

Gowans et al., 1988; Lee et al., 1993; Xia and Lai, 1992; Xia et al., 1992; Zuccola et al., 1998; Figure 1.4A). The extended C-terminal terminal domain of HDAg-L contains a nuclear export signal (NES) and clathrin adaptor homology region which might facilitate clathrin-mediated exocytosis (Huang et al., 2007; Lee et al., 2001). HDAg can function as an RNA chaperone through its leucine-rich coiled-coil region and can stimulate *cis*-cleavage by the *delta* ribozyme (Huang and Wu, 1998). HDAg undergoes much post-translational modification including phosphorylation, methylation, acetylation, and isoprenylation (Figure 1.4B).

Although multifunctional, the protein product of this single-open reading frame is not, by itself, sufficient for viral transcription and replication. Remarkably, HDV lacks an RNA-dependent RNA polymerase (RdRp) which is traditionally required for RNA-directed RNA synthesis; this observation has prompted several questions regarding how HDV replicates in the human host in the absence of such enzyme. The present studies are thus aimed at determining the host polymerases and accessory factors involved in HDV replication in the human host and at determining the sequence and structural elements that constitute an RNA promoter for this unique pathogen.

1.2 HDV coercion of host cellular proteins: host-pathogen interactions

All viruses must subvert host cellular machinery for genome expression and virion production. HDV is exquisitely reliant upon the host cell for these processes due to its extremely limited protein coding capacity. Several host proteins interact with HDV during the course of infection, including those in the nucleus that facilitate transcription and replication, and those in the cytoplasm that are principally associated with production of HDAgs. Host proteins that interact with HDV can be divided into two categories – those that

interact with HDAg and those that interact with HDV RNA. While the biological significance of some of these intermolecular interactions is known, the role of most of these interactions in the life cycle of HDV remains elusive.

Genome-wide proteomic analysis of Huh7 hepatocellular carcinoma cells transfected with HDV RNA and both HDAgs have shown that at least 32 proteins demonstrate differential expression in the presence of HDV components (Mota et al., 2008). The affected proteins are involved in diverse processes such as nucleic acid metabolism, protein metabolism, cellular transport, signal transduction, apoptosis, and cell growth. This altered global protein expression profile could provide insight into the cellular factors with which HDV components interact in infected cells and suggest mechanisms of HDV pathogenesis. To date, numerous proteins have been demonstrated to interact specifically with HDAgs and HDV RNA, and some of these intermolecular interactions have been characterized and biological function proposed (summarized in Table 1.1 and Table 1.2). The following sections will discuss recent discoveries in the area of HDV-host interaction with emphasis on those which are significant to my thesis research.

1.2.1 Interaction of HDAgs with host proteins

HDAgs undergo much post-translational modification including serine and threonine phosphorylation, arginine methylation, lysine acetylation, and cysteine isoprenylation (Chang et al., 1988; Glenn et al., 1992; Mu et al., 1999; Mu et al., 2001; Mu et al., 2004; Otto and Casey, 1996; Li et al., 2004). These modifications contribute to processes such as subcellular localization and efficiency of HDV RNA synthesis, possibly by regulating intermolecular interactions involving HDAgs (Li et al., 2004; Mu et al., 2001; Mu et al. 2004). It has recently been proposed that the modifications present on HDAg-S might dictate

Table 1.1: Host proteins that interact with HDAg.

Host protein	Proposed Function
	Post translational modification
Casein Kinase II (CKII)	Phosphorylation (S2 and S213)
Double-stranded RNA-activated protein kinase R (PKR)	Phosphorylation (S177, S180, T182)
Extracellular signal-related kinases 1 and 2 (ERK1/2)	Phosphorylation (S177)
Protein Kinase C (PKC) [a]	Phosphorylation (S210)
Protein farnesyltransferase (FTase) [a]	Isoprenylation with farnesyl (C211)
Protein arginine methyltransferase 1 (PRMT1)	Methylation (R13)
p300 cellular acetyltransferase	Acetylation (K72)
	Sub-cellular localization
karyopherin (importin) 2α	Nuclear import
Nuclear export signal-interacting protein (NESI)[a]	Nuclear export
Nucleolin (C23)	Nucleolar localization, shuttling, RNA synthesis/accumulation (?)
Nucleophosmin (B23)	Nucleolar localization, shuttling, RNA synthesis/accumulation (?)
Yin Yang 1 (YY1)	RNA synthesis/accumulation (?)
Histone H1e[b]	RNA synthesis/accumulation (?)
	RNA synthesis
RNAP II (Rpb1/2 mobile clamp element)	Genome and mRNA synthesis
	Antigenome synthesis (?)
DRB sensitivity-inducing factor (DSIF)	Relieves transcriptional repression; stimulates elongation by RNAP II
delta interacting protein A	Transcriptional regulation (?)
	Other
MOV10	RNA remodelling (?)

(a) HDAg-L only. (b) HDAg-S only. All data references are cited in text.

Table 1.2: Host proteins that interact with HDV RNA.

Host protein	Proposed Function
Adenosine deaminase acting on RNA (ADAR 1)	Post-transcriptional modification of the antigenome that results in production of HDAg-L
Glyceraldehydes 3-phosphate dehydrogenase (GAPDH)	Enhances *delta* ribozyme activity
Double-stranded RNA-activated protein kinase R (PKR)	Recruitment to HDAg for post-translational modification Repression of antiviral response
RNAP I*	Antigenome synthesis
RNAP II*	Genome synthesis, mRNA synthesis, Antigenome synthesis
RNAP III*	Unknown
Polypyrimidine tract-binding protein associated splicing factor (PSF)*	Suspected involvement in recruitment of HDV RNA to RNAP II
54 kDa nuclear RNA-binding protein ($p54^{nrb}$)*	Unknown
Heterogeneous nuclear ribonucleoprotein L (hnRNPL)*	Unknown
Arginine/serine-rich splicing factor (ASF)*	Unknown
Eukaryotic elongation factor 1A1 (eEF1A1)*	Unknown

* Present study

All data references are cited in text.

which RNA species, such as genomic, antigenomic, or HDAg mRNA, is produced during HDV RNA synthesis (Chen et al., 2008; Tseng et al., 2008). Post-translational modifications are performed by numerous cellular proteins, many of which have been identified (Table 1.1).

Two phosphorylated isoforms of HDAg-S and one phosphorylated isoform of HDAg-L are present in infected cells (Mu et al., 1999). While HDAg-S is phosphorylated on both serine and threonine residues, HDAg-L is phosphorylated only on serine residues (Mu et al., 1999). Multiple kinases have been associated with HDAg phosphorylation including casein kinase II (CKII), protein kinase C (PKC), double-stranded RNA-activated protein kinase R (PKR), and extracellular signal-related kinases 1 and 2 (ERK1/2; Chen et al., 2002; Chen et al., 2008; Mu et al., 1999; Yeh et al., 1996). Phosphorylated residues show a high level of conservation among HDV isolates, and different phosphorylation states of HDAg have been associated with transcriptional enhancement, specifically in the case of genomic RNA synthesis (Chen et al., 2008; Mu et al., 2001; Yeh et al., 1994; Yeh and Lee, 1998). Protein arginine methyltransferase 1 (PRMT1) methylates arginine residues located in the RGGR motif in the RNA-binding domain of HDAg, and it has been shown that replication of HDV RNA, specifically the genomic species, requires methylation of arginine 13 of HDAg (Li et al., 2004). It has also been shown that HDAg methylation promotes nuclear retention of HDAg, which might be the reason for the observed transcriptional enhancement. Lysine acetylation, performed by the histone acetyltransferase (HAT) domain of p300 cellular acetyltransferase, has also been shown to promote nuclear retention of HDAg and increase transcription of the three major HDV RNA species (Huang et al., 2008a; Mu et al., 2004). Acetylation of lysine 72 specifically promotes an increase in early HDAg-L synthesis, possibly mediating the transition from active RNA synthesis early in the life cycle of HDV to

moderate synthesis in preparation for viral export later during infection (Mu et al., 2004). Finally, The 19 aa expansion of the carboxy terminal domain (CTD) of HDAg-L contains the isoprenylation signal CRPQ and is isoprenylated with farnesyl at cysteine 211 (Glenn et al., 1992; Otto and Casey, 1996). Isoprenylation causes a conformational change which masks 32 amino acids of the CTD of HDAg-L, possibly explaining its different functions as compared to those of HDAg-S (Hwang and Lai, 1994). This modification is suggested to be performed by protein farnesyltransferase (FTase), and the isoprenylated form of HDAg-L causes *trans*-repression of HDV transcription in the nucleus and facilitates viral packaging with HBV envelope proteins (Hwang and Lai, 1994; Lee et al., 1994). HDAg modification is thus believed to promote subcellular localization of HDV RNPs and enhance synthesis of specific HDV species at precise times during the HDV life cycle.

In addition to several sites for post-translational modification, HDAg contains a NLS comprised of ten amino acids located at position 66-75 that is necessary for nuclear entry of HDV RNPs (Alves et al., 2008). This sequence represents a classical type 1 NLS which is characterized by a large proportion of basic amino acids (Xia et al., 1992). The arginine-rich RNA binding motifs and NLS of HDAg are required for nuclear entry of HDV RNA (Chou et al., 1998). Nuclear entry of HDV RNPs is facilitated by karyopherin (importin) 2α, part of the nuclear pore complex that specifically binds to canonical NLS motifs to regulate entry of cytoplasmic components into the nucleus (Moroianu et al., 1995). This conventional interaction between the NLS of HDAg and karyopherin 2α thus has a specific role in HDV biology.

Following nuclear entry and replication, HDV RNA, in the form of RNPs, must be exported from the nucleus for virion packaging and cellular export. HDAg-L has an NES spanning residues 198-210 (Lee et al., 2001). Specifically, proline 205 plays an essential role

in HDV RNP trafficking out from the nucleus (Lee et al., 2001). Recently, a novel protein, the nuclear export signal-interacting protein (NESI), was identified as an HDAg-L-interacting partner (Wang et al., 2005). NESI is predominantly expressed in the liver and, when NESI expression is inhibited, HDV RNA fails to accumulate in the cytoplasm of infected cells and packaging of HDV RNPs is prevented. These data corroborate a unique and specific role for HDAg-L in viral export and packaging and show the necessity of another host protein for HDV propagation.

Nuclear proteins provide attractive candidates for HDAg-interacting proteins that regulate transcription and replication. Nucleolin, also known as C23, is a multifunctional nuclear phosphoprotein that has been demonstrated to interact with HDAg (Lee et al., 1998). Nucleolin is localizes to the nucleolus and plays numerous roles including synthesis and maturation of the ribosome, cell proliferation and growth, nuclear trafficking, cytokinesis, nucleogenesis, transcriptional repression, replication, signal transduction, and chromosome remodeling (for a review, see Tuteja and Tuteja, 1998). Although it has many properties including DNA-helicase and DNA-dependent ATPase activity, nucleolin also possesses helicase IV activity and can unwind RNA helices (Ghisolfi et al., 1992; Tuteja and Tuteja, 1998). Both HDAg-S and HDAg-L interact with nucleolin in cells, and accumulation of the antigenome increases concurrently with levels of nucleolin, implicating a role for this protein in HDV replication (Lee et al., 1998). The nucleolin-binding domain of HDAg resides in the N-terminus in a lysine/arginine-rich region which is conserved among HDV isolates, and it was shown that this domain is necessary for nucleolar targeting and accumulation of HDV RNAs. These data are suggestive of the involvement of nucleolin in the sequestration of HDAg for HDV replication, possibly through unwinding and stabilizing HDV RNA.

Nucleophosmin, also known as B23, is another nuclear phosphoprotein that has been shown to interact with both HDAg-S and HDAg-L (Huang et al., 2001). Like nucleolin, nucleophosmin has myriad functions including roles in cell growth and proliferation, nuclear shuttling, and ribosome biogenesis (reviewed by Okuwaki, 2008). Although predominantly associated with the nucleolus, nucleophosmin binds to the NLS of proteins to facilitate shuttling between the nucleus and the cytoplasm; it has recently been shown to mediate nuclear import of the viral protein Rev from HIV-1 (Szebeni et al., 1995; 1997). Cells replicating HDV and cells expressing HDAg were both shown to upregulate nucleophosmin mRNA transcription, and nucleophosmin was shown to co-localize with HDAg in the nucleolus (Huang et al., 2001). Binding of HDAgs to nucleophosmin occurs through the N-terminal of HDAg in a region encompassing the NLS but not the coiled-coil domain necessary for HDAg oligomerization. Furthermore, HDAg exists in a large complex with both nucleophosmin and nucleolin, and increasing the amount of nucleophosmin is associated with an increase in the amount of HDV RNA present in cells. From these data, it would appear that nuclear phosphoproteins such as nucleophosmin and nucleolin form RNP complexes within the nucleoli of infected cells to localize HDV RNPs and enhance HDV replication, although a role for these proteins in active RNA synthesis has not yet been described.

Yin Yang 1 (YY1) is a transcription factor that binds both nucleophosmin and nucleolin and has recently been shown to interact with both HDAg-S and HDAg-L (Huang et al., 2008b). Through interaction with numerous other nuclear and cytoplasmic proteins, YY1 can activate and repress gene expression involved in many cellular functions such as embryogenesis, differentiation, and proliferation (reviewed by Gordon et al., 2006). The central domain of HDAg containing its principal RNA binding domain (RBD) interacts with

the zinc-finger region of YY1 in a manner which is independent of HDV RNA, and these proteins, as well as CBP and p300 acetyltransferases, are found in large complexes (Huang et al., 2008b). A role for YY1 in HDV replication is presumed based on the previously reported putative functions nucleolin and nucleophosmin, and it was observed that levels of both genomic and antigenomic HDV transcripts increase in the presence of increasing amounts of YY1. An increase in YY1-associated acetyltransferases also increases HDV replication; acetyltransferase activity is specifically required for enhancement of HDV RNA accumulation. These data suggest that many nuclear and nucleolar proteins that exist in large RNP complexes could play a role in HDV RNA accumulation and/or synthesis. However, caution should be used in the assignment of a distinct function for these nucleolar proteins in HDV biology since interaction with double-stranded RNA (dsRNA) species and ribosomal RNA (rRNA) precursors might mediate the interactions of HDAg, itself an RNA-binding protein, with other nucleolar proteins without having direct impact on the HDV life cycle.

Histone H1e is another nuclear protein that was identified as an HDAg-S interacting partner (Lee and Sheu, 2008). Histones are proteins involved in DNA condensation and organization into chromatin (Kornberg, 1974; Kornberg and Thomas, 1974). In the nucleus, DNA winds around histone octomers to form core particles which are connected by H1 linker proteins to form nucleosomes, the fundamental units of chromatin architecture (reviewed by Happel and Doenecke, 2009). Linker histones are among the most abundant nuclear proteins and have been implicated in both the organization of higher chromatin structure and the modulation of gene expression by controlling the access of regulatory proteins to their DNA targets. The H1 family is contains eleven variants; the H1e variant, also known as histone H1.4, is ubiquitously and constitutively expressed in somatic cells. HDAg-S interacts with histone H1e through its leucine zipper motif, while HDAg-L is

unable to undergo this interaction due to masking of this region because of determinants located upstream of the isoprenylation signal (Lee and Sheu, 2008). It is believed that histone H1e assists HDV replication based on an observed reduction in accumulation of HDV RNAs in cells harboring H1e mutants defective in HDAg-S binding; this reduction is reversible by complementation with wild-type H1e (Lee and Sheu, 2008). It is speculated that H1e might play a role in HDV replication and transcription by a yet unknown mechanism.

The protein MOV10 was recently shown to interact with HDAg (Haussecker et al., 2008). Interestingly, the plant homologue of MOV10 is a putative RNA helicase that associates with Argonaute proteins to facilitate RNA interference in plants (Meister et al., 2005). When MOV10 expression is prevented, HDV replication is inhibited (Haussecker et al., 2008). Inhibition occurs at the level of transcription and not HDAg translation, and the presence of Argonaute-4 protein is also necessary for HDV transcription. The authors hypothesize that these components, both downstream effectors in the microRNA (miRNA) interference system, could function by remodeling HDV RNA upon cell entry to facilitate transcription initiation competency. These data suggest that HDV RNA could be involved in complex RNA regulatory pathways in host cells which are only now being characterized.

A potential human homologue of HDAg-L known as *delta* interacting protein A (DIPA) has also been identified (Brazas and Ganem, 1996). DIPA shares 24% amino acid identity and 56% sequence similarity with HDAg-L, although the ancestral relationship of DIPA and HDAg is controversial. Overexpression of DIPA results in attenuation of HDV replication, suggesting that this protein has a function similar to that of HDAg-L in HDV biology. DIPA has been shown to interact with various nuclear proteins to repress gene expression, acting as a transcriptional regulator (Bezy et al., 2005; Du et al, 2006). From

these observations, it is possible that DIPA plays role in HDV transcription and replication, however whether or not DIPA constitutes a true HDAg-L homologue remains uncertain.

One of the most important aspects of HDV biology is that HDV must interact with and utilize host DNA-dependent RNA polymerases (RNAPs) for transcription and replication since it does not encode an RdRp. This process involves interaction not only with RNAPs but other components of the host transcriptional machinery. The Handa laboratory discovered that HDAg bears weak homology (27% aa identity) to the A polypeptide chain of negative elongation factor (NELF), a transcriptional regulatory factor that interacts directly with RNAP II (Yamaguchi et al., 2001). NELF and DRB sensitivity-inducing factor (DSIF) repress transcriptional elongation until positive transcription elongation factor B (P-TEFb) alleviates this inhibition by phosphorylating both DSIF and the CTD of RNAP II (Kim and Sharp, 2001). It was thus speculated that HDAg might regulate transcription at the elongation step of RNA synthesis. While both HDAgs interact directly with DSIF, HDAg-S is the predominant antigen responsible of relieving DSIF/NELF-mediated transcriptional repression (Yamaguchi et al., 2001). *In vitro* transcription of RNA using nuclear extracts is repressed by addition of DSIF/NELF, but addition of HDAg reverses this repression and stimulates elongation in the absence of DSIF/NELF. Furthermore, addition of HDAg-S can stimulate elongation of an HDV-derived transcript which is prematurely terminated in the absence of HDAg-S. These observations describe a mechanism of action for HDAg-S-mediated enhancement of HDV RNA synthesis by RNAP II.

During the course of investigation of NELF, the Handa laboratory demonstrated the direct interaction of HDAg with RNAP II during transcriptional elongation (Yamaguchi et al., 2001; 2007). Several lines of evidence also show that RNAP II, and potentially RNAP I, interact with HDV RNA directly to facilitate HDV RNA synthesis. Interaction of the

eukaryotic polymerases RNAP I, RNAP II, and RNAP III with HDV RNA is the subject of the present investigation; RNAP involvement in the HDV life cycle is discussed further in Chapter 1.3.

1.2.2 Interaction of HDV RNA with host proteins

Many proteins have been shown to interact with HDV RNA. HDAg itself is an RNA binding protein; HDAg interacts with HDV RNA and is proposed to shuttle it to the nucleus for replication, to the cytosol for packaging and export, and to modulate its transcription and replication (reviewed by Lai, 2005 and Taylor, 2006). HDAg was first identified as an RNA-binding protein through demonstration of its ability to interact with genomic HDV RNA (Chang et al., 1988). Chromatin immunoprecipitation and ribonucleoprotein immunoprecipitation assays (ChIP and RIPA, respectively) have recently confirmed that both polarities of HDV RNA interact with HDAg in cells replicating HDV (Chang et al., 2008; Niranjanakumari et al., 2002). The N-terminal two-thirds of HDAg are extremely basic and contain two stretches of amino acids with similarity to the leucine zipper motif characteristic of nucleic acid-binding proteins, and it was demonstrated that HDAg binds to HDV-derived RNAs through this region (Lin et al., 1990). Interestingly, RNA binding occurs primarily at the central portion of HDAg and not the extreme N-terminus despite the similar amino acid composition in these regions. The HDAg RNA-binding region was localized to two arginine-rich motifs (ARMs) separated by 29 amino acids within the central region of HDAg; both ARMs, specifically the basic amino acids contained therein, are necessary for the RNA-binding activity of HDAg (Lee et al., 1993). Together, the ARMs constitute the principal RBD of HDAg. Binding of HDV RNA to the ARM regions appears to enhance HDV replication *in vivo* since RNA accumulation is impaired by deletion of these

motifs (Lee et al., 1993). Further studies of HDAg revealed a cryptic RBD located at the extreme N-terminal domain of HDAg between residues 2 and 27 (Poisson et al., 1993). This RBD is capable of binding both the genome and antigenome and was found to play a role in the ability of HDAg to act as a chaperone to modulate the activity of the *delta* ribozyme (Wang et al., 2003).

The binding activity of HDAg to HDV RNA is necessary for RNP particle formation during transcription and virion packaging (Ryu et al., 1993). Within an HDV-containing particle, approximately 70 molecules of HDAg are bound to a single molecule of genomic HDV RNA (Ryu et al., 1993). The antigenome is also present in HDV virions and is similarly bound by HDAg (Ryu et al., 1993). The central, ARM-containing region of HDAg demonstrates a high level of conservation among HDV isolates, possibly due to its role in HDV RNA binding and the necessity of this interaction for HDV virion assembly (Chao, Y.C. et al., 1990; Ryu et al., 1993).

The unbranched, rod-like secondary structure of HDV RNA is necessary for HDAg binding (Chao et al., 1991; Lin et al., 1990). Multiple HDV-derived RNAs bind HDAg while other rod-like RNAs such those found in viroids like the potato spindle tuber viroid (PSTVd) do not, indicating that HDAg could exhibit a preference for HDV RNA binding over other rod-like species (Chao et al., 1991). Double-stranded RNAs, such as those from the mouse hepatitis virus (MHV), are unable to bind HDAg, further attesting to the specificity of the HDV RNA-HDAg interaction (Lin et al., 1990). However, double-stranded RNA derived from the central region of HDV is capable of interaction with HDAg, and certain non-related RNAs such as those from the Rous sarcoma virus and hybridized, linear PSTVd dimers are capable of HDAg interaction, indicating that there is some flexibility in the HDAg binding requirements that enables the binding non-HDV RNA species (Chao et al., 1991). Since

HDV RNA binds HDAg, it is possible to extrapolate HDV RNA binding to host cellular proteins that bear similar structural motifs to those found on HDAg; these interactions might be critical to biological processes necessary for HDV replication and packaging.

The secondary structure of HDV RNA is attractive to RNA-binding proteins, and several host cellular proteins have recently been demonstrated to bind, and in some cases modify, HDV RNA. As aforementioned, modification of the HDV antigenome in the region encoding the HDAg-S mRNA is responsible for the production of HDAg-L. The UAG amber termination codon of HDAg-S is converted to UIG, where inosine is read as guanine, yielding the codon for tryptophan. Since modification occurs on the antigenome, or on an RNA that can be processed into the antigenome, the nucleotide conversion is retained in subsequent rounds of replication and is incorporated into the HDAg-S mRNA, resulting in translational readthrough which generates HDAg-L (Casey et al., 1992, Polson et al., 1996). This conversion of adenosine to inosine is carried out by a double-stranded RNA adenosine deaminase (Casey and Gerin, 1995; Polson et al., 1996). Specifically, the small isoform of adenosine deaminase acting on RNA (ADAR 1) has been shown to edit the HDV RNA antigenome (Wong and Lazinski, 2002). HDAg-S can repress this editing event, subsequently repressing the production of HDAg-L and regulating the stages of the HDV life cycle (Polson et al., 1998). ADAR 1 is thus an essential host protein necessary for the progression of HDV infection.

Another critical step in the replication of HDV is the autocatalytic cleavage of its RNA genome and antigenome by the *delta* ribozyme. It was determined that glyceraldehyde 3-phosphate dehydrogenase (GAPDH) interacts with the antigenome of HDV to enhance *delta* ribozyme activity (Lin et al., 2000). Although conventionally attributed to having a metabolic role in glycolysis, GAPDH was recently shown to exhibit nuclear translocation

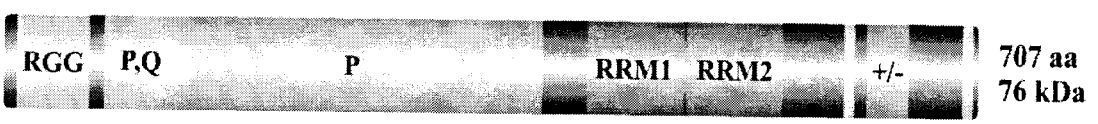
under specific circumstances such as drug treatment, hyperglycemic stress, cell entry into S-phase, and viral infection (reviewed by Sirover, 2005). In the nucleus, GAPDH has been found in association with complexes involved in such diverse functions as apoptotic cell death, DNA proofreading, nuclear fusion, and telomere maintenance (Sirover, 2005). Interestingly, GAPDH plays a role in transcriptional activation of histone H2B, indicating that this protein might be involved in RNA regulation (Zheng et al., 2003). Interaction of GAPDH with HDV RNA is localized to a small, U/C-rich portion of the antigenome corresponding to nucleotides 379 to 414 (Lin et al., 2000). The GAPDH-HDV interaction enables the shuttling of GAPDH to the nucleus and enhances *delta* ribozyme activity almost two-fold in agreement with a previous study which showed that GAPDH increases the rate of *cis*-cleavage of hammerhead ribozymes (Lin et al., 2000; Sioud and Jesperson, 1996). It is thus speculated that GAPDH enhances HDV replication by acting as a molecular chaperone that facilitates the unwinding of HDV RNA to enable autocatalytic cleavage during rolling circle replication.

As aforementioned, PKR is a protein that interacts with and phosphorylates HDAg (Chen et al., 2002). It has also been shown that the genome, antigenome, and subgenomic fragments of HDV RNA bind and activate PKR (Robertson et al., 1996). This observation is unusual since PKR activation typically occurs in the presence of extensively base-paired, dsRNA; binding of highly-structured, single-stranded RNA (ssRNA) such as viral RNA usually inhibits PKR activity (Manche et al., 1992). When activated, PKR mediates the interferon-induced antiviral response and acts as a tumor suppressor by inhibiting cellular translation; paradoxically, while HDV RNA activates PKR, it fails to inhibit protein synthesis in cell-free translation systems (Manche et al., 1992; Robertson et al., 1996). Binding of specific viral ssRNA structures suppresses PKR activation to facilitate viral

propagation, and it can be speculated that HDV interaction with PKR might serve this purpose despite its non-conventional method of translational inhibition (Robertson et al., 1996). It is possible that interaction of HDV RNA with PKR mediates its recruitment to RNP complexes containing HDAg so that PKR can phosphorylate HDAg, although this theory has yet to be tested (Chen et al., 2002; Robertson et al., 1996). Further studies of this interaction will be required to determine the precise role of PKR in the HDV life cycle.

One of the main objectives of my thesis research was to identify novel proteins that interact with HDV RNA and, where possible, propose a role for these interactions in the HDV life cycle. The polypyrimidine tract-binding protein (PTB)-associated splicing factor (PSF) is a multifunctional nuclear protein which I identified as an HDV RNA-interacting partner both *in vitro* and in cells replicating HDV. First identified by Patton and associates (1993) as a component of the PTB-containing complex, PSF is a 76 kDa multifunctional protein involved in many nuclear processes in addition to post-transcriptional mRNA modification (reviewed by Shav-Tal and Zipori, 2002). PSF is comprised of an N-terminal RGG motif, a P/Q-rich region, two canonical RNA recognition motifs (RRMs), and two NLS sequences flanking a highly-charged C-terminal region (Shav-Tal and Zipori, 2002; Figure 1.5). The N-terminal domain of PSF is responsible for protein-protein interactions and possesses DNA-binding activity, while the C-terminal region is responsible for binding RNA through both RRMs. PSF can act independently or form a heterotetramer with the 54 kDa nuclear RNA-binding protein ($p54^{nrb}$), a protein that bears significant homology to the CTD of PSF. PSF facilitates a number of nuclear activities including splicing, polyadenylation, transcriptional regulation, retention of defective RNAs, nucleic acid unwinding and annealing, nuclear shuttling, and pH homeostasis (Shav-Tal and Zipori, 2002). To enable these diverse functions, PSF demonstrates diffuse localization throughout the nucleus and is

Figure 1.5: Schematic representation of the polypyrimidine tract binding protein-associated splicing factor (PSF). PSF is a multifunctional nuclear protein involved in processes such as transcriptional regulation, pre-mRNA processing, DNA repair, and RNA retention. It is comprised of an RGG motif, a proline (P) and glutamine (Q) rich domain involved in protein-protein interaction, two canonical RNA recognition motifs (RRMs), a highly-charged region (+/-), and two nuclear localization signals (yellow). Adapted from Shav-Tal and Zippori (2002).



Adapted from Shav-Tal and Zipori (2002)

also found in association with the nucleolus, nuclear membranes, and paraspeckle subdomains (Shav-Tal and Zipori, 2002). Interestingly, it has been shown that PSF interacts with the CTD of RNAP II during both transcription initiation and elongation, and that PSF can likely interact with RNA and the CTD of RNAP II simultaneously (Emili et al., 2002). These observations demonstrate that PSF could be integral in the coupling of transcription and splicing and have led to my speculation that PSF might provide a physical link between HDV RNA and RNAP II. Due to its RNA-binding capacity, interaction with RNAP II, and involvement in numerous nuclear processes, establishing a role for PSF in the HDV life cycle might provide novel insights into the mechanism of HDV replication and, potentially, the mechanism of HDV pathogenesis.

The rapidly increasing number of host proteins that interact with HDAgs and HDV RNA is testament to the complexity of the host-pathogen interaction upon HDV infection. Such interactions might enable HDV propagation, exert pathogenic effects on the host, or be completely benign; it is thus imperative to determine which host cellular proteins interact with HDV following infection and determine their roles in the HDV life cycle and pathogenesis. Our laboratory is continuing to elucidate and characterize new HDV RNA-interacting partners, and through the course of my investigations novel HDV-interacting proteins have been identified and suspected candidates confirmed. Through identification of HDV-interacting factors and determination of their roles in HDV biology it will become possible to develop novel strategies to combat HDV infection. More generally, characterization of the manner in which HDV RNA is capable of replication and transcription by coercion of host proteins, specifically DNA-dependent RNA polymerases, will provide global insight into polymerase function and potentially reveal novel pathways of RNA synthesis, metabolism, and regulation.

1.3 RNA-templated transcription by DNA-dependent RNA polymerases: transcription and replication of the hepatitis *delta* virus

Traditionally, DNA-dependent RNA polymerases such as viral polymerases, the bacterial RNAP, and eukaryotic RNAPs I, II, and III were thought to transcribe exclusively from DNA templates. However, the discovery of infectious RNA agents that do not encode polymerases, such as viroids and HDV, has challenged this fundamental principle of molecular and cellular biology. *How is it possible that an RNA species replicates and transcribes RNA from an RNA template in the absence of an RNA-dependent RNA polymerase?* This conundrum inspires RNA biochemists to address many questions related to RNA-templated RNA synthesis including which RNAP(s) are involved in RNA-templated transcription and which RNA elements comprise a promoter that can be utilized in a non-canonical manner by RNAPs.

Examples of RNA-templated RNA synthesis can be found in viral, bacterial, and eukaryotic systems, although the significance of these interactions remains poorly understood. RNAP from the T7 bacteriophage is capable of specific, RNA-templated synthesis *in vitro*, although the physiological relevance of this phenomenon is unknown (Konarska and Sharp, 1989). In bacteria, an RNA species known as 6S was recently shown to direct RNA-templated RNA synthesis by bacterial RNAP in order to regulate gene expression in response to changing environmental conditions (Wassarman, 2007; Wassarman and Stortz, 2000). Both of these polymerases are also capable of template-free RNA synthesis and synthesis of RNA from random RNA oligonucleotide polymers (Biebricher and Luce, 1996; Biebricher and Orgel, 1973; Wettich and Biebricher, 2001). Subviral plant pathogens such as viroids do not encode the RdRp necessary for their transcription and replication and are thus reliant on host RNAPs, such as RNAP II, the phage-like nuclear

encoded polymerase (NEP), and bacteria-like plastid encoded polymerase (PEP) for survival (reviewed in Daròs et al., 2006 and Tsagris, 2008). These observations demonstrate that RNA-templated RNA transcription is a phenomenon that transcends polymerases of highly-divergent backgrounds, and this unconventional use of RNAPs might offer insight into novel RNA-mediated cellular pathways and provide clues to the evolution of modern RNAPs from RNA-dependent ancestors (discussed further in Chapter 7.1).

1.3.1 The eukaryotic DNA-dependent RNA polymerases

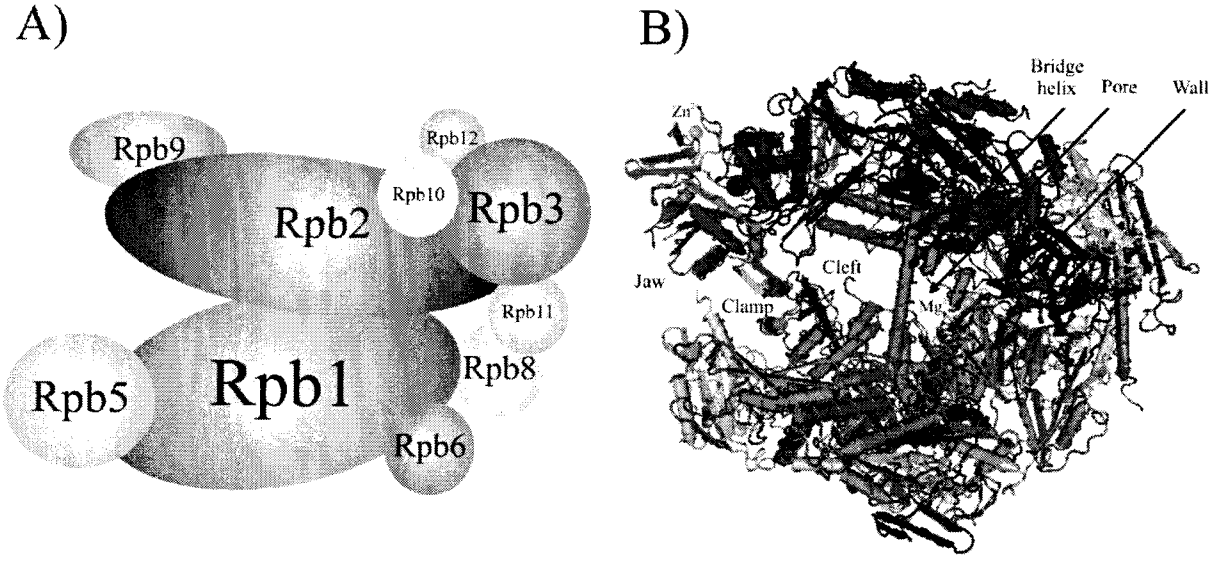
The present study is concerned with examination of how HDV, the only known mammalian subviral RNA pathogen, subverts host polymerases and nuclear proteins to facilitate RNA synthesis. Since HDV does not encode a polymerase, one or more host polymerases must be responsible for synthesis of the HDV genome, antigenome, and mRNA. In eukaryotic systems, transcription from B-DNA templates is carried out by three RNAPs, RNAP I, RNAP II, and RNAP III, which are typically associated with synthesis of rRNA, mRNA, and transfer RNA (tRNA), respectively. These three polymerases are comprised of 14, 12, and 17 subunits, respectively, and their cores and active catalytic centers are highly conserved (reviewed by Cramer, 2002; Cramer et al., 2008). It has been shown that 22 regions of homology exist among the core subunits of bacterial and eukaryotic RNAPs, and that polymerase structure is well conserved despite differences in primary amino acid sequence among homologous RNAP subunits (Cramer, 2002). In the last decade, much information regarding the composition and structure of these polymerases has emerged, providing insight into the similarities and differences among these enzymes.

RNAP II is the most well characterized eukaryotic polymerase despite its high level of complexity. The RNAP II transcriptional machinery contains nearly 60 polypeptides while

the other polymerases require relatively few (reviewed by Hahn, 2004). It is composed of a 10 subunit core enzyme containing Rpb1, Rpb2, Rpb3, Rpb5, Rpb6, Rpb8, Rpb9, Rpb10, Rpb11, and Rpb12, and a peripheral heterodimer comprised of Rpb4 and Rpb7 (Cramer et al., 2001; Figure 1.6). Homologues of Rpb1, Rpb2, Rpb3, and Rpb11 are found in all three eukaryotic polymerases and in those of bacteria and archaea, while Rpb5, Rpb6, Rpb8, Rpb10, and Rpb11 are completely conserved among RNAP I, RNAP II, and RNAP III (Cramer et al., 2008). General transcription factors (GTFs) are also required for promoter binding and nucleation of the pre-initiation complex (PIC), the first step in RNA synthesis in which the transcriptional machinery is assembled in an inactive conformation at the promoter element upstream from the initiation site (Hahn, 2004). GTFs enable recognition of specific promoter features, response to additional regulatory factors, and conformational changes necessary for RNAP II activity (reviewed by Hahn, 2004 and Lee and Young, 2000). Each RNAP has a unique set of GTFs, although some of them are homologous or shared between RNAPs. For example, the TATA-binding protein (TBP) is a common component of GTFs used by RNAPs I, II, and III, and is required for RNA synthesis by all three RNAPs (Cormack and Struhl, 1992). Additional gene-specific regulatory factors, such as those that activate or repress gene expression, serve to finely-tune the transcription process and precisely regulate transcription in response to different temporal and environmental cues (Hahn, 2004). Transcription by RNAP II thus involves a dynamic interaction between nucleic acids and proteins and has consequently been a compelling area of research which has grown substantially within the last decade.

Conformational change of the RNAP II complex following PIC assembly causes unwinding of 11 to 15 base pairs of DNA around the transcription start site, allowing entry of the single-stranded DNA template into the active site of the polymerase, leading to

Figure 1.6: The RNAP II core. A) Schematic representation of the 10 subunit core of RNAP II. Adapted from Cramer (2002). B) Structure of RNAP II core. Subunit colours correspond to (A). The peripheral Rpb4/Rpb7 heterodimer is not shown; it is located beneath the Rpb1 clamp. Components of the active site are indicated with text and arrows. Crystal structure resolved by Cramer et al. (2001; PDB accession number 1I50).


A)
Rpb9
Rpb12
Rpb10
Rpb2
Rpb3
Rpb11
Rpb5
Rpb1
Rpb8
Rpb6
B)
Bridge helix
Pore
Wall
Zn
Jaw
Cleft
Clamp
Mg

transcription initiation (Hahn, 2004). The RNAP II core contains the two largest RNAP II subunits, Rpb1 and Rpb2, which form a deep, acidic cleft into which incoming template DNA is inserted (Cramer et al., 2008; Hahn, 2004). The remaining subunits comprising the core assemble around the periphery of the Rpb1/Rpb2 cleft (Cramer et al., 2008; Hahn, 2004; Figure 1.6). Rpb1 contains a mobile clamp region and bridge helix that spans the cleft region between Rpb1 and Rpb2 (Cramer et al., 2001). Rpb2 is comprised of lobe and protruding domains that form a wall to block the end cleft (Cramer et al., 2001). The active site of RNAP II lies in the base of the cleft at the center of the polymerase and is formed between the clamp, bridge helix, and wall of Rpb2 (Cramer, 2004). Beneath the active center lies a pore that widens towards the outside of the enzyme; a region of the pore formed by loops of Rpb1 and Rpb2 contains two metal ions, termed A and B, which are necessary for RNA polymerization (Cramer, 2004; Cramer et al., 2008; Hahn et al., 2004). The active site also has endogenous RNA cleavage activity that can be stimulated by the GTF TFIIS; this ability facilitates escape from transcriptional arrest and mediates proofreading (Powell et al., 1996). As elongation commences, RNAP II loses contact with components of the PIC which remain at the promoter to indicate regions that have been transcribed and to provide a scaffold for the nucleation of PICs for upcoming rounds of transcription (Hahn, 2004, Yudkovsky et al., 2000).

The mobile clamp element is a structural component of Rpb1 that is necessary for stabilizing the DNA-RNA hybrid during elongation, enabling processivity of the enzyme (Cramer, 2004; Hahn 2004). It is connected to the core through a series of five flexible switches and exists in several conformations (Gnatt et al., 2001). The Rpb4/Rpb7 heterodimer serves as wedge to maintain the clamp in the closed position prior to initiation, but the clamp must then loosen to permit template entry (Cramer, 2001; Gnatt et al., 2001).

During elongation, the clamp rotates over the acidic cleft to keep the DNA template and nascent transcript in place and is considered to adopt a closed conformation (Cramer et al., 2001; Gnatt et al., 2001). Interaction with clamp switches and the bridge helix stabilize and position the nucleic acid-polymerase interaction. During elongation, an eight nucleotide duplex of DNA-RNA is found at the core of the polymerase bounded by the bridge helix and wall of the Rpb2 subunit (Gnatt et al., 2001; Westover, 2004). Three loops, the lid, rudder, and fork, protrude from the clamp and facilitate polymerase processivity (Gnatt et al., 2001; Westover, 2004). The lid separates the DNA-RNA hybrid and guides the nascent RNA towards the exit grove, the rudder interacts with the single-stranded template DNA to prevent re-association with nascent RNA, and the fork stabilizes the hybrid duplex (Westover, 2004). During RNA synthesis, a nucleoside triphosphate (NTP) enters and binds the elongation complex, becomes incorporated at the 3' end of the nascent mRNA (releasing a pyrophosphate ion), and causes translocation of the DNA-RNA hybrid to facilitate incorporation of the next nucleotide (Cramer et al., 2008). This basic mechanism is employed by all known RNAPs.

Another interesting feature of RNAP II is the extended CTD of Rpb1 which couples transcription with post-transcriptional RNA processing. The CTD is attached to Rpb1 by a flexible linker helix located near the exit pore of the polymerase and makes contacts with the peripheral Rpb4/Rpb7 complex (reviewed by Meinhart et al., 2005). Human Rpb1 contains a CTD comprised of 52 heptapeptide repeats of the consensus sequence YSPTSPS. Although variable in number, this repeat is conserved among eukaryotic polymerases and the number of heptad repeats is correlated with genome complexity. All seven amino acids of the heptad repeat are subject to dynamic and reversible modification: phosphorylation of tyrosine, threonine, and all serine residues, isomerization of both proline residues, and glycosylation

of threonine and serine residues have been reported (reviewed by Egloff and Murphy, 2008). Serine phosphorylation is the most common modification of the CTD, and the position of phosphorylated serine residues dictates which transcription and transcript-processing factors are associated with the polymerase complex at specific points in the transcription cycle. Modifications of the serine residues at positions two and five of the heptad repeat have been extensively studied and have been shown to be integral to events such transcriptional elongation and 5' methylguanosine capping, respectively; dephosphorylation of these residues has been implicated in polymerase recycling (Komarnitsky et al., 2000). The CTD is hypophosphorylated upon promoter binding and transcription initiation, and transition to the elongation phase of transcription is associated with hyperphosphorylation of the CTD (Meinhart et al., 2005). Phosphorylation of the CTD has been demonstrated to mediate the association of RNA-modifying factors such as those associated with capping, splicing, cleaving, and polyadenylation; pre-mRNA modification begins prior to completion of RNA synthesis and occurs concurrently with transcription (Egloff and Murphy, 2008; Meinhart et al., 2005). The unique pattern of phosphorylation of CTD residues, also known as the "CTD code", has been implicated in additional events such as chromatin remodeling, DNA repair, and RNA editing, packaging, and nuclear export (Egloff and Murphy, 2008). Many proteins with the capacity to modify the CTD of RNAP II exist, but the full extent of CTD modification and subsequent modulation of RNAP II activity is currently the subject of investigation.

RNAP I and RNAP III have not been as well characterized as their counterpart, RNAP II. Although the core elements and active centers of these RNAPs are very similar, small variations in amino acid composition can significantly modify their properties, such as susceptibility to inhibitors of transcription (Bushnell et al., 2001). While the ten core

subunits of the three RNAPs are conserved to a large extent, certain structural differences give rise to gene-specific functions of these three enzymes (reviewed by Cramer, 2008). Although the existence of a peripheral heterodimer such as Rpb4/Rpb7 is maintained, little homology exists between these subunits, and RNAP I and RNAP III contain two and five additional subunits, respectively.

RNAP I is responsible for rRNA synthesis, the main component of ribosome biogenesis, and is consequently considered to be an indicator of cell growth. In yeast, RNAP I-mediated transcription accounts for 60% of all RNAP activity and rRNA is the most abundant cellular RNA with over 80% of total RNA being comprised of this species (Warner, 1999). As aforementioned, RNAP I contains 14 subunits; the three RNAPs have five shared subunits, and RNAP I and RNAP III additionally share two subunits not found in RNAP II. Rpb1 and Rpb2 are replaced with A190 and A135 homologues in RNAP I, and the A12.2 subunit of RNAP I contains endogenous cleavage stimulatory activity found in the TFIIS GTF of RNAP II (Kuhn et al., 2007). Another unique heterodimer comprised of A49 and A34.5 is located over the acidic cleft and resembles components of TFIIF. This A49/A34.5 heterodimer is necessary for elongation, demonstrating that some indispensable GTFs are basal components of the RNAP I complex and their presence could enhance the processivity of this polymerase. Finally, a complex similar to the Rpb4/Rpb7 heterodimer, A14/A43, forms a smaller stalk-like component and provides a unique, promoter-specific interface for this RNAP. Although RNAP I and RNAP II demonstrate a high amount of sequence divergence, the overall structure of their core and active site are conserved. Some structural features, such as the clamp knob and clamp head, are not seen with other polymerases, as are various extensions and protrusions of many of the subunits, revealing RNAP I-specific structural interfaces that might result in gene-specific interactions. Since the

upstream interface of RNAP I is significantly different from those of other RNAPs, it might serve to facilitate unique interactions with RNAP I-specific promoter elements.

The tertiary structure of the RNAP III core polymerase was also recently determined (Fernandez-Tornero et al., 2007; Jasiak et al., 2006). Although the core is structurally similar to the other two RNAPs, RNAP III contains five RNAP III-specific subunits that form two subcomplexes: a three subunit complex involved in initiation and a two subunit complex involved in termination and re-initiation. These subcomplexes provide an upstream face that is unique from those presented by RNAP I and RNAP II, providing insight into how this polymerase might recognize its specific promoter targets. These data indicate that the three polymerases operate in a manner that is mechanistically similar and evolutionarily conserved, although slight structural differences in the core, the existence of different accessory subunits, and variation among structures and surfaces that bind promoter elements cause the diversity in genetic elements recognized by these polymerases and are responsible for differences in the genes that they transcribe.

Given the observations that much homology exists among the ten core subunits of these polymerases and the fact that HDV is not a conventional template for RNAP-mediated transcription, it is feasible that one or more eukaryotic RNAPs are capable of transcribing from an HDV template. The following discussion will address evidence for the non-canonical use of RNA templates in RNAP-mediated synthesis of this unique viral pathogen.

1.3.2 Products of HDV RNA synthesis

The potential for multiple host polymerase involvement in the life cycle of HDV is suggested by the fact that at least three final and distinct products of RNA-templated RNA synthesis are required for HDV propagation: the genome, the antigenome, and the HDAg mRNA. One

hypothesis is that HDAg mRNA is transcribed first due to the use of HDAg-S in HDV RNA synthesis early in the life cycle of HDV (Taylor, 1999). Interruption of symmetrical rolling circle replication due to the presence of a polyadenylation signal on the antigenomic RNA could occur, generating the small, poly(A)-tailed transcript of HDAg, presumably by RNAP II. However, it has also been shown that transcription of HDAg mRNA can occur concurrently with and independently of replication, signifying the potential for different temporal and spatial localization of these processes or utilization of more than one host polymerase (Modahl and Lai, 1998).

Additional HDV-directed transcripts have recently been discovered, adding a further layer of complexity to the HDV life cycle. A third isoform of HDAg, referred to as peptide K, has been found in subpopulations (> 2%) of cells replicating HDV (Bichko and Taylor, 1996; Khudyakov and Makhov, 1990). The peptide K ORF occurs at the -1 position relative to the HDAg ORF which gives rise to a peptide of 24 to 63 amino acids due to variability of the location of the termination codon (Bichko and Taylor, 1996; Khudyakov and Makhov, 1990). The peptide K ORF lacks an in-frame initiator codon and thus relies on mutation or frameshift error for mRNA transcription (Bichko and Taylor, 1996). Subcellular localization of peptide K is similar to that observed for HDAg-S and HDAg-L, though it is unlikely to be necessary for HDV survival due to its low prevalence in infected cells (Bichko and Taylor, 1996).

Two small RNA species were recently found in cells replicating HDV (Haussecker et al., 2008). An antigenomic species of 24 to 25 nt was isolated whose location coincides with the start site of HDAg mRNA synthesis, while another species of 18 to 19 nt was recovered corresponding to the same location on the genomic polarity. Although it is common for RNAPs to synthesize several small, abortive RNA oligonucleotides prior to promoter escape

and elongation of a *bona fide* transcript, these HDV-derived small RNAs are longer than typical abortive products and were consistently observed throughout the course of experimentation (Haussecker et al., 2008; Hsu et al. 2003; Luse and Jacobs, 1987). Several other small, HDV-derived species of both polarities were also found, and it is hypothesized that they might serve as small priming RNAs (spRNAs) for the synthesis of the HDV genome and antigenome (Haussecker et al., 2008). These small RNAs would thus serve a vital function in HDV biology and would constitute veritable HDV-directed transcripts. Since there are at least three other necessary products of HDV replication and transcription and several circular and linear products of various lengths exist in infected cells, multiple mechanisms of HDV RNA synthesis might exist involving various intermediates and utilizing different host polymerases.

1.3.3 Evidence for RNAP II-directed synthesis of HDV

After discovery of HDV in the liver of HBV-infected patients and its subsequent molecular characterization, several lines of evidence began to emerge implicating RNAP II as the polymerase principally responsible for HDV RNA synthesis. Like members of the *Pospiviroidae* family of viroids, HDV RNA accumulates in the nucleus of infected cells. Pospiviroid replication is proposed to be performed by host nuclear RNAP II based on physical association with RNAP II, ability to be transcribed by this polymerase *in vitro*, and sensitivity of this transcription to α-amanitin (Goodman et al. 1984; Mühlbach and Sänger, 1979). The cyclic peptide α-amanitin is a mycotoxin produced by members of the *Amanita* family of poisonous mushrooms that specifically inhibits DNA-dependent RNA transcription by binding directly to the polymerase (Bushnell et al., 2002; Wieland and Faulstich, 1991). RNAP II is extremely sensitive to this compound since α-amanitin lodges under the bridge

helix between the Rpb1 and Rpb2 subunits and forms hydrogen bonds with the structure, buttressing the bridge helix and preventing the conformational change necessary for translocation (Bushnell et al., 2002). In eukaryotes, RNAP II is sensitive to extremely doses of α-amanitin (10^{-9} M to 10^{-8} M), whereas RNAP III is only sensitive to higher doses (10^{-5} M to 10^{-4} M) and RNAP I is almost completely resistant (10^{-3} M; reviewed by Chambon, 1975). Sensitivity to α-amanitin is thus a presumptive indicator of RNAP II-mediated transcription. Due to the similar genomic composition and nuclear localization of pospiviroids and HDV, RNAP II was suspected to be the host polymerase responsible for HDV RNA synthesis. Indeed, HDV RNA accumulation in infected cells was shown to be sensitive to α-amanitin, and HDV synthesis could be restored in the presence of α-amanitin in cells harbouring an α-amanitin-resistant allele of RNAP II (Fu and Taylor, 1993; Macnaughton et al, 1991; Modahl et al., 2000). Based on these observations, RNAP II became the preferred candidate polymerase for HDV RNA synthesis.

Certain forms of post-transcriptional modification are unique to RNAP II transcripts. HDAg mRNA is post-transcriptionally modified with a 5' methylguanine cap and 3' poly(A) tail; these features are hallmark characteristics of RNAP II transcripts (Gudima et al., 2000; Hseih et al., 1990; Nie et al., 2004). Additionally, the recently-discovered HDV spRNAs are predominantly modified with a 5' cap (Haussecker et al., 2008). Since neither cap nor tail structures are associated with RNAP I and RNAP III transcripts, it would appear likely, from these observations, that transcription of these HDV RNAs is mediated by RNAP II.

Additional support for RNAP II-directed synthesis of HDV RNA comes from the observation that HDAg-S co-localizes with SC35, an essential non-small nuclear ribonucleoprotein splicing factor, in speckles in the nucleoli (Bichko and Taylor, 1996). SC35-containing nucleolar subdomains are active sites of RNAP II transcription, and co-

localization with HDAg-S, which is used in HDV RNA synthesis, would suggest RNAP II involvement. The discrete localization of HDAg-S in nuclear speckles changes throughout the course of HDV infection, corresponding to the shift from active HDV transcription involving HDAg-S to transcriptional repression by HDAg-L later in HDV life cycle, further suggesting RNAP II involvement in these processes.

HDAg-S has been shown to be necessary for HDV RNA synthesis and accumulation in HDV-infected cells (Kuo et al., 1989). The Handa laboratory has recently demonstrated that HDAg-S binds directly to RNAP II and promotes translocation of the RNAP II complex during transcriptional elongation, supporting involvement of RNAP II in HDV RNA synthesis (Yamaguchi et al., 2001; 2007). HDAg-S can bind both the hypo- and hyperphosphorylated forms of RNAP II found during transcriptional initiation and elongation, respectively. Specifically, HDAg-S binds the mobile clamp element of RNAP II, and makes direct contact with the Rpb1 and Rpb2 subunits (Yamaguchi et al., 2007). The C-terminal amino acids 130-195 of HDAg-S are responsible for the HDAg-S-RNAP II interaction; deletion of only eight amino acids from the C-terminus of HDAg-S impairs the ability of HDAg-S to stimulate transcriptional elongation. HDAg-S residues that are conserved among HDV isolates such as D187, I188, and F194 are also necessary for the HDAg-S-RNAP II interaction, as is phosphorylation of S177 (Chen et al., 2008; Yamaguchi et al., 2007). Mechanistically, HDAg-S can compete with the transcription inhibitor NELF for RNAP II binding, and HDAg-S binding to RNAP II relieves NELF-mediated transcriptional repression and stimulates elongation (Yamaguchi et al., 2001). HDAg-S also accelerates forward translocation of the polymerase complex; this presumably occurs by loosening the clamp module, causing a concurrent loss of transcriptional fidelity (Nedialkov et al., 2003; Yamaguchi et al. 2007). Since binding of HDAg-S affects the recognition of

incoming nucleotides and influences the choice of nucleotide incorporated into the transcript, it is possible that template recognition (i.e. RNA versus DNA) is also affected, providing a mechanism by which RNAP II is capable of utilizing this HDV RNA template. The involvement of HDAg-S in these specific processes indicates a role for HDV in RNAP II-mediated transcription, but does not exclude other polymerase involvement in HDV biology.

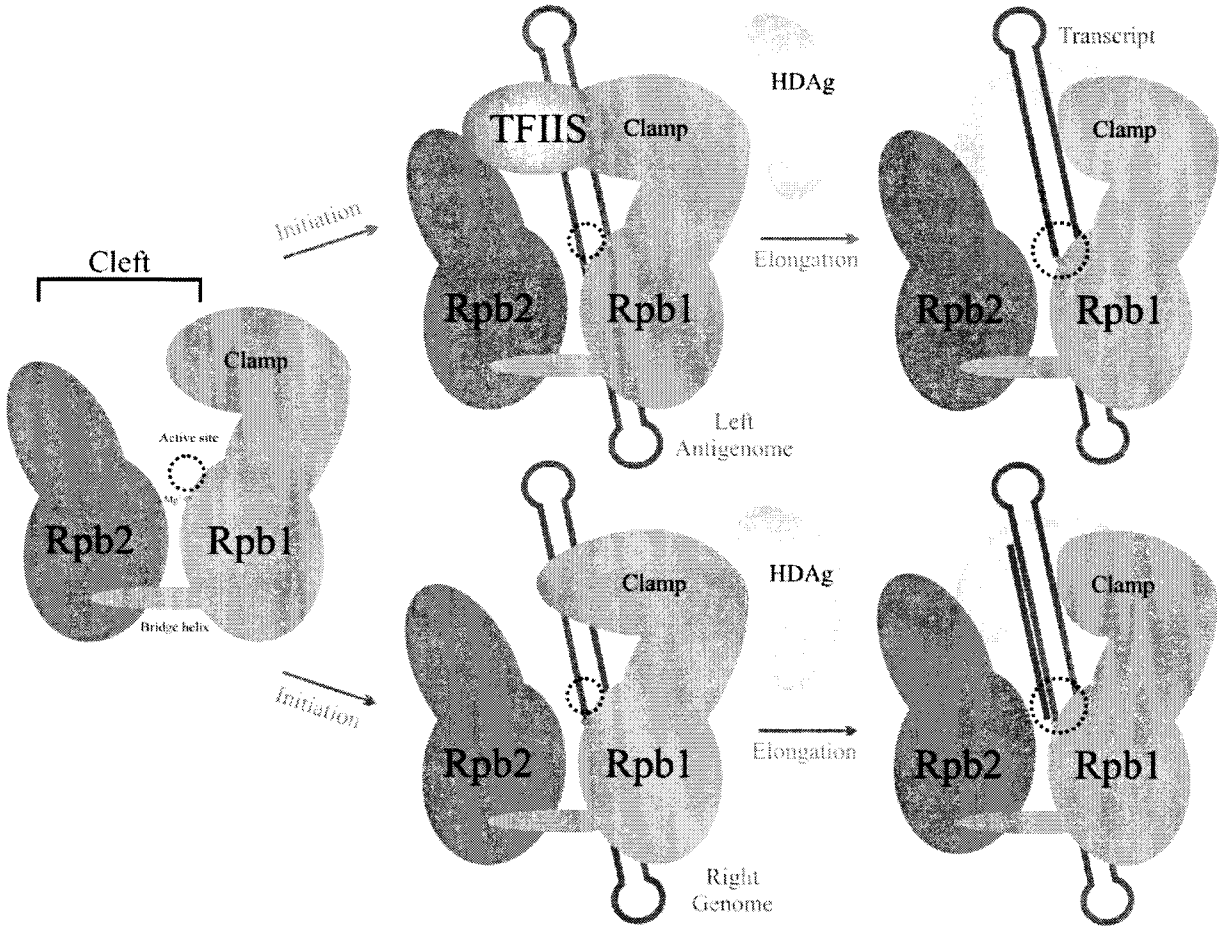
In vitro transcription systems using HeLa cell nuclear extract (NE) also support a role for RNAP II in HDV RNA synthesis. Beard and associates (1996) used a segment of the HDV genome corresponding to a 199 nt portion of the right terminal stem-loop domain of genomic polarity (hereafter referred to as R199G; Figure 1.1B) to transcribe an RNA product using this *in vitro* system. A product of approximately 100 nt was produced by *de novo* synthesis from R199G. Primer extension experiments localized the initiation site to the region of nucleotide A_{1631} which lies immediately upstream of the presumptive initiation site for HDAg mRNA synthesis at nucleotide U_{1630}. Later experiments would further refine the transcription start site to U_{1630}, causing the first transcribed nucleotide to be adenosine, a common feature of RNAP II transcripts (Gudima et al., 1999). Recent *in vitro* transcription experiments further demonstrate transcriptional initiation from the R199G region of HDV RNA and show that a functional PIC can form on this RNA species (Abrahem and Pelchat, 2008). Together, these observations strongly support the existence of a promoter element located at the tip of the right terminal stem-loop domain of the HDV genome.

Evidence supporting the existence of an additional promoter region located on the HDV antigenome has also been reported. Filipovska and Konarska (2000) used the *in vitro* transcription method to characterize a putative promoter element comprised of 280 nt of the left terminal stem-loop domain of antigenomic polarity. Transcription of this species is sensitive to α-amanitin, and transcription in the presence of α-amanitin can be partially

restored using NE prepared from cells containing an α-amanitin-resistant allele of RNAP II. The transcript generated in this system was a chimeric species comprised of the 5' end of the HDV terminal hairpin template joined with the 3' end of nascent transcript. It would thus appear that transcription did not occur by *de novo* RNAP II-directed synthesis but by a cleavage/primer extension reaction. These observations were corroborated by data from Lehmann and colleagues (2007) who recently elucidated the crystal structure of purified RNAP II engaged in transcription from an RNA species derived from the extreme tip of the left terminal stem-loop domain of the HDV antigenome. In these experiments, RNA-dependent transcription by RNAP II was shown to be template-driven and efficient, although slower and less processive than that observed with a canonical DNA template. The GTF TFIIS is speculated to enhance cleavage of the terminal stem-loop region to facilitate extension from the new 3' end. While this portion of the left terminal stem-loop domain of the HDV antigenome can serve as a template for RNAP II *in vitro*, its role in HDV replication *in vivo* remains controversial since transcription from this region proceeds by a non-conventional 3' end extension mechanism (Filipovska and Konarska, 2000; Lehmann et al., 2007). Furthermore, previous studies had demonstrated that HDV RNA can accumulate in cells in the absence of this region (Filipovska and Konarska, 2000; Netter et al., 1993). Additional studies will thus be required to determine whether or not this region of the HDV antigenome constitutes a *bona fide* promoter element used in HDV replication.

In summary, *in vitro* transcription experiments suggest that RNAP II is responsible for HDV transcription, that transcription by RNAP II might occur from genomic and antigenomic templates using two different promoters, and that transcription of short RNA species from an HDV template can occur in the absence of HDAg-S *in vitro*. Potential mechanisms of RNAP II-directed HDV synthesis are shown in Figure 1.7. Though

Figure 1.7: Interaction of HDV with RNAP II. Two putative promoters have been identified for RNA-templated synthesis by RNAP II. For left terminal antigenome synthesis, TFIIS cleaves the HDV RNA template and a chimeric product is synthesized by RNAP II by extension of the template RNA (Filipovska and Konarska, 2000; Lehmann et al., 2007). For the right terminal genome, cleavage of the HDV template does not occur and transcription occurs by *de novo* synthesis (Beard et al., 1996). HDAg binds the clamp region of the Rpb1 subunit of RNAP II and loosens the clamp to facilitate elongation (Yamaguchi et al., 2001; 2007).


Cleft
Clamp
Active site
Mg
Rpb2
Rpb1
Bridge helix
Initiation
TFIIS
Clamp
Rpb2
Rpb1
HDAg
Elongation
Transcript
Clamp
Rpb2
Rpb1
Left
Antigenome
Initiation
Clamp
Rpb2
Rpb1
HDAg
Elongation
Clamp
Rpb2
Rpb1
Right
Genome

intriguing, these data can neither confirm the use of both putative promoters *in vivo* for HDV RNA synthesis nor exclude the existence of other promoters scattered across the HDV genome, particularly given the non-canonical nature of transcription from the antigenomic template. Furthermore, these data do not address the possibility of the involvement of additional cellular polymerases in HDV biology.

1.3.4 The template switching hypothesis

Although HDV RNA acquires a circular, rod-shaped structure, it has been observed that the linear HDV species of both genomic and antigenomic polarity can serve as templates that ultimately yield circular products (Chang and Taylor, 2002). However, the linear template must be greater than unit length to initiate replication, and the circular template is much more efficient than the linear concatamers (Gudima et al., 2004). Due to the ability of the host polymerase to use linear RNA as a template, a phenomenon known as **template switching** has been proposed for HDV RNA synthesis in addition to the symmetrical rolling circle mechanism (Chang and Taylor, 2002). Template switching is commonly used by viral RdRps and reverse transcriptases (RTs) as a method of homologous recombination since it involves discontinuous transcription across a gap in templates, enabling two linear templates to yield a single, recombinant product. In the case of HDV, synthesis of the antigenome can occur across the gap in the folded genomic linear transcript when it assumes its rod-shaped structure (Chang and Taylor, 2002). Template switching could also explain the mechanism by which HDV mutants harboring large inserts of foreign RNA that cause reduced replication fitness are capable of ejecting these inserts (Gudima et al., 2006; Wang et al., 1997). Proponents of the template switching hypothesis believe that RNAP II is the cellular polymerase capable of performing this procedure although, to date, only RNA-dependent

polymerases such as viral RdRps and RTs have been shown to have the capacity to perform template switching.

It is further speculated that HDV might use template switching as a mechanism of homologous recombination between circular monomers (Chao, 2007). It is proposed that the polymerase locally unwinds the rod-like structure of HDV template, generating small stem-loops that cause stalling of the polymerase. Dissociation of the polymerase-nascent transcript complex from the template occurs, and the complex associates with a secondary HDV template and continues synthesis of a chimeric product. HDV might thus use template switching as not only a mechanism for utilizing linear templates but also as a means of increasing genetic diversity among isolates. While template switching might occur concurrently with symmetrical rolling circle replication in HDV RNA synthesis, it cannot be conclusively stated, based on these observations alone, that RNAP II is the sole polymerase responsible for these methods of replication.

1.3.5 Evidence for the involvement of additional RNAPs in HDV RNA synthesis

Much data has recently emerged regarding the involvement of additional host polymerases, specifically RNAP I, in the HDV lifecycle. The observation that replication of HDV, under some circumstances, is resistant to high levels of α-amanitin has prompted further investigation into the involvement of RNAP I in HDV replication (Macnaughton et al., 2002; Modahl et al., 2000). The Lai laboratory has performed several lines of experimentation to corroborate these findings (Chen et al., 2008; Li et al., 2006). Metabolic labeling experiments have demonstrated that the genomic and antigenomic strands of HDV RNA are transcribed in separate nuclear compartments: antigenome synthesis takes place in the nucleolus and is largely resistant to inhibition by to α-amanitin, whereas genome synthesis

occurs throughout the nucleoplasm and is sensitive to α-amanitin (Li et al., 2006). Furthermore, HDAg-S co-localizes with PML bodies during transcription of the genome but not the antigenome. PML bodies are dynamic subnuclear domains associated with numerous cell functions such as cell cycle regulation, apoptotic cell death, and antiviral response through modulation of transcription, post-translational modification, and protein trafficking (reviewed by Bernardi and Pandolfi, 2007). Association of nascent HDV genome with PML bodies indicates discrete subnuclear compartmentalization of the synthesis of these two species. HDAg-S was also shown to co-immunoprecipitate with both RNAP II and SL1, a component of the RNAP I transcriptional machinery (Li et al., 2006). Furthermore, when *in vitro* transcription was performed using purified HDV particles in HeLa NE, synthesis of the antigenome was largely diminished upon immunodepletion with α-SL1 antibodies, whereas immunodepletion with α-RNAP II or α-PML antibodies had little effect. Genome synthesis was inhibited by all three antibodies including SL1, suggesting the utilization of different polymerases for transcription but also illustrating the need for further characterization of HDV-RNAP interactions.

Additional data supporting a role for RNAP I in HDV RNA synthesis have come from experiments in which recombinant HDAg-S was fused to a nucleolar localization signal to restrict its distribution to the nucleolus (Huang et al. 2008a). It was shown that genome synthesis was undetectable when HDAg-S was sequestered to the nucleolus while antigenome synthesis was unaffected. It is thus suspected that initiation of synthesis of the antigenome occurs in the nucleolus, and that the observations made by Li and associates (2006) with respect to the nucleolar localization of nascent antigenome are the result of active transcription and not transcript accumulation following synthesis. Synthesis of the genome could be restored upon actinomycin D treatment to release nucleolar proteins to the

nucleoplasm. Although actinomycin D is a general transcription inhibitor of DNA-dependant RNA synthesis by all three eukaryotic polymerases, it does not affect HDV RNA synthesis (Macnaughton et al., 2002). RNAP II failed to co-immunoprecipitate with nucleolar-localized HDAg-S until actinomycin D treatment released HDAg-S into the nucleoplasm (Huang et al., 2008a). Surprisingly, RNAP I did not co-immunoprecipitate with HDAg-S as would be expected based on the observations of Li and associates (2006). Although it is possible that RNAP I-directed HDV synthesis does not require HDAg-S, additional studies will be required to further characterize the role of RNAP I in HDV replication.

Further evidence of RNAP I involvement in HDV biology has been achieved using a sensitive, quantitative RT-PCR strategy to detect genomic, antigenomic, and HDAg mRNA transcripts in an HDV cDNA-free transfection system (Chen et al., 2008). Experiments using this model have shown that post-translational modifications of HDAg, including methylation of R13, acetylation of K72, and phosphorylation of S177 have differing effects on synthesis of these three transcripts. Synthesis of the genome and HDAg mRNA requires these modifications while they are unnecessary for antigenome synthesis. Serine 177 phosphorylation was shown to be important in RNAP II interaction with HDAg; specifically, synthesis of the genome from the antigenome requires this modification. The acetylation and deacetylation of HDAg-S is proposed to serve as a molecular switch that dictates which transcripts are generated by cellular polymerases. These data contribute to the body of evidence supporting transcription of HDAg mRNA and synthesis of the genome by a separate polymerase than that responsible for antigenome synthesis. Synthesis of HDAg mRNA and the genome share similar features including a requirement for a modified HDAg, nucleoplasmic localization, export to the cytoplasm, and sensitivity to α-amanitin, and these requirements differ from those of antigenome synthesis (Chen et al., 2008; Li et al., 2006;

Macnaughton and Lai, 2002b; Macnaughton et al., 2002; Modahl et al., 2000). These observations are summarized in Table 1.3. Together, it can be speculated that at least two cellular polymerases are responsible for the synthesis of HDV RNA and that each polymerase has a discrete role in the HDV life cycle.

Despite lines of evidence supporting RNAP I involvement in HDV antigenome synthesis, involvement of RNAPs other than RNAP II in HDV RNA synthesis remains controversial. Recent developments from the Taylor laboratory have refuted the involvement of RNAP I as a significant participant in HDV transcription (Chang et al., 2008). Using an inducible HDV replication system, the sensitivity of transcription of HDV RNA species to α-amanitin was determined by transcriptional run-on assay. It was shown that treatment with small amounts of α-amanitin reduced transcription of both polarities of HDV to approximately 14% of their pre-treatment levels while RNAP I and III cellular transcripts were largely unaffected by α-amanitin treatment. Modified ChIP assays confirmed that both polarities of unit-length HDV RNA interact with RNAP II in cells replicating HDV, and both polarities were shown to be present in large RNAP II- and HDAg-S-containing complexes. HDAg-S and RNAP II co-localize in SC35 speckles and in the nucleoplasm of cells actively replicating HDV, indicating that transcription or transcript processing is likely occurring. These data would suggest that RNAP II is the principal mediator of HDV RNA synthesis despite convincing evidence for RNAP I involvement.

Although HDV is the smallest known mammalian pathogen, data emerging regarding polymerase utilization and regulation of RNA synthesis by this virus are attaining a level of previously-unimagined complexity. A comprehensive picture of the HDV life cycle, the polymerases involved, and promoter elements mediating RNA-templated transcription by RNAPs are rapidly emerging, and these data will provide valuable insight into the intricate

Table 1.3: Features of transcription of various HDV RNA species.

Feature	Genome → antigenome	Antigenome → genome and mRNA
Sensitivity to α-amanitin	Resistant	Sensitive
Localization	Nucleolus	Nucleoplasm; PML bodies; SC35 speckles
Export	Remains in nucleus	Exported to cytoplasm
delta antigen	Unmodified	Methylated, phosphorylated, acetylated
Polymerases involved[a]	RNAP I, II	RNAP II

(a) Polymerase assignment based on current evidence; does not exclude involvement of other polymerases. All data references are cited in text.

world of RNA biology. A large component of my doctoral research is dedicated to confirming the RNAP II-HDV RNA interaction, characterizing this interaction with respect to the RNA species bound by RNAP II, and determining if RNAP I and RNAP III can similarly interact with HDV RNA. Elucidation of host polymerases capable of RNA-templated RNA synthesis will not only provide insight into mechanisms of HDV replication and transcription, but will also suggest the exciting possibility for RNA-templated RNA synthesis within eukaryotic cells and imply a novel RNA-to-RNA flow of genetic information.

1.3.6 Characteristics of putative HDV RNA promoters

HDV exists as a quasi species - a heterogeneous population of viral RNA species differing within and among isolates - presumably due to the infidelity of the host RNAP(s) when utilizing an unconventional, RNA template. However, certain constraints for the size and composition of the HDV genome must exist in order to be recognized and replicated by the nuclear transcriptional machinery. Although at least seven major HDV clades have been identified and substantial sequence variability is found among isolates, a recent study of over 200 global samples of HDV revealed that the genome size of each isolate was within 30 nt of each other (Radjef et al., 2004). This is consistent with the observation that viral genome size is constrained by the fidelity of the polymerase responsible for its replication; if one or more eukaryotic RNAPs mediate HDV transcription, there would be a presumed increase in transcriptional infidelity when utilizing a non-preferred template (Poole and Logan, 2005). While nucleotide sequence is variable among HDV isolates, base pair co-variation exists to maintain the rod-shaped conformation through base pairing, implicating a strong selective pressure to maintain this structure (Greco-Stewart et al., 2007). Additionally, sequence

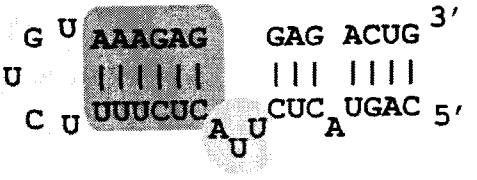
conservation occurs in the region of the *delta* ribozymes and HDAg ORF since these motifs are required for HDV replication (Chao, Y.C. et al., 1990). Smaller structural motifs at the terminal hairpin domains of the HDV RNA genomic and antigenomic species also show a high level of sequence conservation and co-variation, suggesting that these regions might also have an essential role in HDV biology (Greco-Stewart et al., 2007).

Evidence for multiple HDV promoters arose from studies in which small insertions were made in HDV RNA at various locations (Wang et al., 1997). It was shown that insertion into almost any region of HDV RNA was deleterious to RNA synthesis and that replication-competent mutants lost the inserted sequence upon replication. Insertion into the right terminal domain of the genome predominantly inhibited synthesis of the antigenome, and insertion into the left terminal domain of the antigenome inhibited genome synthesis. Furthermore, mutants were identified that had the capacity to synthesize the genome but not HDAg mRNA, indicating that unique elements are required for synthesis of different HDV RNA species. These data support the existence of multiple promoters in the HDV genome and antigenome.

Several small structural features of HDV RNA that are important for transcription and replication have been identified using *in vitro* transcription and cell infection models. Hairpin structures located at the terminal domains of the rod-shaped HDV genome and antigenome have been reported to represent sites of transcriptional initiation (Beard et al., 1996; Filipovska and Konarska, 2000). One such putative promoter, R199G, is located at the right terminal domain of genomic polarity (Beard et al., 1996; extreme terminal tip shown in Figure 1.8A). HDV mutants containing large deletions in which 512, 1049, or 1329 nt are removed from the right terminal domain encompassing the R199G region retain *delta* ribozyme activity but are incapable of replication (Lazinski and Taylor, 1994).

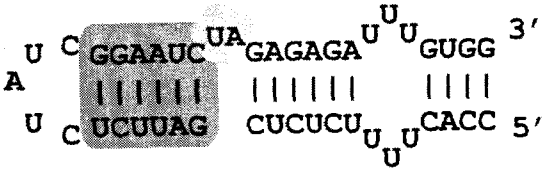
Figure 1.8: Extreme terminal domains of HDV RNA. A) Extreme tip of the right terminal domain of genomic polarity (reported to have promoter activity; Beard et al., 1996). B) Extreme tip of left terminal domain of antigenomic polarity (reported to have promoter activity; Filipovska and Konarska, 2000; Lehmann et al., 2007). C) Extreme tip of left terminal domain of genomic polarity. Important feature are highlighted including terminal loops (yellow), stems (red), and external bulges (green).

A)



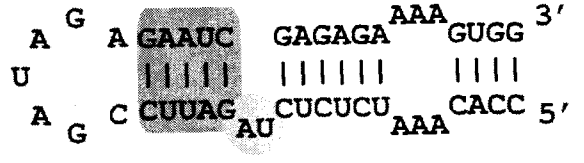
Genome: Right terminal

B)



Antigenome: Left terminal

C)



Genome: Left terminal

Computational analysis of HDV isolates shows the conservation of a high G/C content in the R199G region and the preservation of several small motifs at the extreme terminal tip including a 3 nt external bulge and 5 nt terminal loop with a 6 nt stem (Gudima et al., 1999; Figure 1.8A). These structures have all been shown to be necessary for HDV RNA synthesis.

The 3 nt external bulge at the extreme tip of R199G contains the putative U_{1630}/A_{1631} transcription start site, and mutation in this region results in a significant decrease in HDV accumulation (Beard et al., 1996; Gudima et al., 1999). When mutation is introduced to this site, compensatory sequence alteration to restore an adenosine residue as the start site occurs following HDV replication (Gudima et al., 1999). Mutation of this 3 nt bulge is poorly tolerated if its secondary structure is not preserved (Gudima et al., 1999; Wu et al., 1997). Removal of the 6 nt stem structure results in the inability of HDV to replicate, implicating the stem in the stabilization of the RNA secondary structure. Addition of only one or two nucleotides to this stem with maintenance of base pairing is deleterious to HDV accumulation, and deletion of one to three nucleotides results in almost undetectable levels of HDV accumulation (Gudima et al., 1999; Wu et al., 1997). HDV mutants harboring smaller deletions in this region in which upper and lower strand base pairing is preserved also fail to replicate, indicating a low tolerance for sequence alteration in this region despite preservation of secondary structure (Lazinski and Taylor, 1994). Inversion of the two strands of the stem, known as the "flip" mutation, results in diminished HDV replication to the point that less than 2% of the wild-type levels of HDV RNA accumulates (Wu et al., 1997). Mutation of the 5 nt terminal loop is also deleterious to replication if the secondary structure is grossly altered; randomization of the loop sequence and addition of two to four nucleotides to the loop has little effect on HDV accumulation (Gudima et al., 1999; Wu et al., 1997). In general, single nucleotide substitutions in R199G are well tolerated, although transitions are

more common than transversions among HDV isolates (Gudima et al., 1999; Netter et al., 1995; Wu et al., 1997). Together, these data indicate that the structure of the promoter region is more significant that sequence composition, particularly in the region of the external bulge and terminal loop, although the 6 nt stem sequence is requisite for promoter function, both in primary sequence and secondary structure.

The left terminal stem-loop domain of antigenomic polarity has also been characterized with respect to its ability to act as an RNA template for RNAP-mediated transcription. Mutational analysis showed that a 280 nt segment from this region can initiate rapid and efficient transcription (Filipovska and Konarska, 2000; extreme tip of this domain shown in Figure 1.8B). However, this RNA species loses promoter activity when deletions interfere with the hairpin secondary structure; restoration of the structure by compensatory deletions or alteration of downstream sequences rescues promoter activity. The transcriptional start site was localized to the extreme tip of the hairpin, and segments as small as 62 nt at the end of this hairpin can support transcription. While alteration of the primary sequence of this region is tolerated, alteration of the bulge region adjacent to the putative start site prevents RNA product from accumulating. These data further support the observation that structure, rather than sequence, is preferentially recognized by host RNAPs when transcribing from a non-DNA template.

Recent experiments have confirmed that the extreme tip of the left terminal domain of the antigenome of HDV can act as a template for RNA-templated synthesis by RNAP II and provide insight into the RNA requirements for the RdRp activity of RNAP II (Lehmann et al., 2007). The HDV stem-loop is required for RNAP II recognition, and it binds the active site of RNAP II where the DNA-nascent RNA hybrid usually resides. Many short RNA stem-loop structures with 5' extensions can also serve as templates for the RdRp activity of

RNAP II. The GTF TFIIS stimulates cleavage of the UUA bulge of the HDV template to expose a 6 nt stem-loop which is subsequently elongated at the 3' end to generate the observed chimeric template/nascent RNA product. The HDV stem acquires an A-helical structure, and it appears that this is the key feature recognized by RNAP II. As the transcript grows during transcription, the length of the stem begins to impede polymerase procession over the template by clashing with the polymerase lid; since HDV is comprised of a long stem-loop structure, it is believed that binding of HDAg-S to the polymerase clamp results in conformational change and allows elongation to progress *in vivo* (Yamaguchi et al., 2007). These observations indisputably show the ability of RNAP II to transcribe from an RNA template and reveal the secondary structure elements of RNA templates that are necessary for this unconventional transcription to occur.

Alteration of additional regions of the HDV genome can adversely affect HDV replication although promoter activity for these regions has not been reported. Studies have been performed in which mutations have been introduced to the left terminal of the HDV genome (Gudima et al., 2006; Wu et al., 1997; Figure 1.8C). This region is more variable among quasi-species; it is comprised of a 5 to 6 nucleotide stem flanked by a 5 to 7 nt terminal loop (Greco-Stewart et al., 2007). Accordingly, this region is more tolerant to single nucleotide changes within the terminal loop and deletions and insertions affecting the stem structure. Randomization of the terminal loop and insertion of up to six nucleotides therein have negligible effect on HDV accumulation (Wu et al., 1997). Deletion of the terminal loop and 3 adjacent base pairs diminishes RNA accumulation by over 50%, and deletion of 16, 26, and 36 nucleotides has increasingly more pronounced effects. After deletion of 36 nucleotides HDV accumulation is barely detectable. It was further observed that insertion of large sequences including the GNRA tetraloop, bacterial terminator structure, and 55 to 57 nt

segments of foreign RNA into this region almost completely abrogates RNA accumulation when the insert exceeds 4 nt in length (Gudima et al., 2006; Wu et al., 1997). Mutants retaining any replication competency revert to wild-type size following three consecutive rounds of transfection (Gudima et al., 2006). Large insertions in the center of the HDV genome that disrupt the rod-shaped structure are also poorly tolerated, and reversion to wild-type size similarly occurs. These observations demonstrate a high amount of selective pressure for maintenance of the rod-shaped structure and finite genome size for HDV replication, and show the necessity of several regions of HDV RNA for HDV survival. Due to the increased tolerance for sequence flexibility at the extreme tip of the hairpin, the significance of this region in HDV replication requires further characterization. However, the possibility that it contains an uncharacterized promoter or enhancer element for HDV RNA synthesis cannot be excluded. It is thus possible that multiple uncharacterized promoters for the RdRp activity of RNAP II, or other host polymerases, exist in other regions of the HDV genome.

1.4 Statement of purpose and hypothesis

Much remains to be discovered regarding the host polymerases involved in HDV replication and transcription and the manner in which HDV subverts these polymerases to fulfill its life cycle. The present study investigates these questions based on the following hypothesis:

Host proteins interact with HDV RNA to facilitate transcription and replication by processes that coerce DNA-dependent RNA polymerases to transcribe from a non-conventional RNA template.

The objectives of the present study are thus:

- To determine which host RNAPs interact with HDV RNA both *in vitro* and in cells replicating HDV
- To characterize the specific RNA features, such as sequence and secondary structure, required for RNAP binding
- To identify additional host proteins that interact with HDV RNA and characterize their binding with respect to requisite RNA features
- To propose a role for novel HDV RNA-interacting proteins in HDV RNA synthesis

Achievement of these objectives will not only broaden the global knowledge of HDV RNA synthesis but also contribute to the rapidly-expanding field of RNA biology by providing insight into the requirements for RNA-templated transcription by DNA-dependent RNA polymerases – a phenomenon which could help explain the evolution of ancestral polymerases and provide an additional layer of complexity to the story of RNA regulation in eukaryotic cells.

1.5 References

Abrahem, A., Pelchat, M., 2008. Formation of an RNA polymerase II preinitiation complex on an RNA promoter derived from the hepatitis *delta* virus RNA genome. Nucleic Acids Res. 36, 5201-5211.

Alves, C., Freitas, N., Cunha, C., 2008. Characterization of the nuclear localization signal of the hepatitis delta virus antigen. Virology 370, 12-21.

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Bernardi, R., Pandolfi, P.P., 2007. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat. Rev. Mol. Cell. Biol. 8, 1006-1016.

Bezy, O., Elabd, C., Cochet, O., Petersen, R.K., Kristiansen, K., Dani, C., Ailhaud, G., Amri, E.Z., 2005. *Delta*-interacting protein A, a new inhibitory partner of CCAAT/enhancer-binding protein *beta*, implicated in adipocyte differentiation. J. Biol. Chem. 280, 11432-11438.

Bichko, V.V., Taylor, J.M., 1996. Redistribution of the *delta* antigens in cells replicating the genome of hepatitis *delta* virus. J. Virol. 70, 8064-8070.

Biebricher, C.K., Luce, R., 1996. Template-free generation of RNA species that replicate with bacteriophage T7 RNA polymerase. EMBO J., 15, 3458-3465.

Biebricher, C.K., Orgel, L.E., 1973. An RNA that multiplies indefinitely with DNA-dependent RNA polymerase: selection from a random copolymer. Proc. Natl. Acad. Sci. 70, 934-938.

Bonino, F., Hoyer, B., Shih, J.W., Rizettto, M., Purcell, R.H., Gerin, J.L., 1984. *Delta* hepatitis agent: structural and antigenic properties of the *delta*-associated particle. Infect. Immun. 43, 1000-1005.

Branch, A.D., Robertson, H.D., 1984. A replication cycle for viroids and other small infectious RNAs. Science 223, 450-455.

Brazas, R., Ganem, D., 1996. A cellular homolog of hepatitis *delta* antigen: implications for viral replication and evolution. Science 274, 90-94.

Bushnell, D.A., Cramer, P., Kornberg, R.D., 2002. Structural basis of transcription: α-amanitin-RNA polymerase II cocrystal at 2.8 angstrom resolution. Proc. Natl. Acad. Sci. 99, 1218-1222.

Casey, J.L., Bergmann, K.F., Brown, T.L., Gerin, J.L., 1992. Structural requirements for RNA editing in hepatitis *delta* virus: evidence for a uridine-to-cytidine editing mechanism. Proc. Natl. Acad. Sci. 89, 7149-7153.

Casey, J.L., Gerin, J.L., 1995. Hepatitis D virus RNA editing: specific modification of adenosine in the antigenomic RNA. J. Virol. 69, 7593-7600.

Chadalavada, D.D., Cerrone-Szakal, A.L., Bevilacqua, P.C., 2007. Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions. RNA 13, 2189-2201.

Chambon, P. 1975. Eukaryotic nuclear RNA polymerases. Annu. Rev. Biochem. 44, 613-638.

Chang, F.L., Chen, P.J., Tu, S.J., Wang, C.J., Chen, D.S., 1991. The large form of hepatitis *delta* antigen is crucial for the assembly of hepatitis *delta* virus. Proc. Natl. Acad. Sci. 88, 8490-8494.

Chang, J., Taylor, J.M., 2002. *In vivo* RNA-directed transcription, with template switching, by a mammalian RNA polymerase. EMBO J. 21, 157-164.

Chang, J., Provost, P., Taylor, J.M., 2003. Resistance of human hepatitis *delta* virus RNAs to dicer activity. J. Virol. 77, 11910-11917.

Chang, J., Gudima, S.O., Tarn, C., Nie, X., Taylor, J.M., 2005. Development of a novel system to study hepatitis *delta* virus genome replication. J. Virol. 79, 8182-8188.

Chang, J., Nie, X., Chang, H.E., Han, Z., Taylor, J., 2008. Transcription of hepatitis *delta* virus RNA by RNA polymerase II. J. Virol. 82, 1118-1127.

Chang, M.F., Baker, S.C., Soe, L.H., Kamahora, T., Keck, J.G., Makino, S., Govindarajan, S., Lai, M.M., 1988. Human hepatitis *delta* antigen is a nuclear phosphoprotein with RNA-binding activity. J. Virol. 62, 2403-2410.

Chao, M., 2007. RNA recombination in hepatitis *delta* virus: implications regarding the abilities of mammalian RNA polymerases. Virus Res. 127, 208-215.

Chao, M., Hsieh, S.Y., Taylor, J., 1990. Role of two forms of the hepatitis *delta* virus antigen: evidence for a mechanism of self-limiting genome replication. J. Virol. 64, 5066-5069.

Chao, M., Hsieh, S.Y., Taylor, J., 1991. The antigen of hepatitis *delta* virus: examination of *in vitro* RNA-binding specificity. J. Virol. 65, 4057-4062.

Chao, Y.C., Chang, M.F., Gust, I., Lai, M.M., 1990. Sequence conservation and divergence of hepatitis *delta* virus RNA. Virology 178, 384-392.

Chen, C.W., Tsay, Y.G., Wu, H.L., Lee, C.H., Chen, D.S., Chen, P.J., 2002. The double-stranded RNA-activated kinase, PKR, can phosphorylate hepatitis D virus small *delta* antigen at functional serine and threonine residues. J. Biol. Chem. 277, 33058-33067.

Chen, P.J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., Taylor, J., 1986. Structure and replication of the genome of hepatitis δ virus. Proc. Natl. Acad. Sci. 83, 8774-8778.

Chen, Y.S., Huang, W.H., Hong, S.Y., Tsay, Y.G., Chen, P.J., 2008. ERK1/2-mediated phosphorylation of small hepatitis *delta* antigen at serine 177 enhances hepatitis *delta* virus antigenomic RNA replication. J. Virol. 82, 9345-9358.

Chou, H.C., Hsieh, T.Y., Sheu, G.T., Lai, M.M., 1998. Hepatitis *delta* antigen mediates the nuclear import of hepatitis *delta* virus RNA. J. Virol. 72, 3684-3690.

Cochrane, J.C., Strobel, S.A., 2008. Catalytic strategies of self-cleaving ribozymes. Acc. Chem. Res. 41, 1027-1035.

Cormack, B.P., Struhl, K., 1992. The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. Cell, 69, 685-696.

Cramer, P., 2002. Multisubunit RNA polymerases. Curr. Opin. Struct. Biol. 12, 89-97.

Cramer, P., 2004. RNA polymerase II structure: from core to functional complexes. Curr. Opin. Genet. Dev. 14, 218-226.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengle, S., Geiger, S.R., Jasiak, A.J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C.D., Lehmann, E., Leike, K., Sydow, J.F., Vannini, A., 2008. Structure of eukaryotic polymerases. Annu. Rev. Biophys. 37, 337-352.

Cramer, P., Bushnell, D.A., Kornberg, R.D., 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 292: 1863-1876.

Daròs, J.A., Elena, S.F., Flores, R., 2006. Viroids: an Ariadne's thread into the RNA labyrinth. EMBO Rep. 7, 593-598.

Davies, J.W., Kaesberg, P., Diener, T.O., 1974. Potato spindle tuber viroid XII. An investigation of viroid RNA as a messenger for protein synthesis. Virology 61, 281-286.

Diener, T.O., 1974. Viroids: the smallest known agents of infection. Annu. Rev. Microbiol. 28, 23-29.

Du, X., Wang, Q., Hirohashi, Y., Greene, M.I., 2006. DIPA, which can localize to the centrosome, associates with p78/MCRS1/MSP58 and acts as a repressor of gene transcription. Exp. Mol. Pathol. 81, 184-190.

Egloff, S., Murphy, S. 2008. Cracking the RNA polymerase II CTD code. Trends Genet. 24, 280-288.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Farci, P., Chessa, L., Balestrieri, C., Serra, G., Lai, M.E., 2007. Treatment of chronic hepatitis D. J. Viral Hepat. Suppl. 1, 58-63.

Fernandez-Toreno, C., Bottcher, B., Riva, M., Carles, C., Steurwald, U., Ruigrok, R.W., Sentenac, A., Müller, C.W., Schoehn, G., 2007. Insights into transcription and termination from the electron microscopy structure of yeast RNA polymerase III. Mol. Cell 25, 813-823.

Ferré-D'Amaré, A.R., Zhou, K., Doudna, J.A., 1998. Crystal structure of a hepatitis *delta* virus ribozyme. Nature 279, 621-631.

Filipovska, J., Konarska, M., 2000. Specific HDV RNA-templated transcription by pol II *in vitro*. RNA 6, 41-54.

Flores, R., Randles, J. Bar-Joseph, M., Diener, T., 1998. A proposed scheme for viroid classification and nomenclature. Arch. Virol. 143, 623-629.

Fu, T., Taylor, J., 1993. The RNAs of hepatitis *delta* virus are copied by RNA polymerase II in nuclear homogenates. J. Virol. 67, 6965-6972.

Ghisolfi, L., Joseph, G., Amalric, F., Erard, M., 1992. The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. J. Biol. Chem. 267, 2955-2959.

Glenn, J.S., Watson, J.A., Havel, C.M., White, J.M., 1992. Identification of a prenylation site in *delta* virus large antigen. Science 256, 1331-1333.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., Kornberg, R.D., 2001. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 angstrom resolution. Science 292, 1876-1872.

Goodman, T.C., Nagel, L., Rappold, W., Klotz, G., Reisner, D., 1984. Viroid replication: equilibrium association constant and comparative activity measurements for the viroid-polymerase interaction. Nucleic Acids Res. 12, 6231-6246.

Gordon, S., Akopyan, G., Garban, H.M, Bonavida, B., 2006. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. Oncogene 25, 1125-1142.

Govindarajan, S., Chin, K.P., Redeker, A.G., Peters, R.L., 1984. Fulminant B viral hepatitis: role of the *delta* agent. 1984. Gastroenterology 86, 1417-1420.

Gowans, E.J., Baroudy, B.M., Negro, F., Ponzetto, A., Purcell, R.H., Gerin, J.L., 1988. Evidence for replication of hepatitis *delta* virus RNA in hepatocyte nuclei after *in vivo* infection. Virology 167, 274-2478.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology 357, 68-78.

Gudima, S., Dingle, K., Wu, T.T., Moraleda, G., Taylor, J., 1999. Characterization of the 5' ends for polyadenylated RNAs synthesized during the replication of hepatitis *delta* virus. J. Virol. 73, 6533-6539.

Gudima, S., Wu, S.Y., Chiang, C.M., Moraleda, G., Taylor, J., 2000. Origin of hepatitis *delta* virus mRNA. J. Virol. 74, 7204-7210.

Gudima, S.O., Chang, J., Taylor, J.M., 2004. Features affecting the ability of hepatitis *delta* virus RNAs to initiate RNA-directed RNA synthesis. J. Virol. 78, 5737-5744.

Gudima, S.O., Chang, J., Taylor, J.M., 2006. Restoration *in vivo* of defective hepatitis *delta* virus RNA genomes. RNA. 12, 1061-1073.

Hahn, S., 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat. Struct. Mol. Biol. 11, 394-403.

Hall, T.C., Wepprich, R.K., Davies, J.W., Weathers, L.G., Semancik, J.S., 1974. Functional distinctions between the ribonucleic acids from citrus exocortis viroid and plant viruses: cell-free translation and aminoacylation reactions. Virology 61, 486-492.

Happel, N., Doenecke, D., 2009. Histone H1 and its isoforms: contribution to chromatin structure and function. Gene 431, 1-12.

Haussecker, D., Cao, D., Huang, Y., Parameswaran, P., Fire, A.Z., Kay, M.A., 2008. Capped small RNAs and MOV10 in human hepatitis *delta* virus replication. Nat. Struct. Mol. Biol. 15, 714-721.

Hsieh, S.Y., Chao, M., Coates, L., Taylor, J., 1990. Hepatitis *delta* virus genome replication: a polyadenylated mRNA for *delta* antigen. J. Virol. 64, 3192-3198.

Hsu, L.M., Vo, N.V., Kane, C.M., Chamberlin, M.J., 2003. *In vitro* studies of transcript initiation by *Escherichia coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. Biochemistry 42, 3777-3786.

Huang, C., Chang, S.C., Yu, I.C., Tsay, Y.G., Chang, M.F., 2007. Large hepatitis *delta* antigen is a novel clathrin adaptor-like protein. J. Virol. 81, 5985-5994.

Huang, W.H., Yung, B.Y., Syu, W.J., Lee, Y.H., 2001. The nucleolar phosphoprotein B23 interacts with hepatitis *delta* antigens and modulates the hepatitis *delta* virus RNA replication. J. Biol. Chem. 276, 25466-25175.

Huang, W.H., Chen, Y.S., Chen, P.J., 2008a. Nucleolar targeting of hepatitis *delta* antigen abolishes its ability to initiate viral antigenomic RNA replication. J. Virol. 82, 692-699.

Huang, W.H., Mai, R.T., Lee, Y.H., 2008b. Transcription factor YY1 and its associated acetyltransferase CBP and p300 interact with hepatitis *delta* antigens and modulate hepatitis *delta* virus RNA replication. J. Virol. 82, 7313-7324.

Huang, Z.S., Wu, H.N., 1998. Identification and characterization of the RNA chaperone activity of hepatitis *delta* antigen peptides. J. Biol. Chem. 273, 26455-26451.

Hwang, S.B., Lai, M.M., 1994. Isoprenylation masks a conformational epitope and enhances *trans*-dominant inhibition of the large hepatitis *delta* antigen. J. Virol. 68, 2958-2964.

Jacobson, I.M., Dienstag, J.L., Werner, B.G., Brettler, D.B., Levine, P.H., Mushahwar, I.K., 1985. Epidemiology and clinical impact of hepatitis D virus (*delta*) infection. Hepatology 5, 188-191.

Jasiak, A.J., Armache, K.J., Martens, B., Jansen, R.P., Cramer, P., 2006. Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. Mol. Cell 23, 71-81.

Kanel, G.C, Govindarajan, S., Peters, R.L., 1984. Chronic *delta* infection and liver biopsy changes in chronic acitive hepatitis B. Ann. Intern. Med. 101, 51-54.

Ke, A., Zhou, K., Ding, F., Cate, J.H., Doudna, J.A., 2004. A conformational switch controls hepatitis *delta* virus ribozyme catalysis. Nature 429, 201-205.

Khudyakov, Y.E., Makhov, A.M., 1990. Amino acid sequence similarity between the terminal protein of hepatitis B virus and predicted hepatitis *delta* virus gene product. FEBS Lett. 262, 345-348.

Komarnitsky, P., Cho, E.J., Buratowski, S., 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. 14, 2452-2460.

Kornberg, R.D., 1974. Chromatin structure: a repeating unit of histones and DNA. Science 184: 868-871.

Kornberg, R.D., Thomas, J.O., 1974. Chromatin structure: oligomers of the histones. 184: 865-868.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Garber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., Cramer, P., 2007. Functional architecture of RNA polymerase I. 131, 1260-1272.

Kim, J.B., Sharp, P.A., 2001. Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. J. Biol. Chem. 276, 12317-12323.

Konarska, M.M., Sharp, P.A., 1989. Replication of RNA by the DNA-dependent RNA polymerase of phage T7. Cell 57, 423-431.

Koytak, E.S., Yurdaydin, C., Glenn, J.S., 2007. Hepatitis D. Curr. Treat. Options. Gastroenterol. 10, 456-463.

Kuo, M.Y., Goldberg, J., Coates, L., Mason, W., Gerin, J., Taylor, J.M., 1988a. Molecular cloning of hepatitis *delta* virus RNA from an infected woodchuck liver: sequence, structure, and applications. J. Virol. 62, 1855-1861.

Kuo, M.Y., Sharmeen, L., Dinter-Gottlieb, G., Taylor, J., 1988b. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis *delta* virus. J. Virol. 62, 4439-4444.

Kuo, M.Y., Chao, M., Taylor, J., 1989. Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of the *delta* antigen. J. Virol. 63, 1945-1950.

Lai, M.M., 2005. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis *delta* virus. J. Virol. 79, 7951-7958.

Lazinski, D.W., 1999. Viral Hepatitis. In: Schaechter, M., Engleberg, N.C., Eisenstein, B.I., and Medoff, G. (Eds), Mechanisms of microbial disease, 3rd edition. Lippincott, Williams &Wilkins, Baltimore, pp. 391-402.

Lazinski, D.W., Taylor, M.J., 1994. Expression of hepatitis *delta* virus RNA deletions: *cis* and *trans* requirements for self-cleavage, ligation, and RNA packaging. J. Virol. 61, 2879-2888.

Lee, C.H., Chang, S.C., Chen, C.J., Chang, M.F., 1998. The nucleolin binding activity of hepatitis *delta* antigen is associated with nucleolus targeting. J. Biol. Chem. 273, 7650-7656.

Lee, C.H., Chang, S.C., Wu, C.H., Chang, M.F., 2001. A novel chromosome region maintenance 1-independent nuclear export signal of the large form of the hepatitis *delta* antigen that is required for viral assembly. J. Biol. Chem. 276, 8142-8148.

Lee, C.Z., Lin, J.H., Chao, M., McKnight, K., Lai, M.M., 1993. RNA-binding activity of hepatitis delta *antigen* involves two arginine-rich motifs and is required for hepatitis *delta* virus RNA replication. J. Virol. 67, 2221-2227.

Lee, C.Z., Chen, P.J., Lai, M.M., Chen, D.S., 1994. Isoprenylation of large hepatitis *delta* antigen is necessary but not sufficient for hepatitis *delta* virus assembly. Virology 199, 169-175.

Lee, C.Z., Chen, P.J., Chen, D.S., 1995. Large hepatitis *delta* antigen in packaging and replication inhibition: role of the carboxyl-terminal 19 amino acids and amino-terminal sequences. J. Virol. 69, 5332-5336.

Lee, C.Z., Sheu, J.C., 2008. Histone H1e interacts with small hepatitis *delta* antigen and affects hepatitis delta virus replication. Virology 375, 197-204.

Lee, T.I., Young, R.A., 2000. Transcription of eukaryotic protein-coding genes. Ann. Rev. Genet. 34, 77-137.

Lehmann, E., Brueckner, F., Cramer, P., 2007. Molecular basis of RNA-dependent RNA polymerase II activity. Nature 450, 445-449.

Li, Y.J., Stallcup, M.R., Lai, M.M., 2004. Hepatitis *delta* virus antigen is methylated at arginine residues and methylation regulates subcellular localization and RNA replication. J. Virol. 78, 13325-136334.

Li, Y.J., Macnaughton, T., Gao, L., Lai, M.M., 2006. RNA-templated replication of hepatitis *delta* virus: genomic and antigenomic RNAs associate with different nuclear bodies. J. Virol. 80, 6478-6486.

Lin, J.H., Chang, M.F., Baker, S.C., Govindarajan, S., Lai, M.M., 1990. Characterization of hepatitis *delta* antigen: specific binding to hepatitis *delta* virus RNA. J. Virol. 64, 4051-4058.

Lin, S.S., Chang, S.C., Wang, Y.H., Sun, C.Y., Chang, M.F., 2000. Specific interaction between the hepatitis *delta* virus RNA and glyceraldehyde 3-phosphate dehydrogenase: an enhancement on ribozyme catalysis. Virology 271, 46-57.

Poole, A.M., Logan, D.T., 2005. Modern mRNA proofreading and repair: clues that the last universal common ancestor possessed an RNA genome? Mol. Biol. Evol. 22, 1444-1455.

Luse, D.S., Jacob, G.A., 1987. Abortive initiation by RNA polymerase II *in vitro* at the adenovirus 2 major late promoter. J. Biol. Chem. 262, 14990-14997.

Macnaughton, T.B., Gowans, E.J., MacNamara, S.P., Burrell, C.J., 1991. Hepatitis *delta* antigen is necessary for access for hepatitis *delta* virus RNA to the cell transcriptional machinery but is not part of the transcriptional complex. Virology 184, 387-390.

Macnaughton, T.B., Lai, M.M., 2002a. Large hepatitis *delta* antigen is not a suppressor of hepatitis *delta* virus RNA synthesis once RNA replication is established. J. Virol. 76, 9910-9919.

Macnaughton, T.B., Lai, M.M., 2002b. Genomic but not antigenomic hepatitis *delta* virus is preferentially exported from the nucleus immediately after synthesis and processing. J. Virol. 76, 3928-3935.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M.M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

Manche, L., Green, S.R., Schmedt, C., Matthews, M.B., 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. Mol. Cell. Biol. 12, 5238-5248.

Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., Cramer, P., 2005. A structural perspective of CTD function. Genes Dev. 19, 1401-1415.

Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Lührmann, R., Tuschi, T., 2005. Identification of novel argonaute-associated proteins. Curr. Biol. 15, 2149-2155.

Modahl, L.E., Lai, M.M., 1998. Transcription of hepatitis *delta* antigen mRNA continues throughout hepatitis *delta* virus (HDV) replication: a new model of HDV RNA transcription and replication. J. Virol. 72, 5449-5446.

Modahl, L.E., Macnaughton, T.B., Zhu, N., Johnson, D.L., Lai, M.M., 2000. RNA-dependent replication and transcription of hepatitis *delta* virus RNA involve distinct cellular RNA polymerases. Mol. Cell. Biol. 20, 6030-6039.

Moroianu, J., Blobel, G., Radu, A., 1995. Previously identified protein of uncertain function is karyopherin *alpha* and together with karyopherin *beta* docks import substrate at nuclear pore complexes. Proc. Natl. Acad. Sci. 92, 2008-2011.

Mota, S., Mendes, M., Penque, D., Coelho, A.V., Cunha, C., 2008. Changes in the proteome of Huh7 cells induced by transient expression of hepatitis D virus RNA and antigens. J. Proteomics 71, 71-79.

Mu, J.J., Wu, H.L., Chiang, B.L., Chang, R.P., Chen, D.S., Chen, P.J., 1999. Characterization of the phosphorylated forms and the phosphorylated residues of hepatitis *delta* virus *delta* antigens. J. Virol. 73, 10540-10545.

Mu, J.J., Chen, D.S., Chen, P.J., 2001. The conserved serine 177 in the *delta* antigen of hepatitis *delta* virus is one putative phosphorylation site and is required for efficient viral RNA replication. J. Virol. 75, 9087-9095.

Mu, J.J., Tsay, Y.G., Juan, L.J., Fu, T.F., Huang, W.H., Chen, D.S., Chen, P.J., 2004. The small *delta* antigen of hepatitis *delta* virus is an acetylated protein and acetylation of lysine 72 may influence its cellular localization and viral RNA synthesis. Virology 319, 60-70.

Mühlbach, H.P., Sänger, H.L., 1979. Viroid replication is inhibited by α-amanitin. Nature 278, 185-188.

Nedialkov, Y.A., Gong, X.Q., Hovde, S.L., Yamaguchi, Y., Handa, H., Geiger,J.H., Yan, Y., Burton, Z.F., 2003. NTP-driven translocation by human RNA polymerase II. J. Biol. Chem. 278, 18303-18312.

Netter, H.J., Hsieh, S.Y., Lazinski, D., Taylor, J., 1993. Modified HDV as a vector for the delivery of biologically-active RNAs. Prog. Clin. Biol. Res. 382, 373-376.

Netter, H.J., Wu, T.T., Bockol, M., Cywinski, A., Ryu, W.S., Tennant, B.C., Taylor, J.M., 1995. Nucleotide sequence stability of the genome of hepatitis *delta* virus. J. Virol. 69, 1687-1692.

Nie, X., Chang, J., Taylor, J.M., 2004. Alternative processing of hepatitis delta virus antigenomic RNA transcripts. J. Virol. 78, 4517-4524.

Niranjanakumari, S., Lasda, E., Brazas, E., Garcia-Blanco, M.A., 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods 26, 182-190.

Okuwaki, M., 2008. The structure and functions of NPM1/nucleophosmin/B23, a multifunctional nucleolar acidic protein. J. Biochem. 143, 441-448.

Otto, J.C., Casey, P.J., 1996. The hepatitis *delta* virus large antigen is farnesylated both *in vitro* and in animal cells. J. Biol. Chem.2 71, 4569-4572.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P., Nadal-Ginard, B., 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev. 7, 393-406.

Poisson, F., Roingeard, P., Baillou, A., Dubois, F., Bonelli, F., Calogero, R.A., Goudeau, A., 1993. Characterization of RNA-binding domains of hepatitis *delta* antigen. J. Gen. Virol. 74, 2473-2478.

Polson, A.G., Bass, B.L., Casey, J.L., 1996. RNA editing of hepatitis *delta* virus antigenome by dsRNA-adenosine deaminase. Nature 380, 454-456.

Polson, A.G., Ley, H.L. 3rd, Bass, B.L., Casey, J.L., 1998. Hepatitis *delta* virus RNA editing is highly specific for the amber/W site and is suppressed by hepatitis *delta* antigen. Mol. Cell. Biol. 18, 1919-1926.

Ponzetto, A., Coté, P.J., Popper, H., Hoyer, B.H., London, W.T., Ford, E.C., Bonino, F., Purcell, R.H., Gerin, J.L., 1984. Transmission of the hepatitis B-associated *delta* agent to the eastern woodchuck. Proc. Natl. Acad. Sci. 81, 2208-2212.

Powell, W., Bartholomew, B., Reines, D., 1996. Elongation factor SII contacts the 3'-end of RNA in the RNA polymerase II elongation complex. J. Biol. Chem. 271, 22301-22304.

Radjef, N., Gordien, E., Ivaniushina, V., Gault, E., Anaïs, P., Drugan, T., Trinchet, J.C., Roulot, D., Tamby, M., Milinkovitch, M.C., Dény, P., 2004. Molecular phylogenetic analyses indicate a wide and ancient radiation of African hepatitis *delta* virus, suggesting a *delta* virus genus of at least seven major clades. J. Virol. 78, 2537-2544.

Reid, C.E., Lazinski, D.W., 2000. A host-specific function is required for ligation of a wide variety of ribozyme-processed RNAs. Proc. Natl. Acad. Sci. 97, 424-429.

Rizzetto, M., Canese, M.G., Arico, S., Crivelli, O., Trepo, C., Bonino, F., Verme, G., 1977. Immunofluorescence detection of new antigen-antibody system (*delta*/anti-*delta*) associated to hepatitis B virus in liver and in serum of HBsAg carriers. Gut 18, 997-1003.

Robertson, H.D., Manche, L., Matthews, M.B., 1996. Paradoxical interactions between human *delta* hepatitis agent RNA and the cellular protein kinase PKR. J. Virol. 70, 5611-5617.

Romeo, R., Ninno, E.D., Rumi, M., Russo, A., Sangiovanni, A., Franchis, R.D., Ronchi, G., Columbo, M., 2009. A 28-year study of the course of hepatitis *delta* infection: a risk factor for cirrhosis and hepatocellular carcinoma. Gastroenterology 136, 1629-1638.

Rosenstein, S.P., Been, M.D. 1990. Self-cleavage of hepatitis *delta* virus genomic strand RNA is enhanced under partially denaturing conditions. Biochemistry 29, 8011-8016.

Ryu, W.S., Bayer, M., Taylor J., 1992. Assembly of hepatitis *delta* virus particles. J. Virol. 66, 2310-2315.

Ryu, W.S., Netter, H.J., Bayer, M., Taylor, J., 1993. Ribonucleoprotein complexes of hepatitis *delta* virus. J. Virol. 67, 3281, 3287.

Salehi-Ashtiani, K., Lupták, A., Szotak, J.W., 2006. A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. Science 313, 1788-1792.

Shav-Tal, Y., Zipori, D., 2002. PSF and p54(nrb)/NonO--multi-functional nuclear proteins. FEBS Lett. 531, 109-114.

Sharmeen, L., Kuo, M.Y., Dinter-Gottlieb, G., Taylor, J., 1988. Antigenomic RNA of human hepatitis *delta* virus can undergo self-cleavage. J. Virol. 62, 2674-2679.

Sioud, M., Jesperson, L., 1996. Enhancement of hammerhead ribozyme catalysis by glyceraldehyde 3-phosphate dehydrogenase. J. Mol. Biol. 257, 775-789.

Sirover, M.A., 2005. New nuclear functions of the glycolytic protein, glyceraldehyde 3-phosphate dehydrogenase, in mammalian cells. J. Cell Biochem. 95, 45-52.

Sureau, C., 2006. The role of HBV envelope proteins in the HDV replication cycle. Curr. Top. Microbiol. Immunol. 307, 113-131.

Sureau, C., Moriarty, A.M., Thornton, G.B., Lanford, R.E., 1992. Production of infectious hepatitis *delta* virus *in vitro* and neutralization with antibodies directed against the hepatitis B virus pre-S antigen. J. Virol. 66, 1241-1245.

Szebeni, A., Herrera, J.E., Olson, M.O., 1995. Interaction of nucleolar protein B23 with peptides related to nuclear localization signals. Biochemistry 34, 8037-8042.

Szebeni, A., Mehorotra, B., Baumann, A., Adam, S.A., Wingfield, P.T., Olson, M.O., 1997. Nucleolar protein B23 stimulates nuclear import of the HIV-1 Rev protein and NLS-conjugated albumin. Biochemistry 36, 3941-3949.

Taylor, J.M., 1999. Hepatitis *delta* virus. Intervirology 42, 173-178.

Taylor, J.M., 2006. Hepatitis *delta* virus. Virology. 344, 71-76.

Tsagris, E.M. Martinez de Alba, A.E. Gozmanova, M., Kalantidis, K., 2008. Viroids. Cell. Microbiol. 10, 2168-2179.

Tseng, C.H., King-Song, Jeng, Lai, M.M., 2008. Transcription of subgenomic mRNA of hepatitis *delta* virus requires a modified hepatitis *delta* antigen that is distinct from antigenomic RNA synthesis. J. Virol. 82, 9409-9416.

Tuteja, R., Tuteja, N., 1998. Nucleolin: a multifunctional nucleolar phosphoprotein. Crit. Rev. Biochem. Mol. Biol. 33, 407-436.

Wang, C.C., Chang, T.C., Lin, C.W., Tsui, H.L., Chu, P.B., Chen, B.S., Huang, Z.S., Wu, H.N., 2003. Nucleic acid binding properties of the nucleic acid chaperone domain of hepatitis *delta* antigen. Nucleic Acids Res. 15, 6481-6492.

Wang, H.W., Chen, P.J., Lee, C.Z., Wu, H.L., Chen, D.S., 1994. Packaging of hepatitis *delta* virus RNA via the RNA-binding domain of hepatitis *delta* antigens: different roles for the small and large *delta* antigens. J. Virol. 68, 6323-6371.

Wang, H.W., Wu, H.L., Chen, D.S., Chen, P.J., 1997. Identification of the functional regions required for hepatitis D virus replication and transcription by linker-scanning mutagenesis of viral genome. Virology 239, 119-131.

Wang, K.S., Choo, Q.L., Weiner, A.J., Ou, J.H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L., Houghton, M., 1986. Structure, sequence, and expression of the hepatitis *delta* (δ) viral genome. Nature 323, 508-514.

Wang, Y.H., Chang, S.C., Huang, C., Li, Y.P., Lee, C.H., Chang, M.F., 2005. Novel nuclear export signal-interacting protein, NESI, critical for the assembly of hepatitis *delta* virus. J. Virol. 79, 8113-8120.

Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. Trends Biochem. Sci. 24, 437-440.

Wassarman, K.M., 2007. 6S RNA: a regulator of transcription. Mol. Microbiol. 65, 1425-1431.

Wassarman, K.M., Stortz, G., 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell 101, 613-623.

Weiner, A.J., Choo, Q.L., Wang, K.S., Govindarajan, S., Redeker, A.G., Gerin, J.L., Houghton, M., 1988. A single antigenomic open reading frame of the hepatitis *delta* virus encodes the epitope(s) of both hepatitis *delta* antigen polypeptides p24 *delta* and p27 *delta*. J. Virol. 62, 594-599.

Westover, K.D., Bushnell, D.A., Kornberg, R.D., 2004. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science 303, 1014-1016.

Wettich, A., Biebricher, C.K., 2001. RNA species that replicate with DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 40, 3308-3315.

Wieland, T., Faulstich, H., 1991. Fifty years of amanitin. Experimentia 47, 1186-1193.

Wong, S.K., Lazinski, D.W., 2002. Replicating hepatitis *delta* virus RNA is edited in the nucleus by the small form of ADAR1. Proc. Natl. Acad. Sci. 99, 15118-15123.

Wu, H.N., Lin, Y.J., Lin, F.P., Makino, S., Chang, M.F., Lai, M.M., 1989. Human hepatitis *delta* virus RNA subfragments contain and autocatalytic cleavage activity. Proc. Natl. Acad. Sci. 86, 1831-1835.

Wu, T.T., Netter, H.J., Lazinski, D.W., Taylor, J.M., 1997. Effects of nucleotide changes on the ability of hepatitis *delta* virus to transcribe, process, and accumulate unit-length, circular RNA. J. Virol. 71, 5408-5414.

Xia, Y.P, Lai, M.M., 1992. Oligomerization of hepatitis *delta* antigen is required for both the *trans*-activating and *trans*-dominant inhibitory activities of the *delta* antigen. J. Virol. 66, 6641-6648.

Xia, Y.P., Yeh, C.T., Ou, J.H., Lai, M.M., 1992. Characterization of nuclear targeting signal of hepatitis *delta* antigen: nuclear transport as a protein complex. J. Virol. 66, 914-921.

Yamaguchi, Y., Filipovska, J., Yano, K., Furuya, A., Inukai, N., Narita, T., Wada, T., Sugimoto, S., Konarska, M.M., Handa, H., 2001. Stimulation of RNA polymerase II elongation by hepatitis *delta* antigen. Science 293, 124-217.

Yamaguchi, Y., Mura, T., Chanarat, S., Okamoto, S., Handa, H., 2007. Hepatitis *delta* antigen binds to the clamp of RNA polymerase II and affects transcriptional fidelity. Genes Cells 12, 863-875.

Yeh, T.S, Lo, S.J., Chen, P.J., Lee, Y.H., 1996. Casein kinase II and protein kinase C modulate hepatitis *delta* virus RNA replication but not empty viral particle assembly. J. Virol. 70, 6190-6198.

Yeh, T.S., Lee, Y.H., 1998. Assembly of hepatitis *delta* virus particles: package of multimeric hepatitis *delta* virus genomic RNA and role of phosphorylation. Virology 249, 12-20.

Yudkovsky, N., Ranish, J.A., Hahn, S., 2000. A transcription reinitiation intermediate that is stabilized by activator. Nature 408, 225-229.

Zheng, L., Roeder, R.G., Luo, Y., 2003. S phase activation of the histone H2B promoter by OCA-S, a coactivator complex that contains GAPDH as a key component. Cell 114, 255-266.

Zuccola, H.J., Rozzelle, J.E., Lemon, S.M., Erickson, B.W., Hogle, J.M., 1998. Structural basis of the oligomerization of hepatitis *delta* antigen. Structure 6, 821-830.

CHAPTER 2:

THE HUMAN RNA POLYMERASE II INTERACTS WITH THE TERMINAL STEM-LOOP REGIONS OF THE HEPATITIS *DELTA* VIRUS RNA GENOME

Valerie S. Greco-Stewart, Paul Miron, Abrahem Abrahem and Martin Pelchat

2.1 Statement of contribution

The manuscript entitled "The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome" was published in Virology in 2007 (357, 68-78). I contributed all laboratory-related experimental data with the exception of two co-immunoprecipitation experiments of the complete hepatitis *delta* genome in Figure 2.3B which were performed by A. Abrahem. P. Miron synthesized four of the RNA species used for co-immunoprecipitation experiments in Figure 2.3B and contributed to the discussion of Figure 2.5. All bioinformatic analyses were performed by M. Pelchat. Sequence-related images appearing in Figures 2.1, 2.4, and 2.5 were also contributed by M. Pelchat. I wrote this manuscript and editing, revisions, and additional insights were provided by M. Pelchat.

2.2 Abstract

The hepatitis *delta* virus (HDV) is an RNA virus that depends on DNA-dependent RNA polymerase (RNAP) for its transcription and replication. While it is generally accepted that RNAP II is involved in HDV replication, its interaction with HDV RNA requires confirmation. A monoclonal antibody specific to the carboxy terminal domain of the largest subunit of RNAP II was used to establish the association of RNAP II with both polarities of

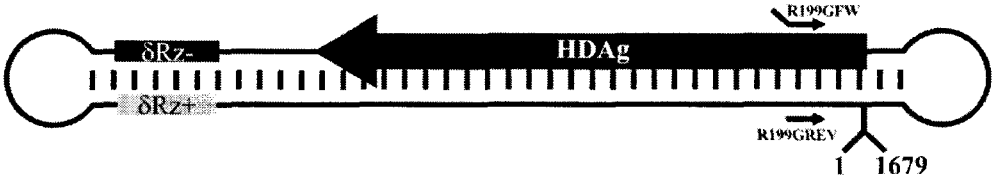
HDV RNA in HeLa cells. Co-immunoprecipitations using HeLa nuclear extract revealed that RNAP II interacts with HDV-derived RNAs at sites located within the terminal stem-loop domains of both polarities of HDV RNA. Analysis of these regions revealed a strong selection to maintain a rod-like conformation and demonstrated several conserved features. These results provide the first direct evidence of an association between human RNAP II and HDV RNA and suggest two transcription start sites on both polarities of HDV RNA.

Keywords: Hepatitis *delta* virus; RNA virus; RNA promoter; RNA polymerase II; co-immunoprecipitation.

2.3 Introduction

The hepatitis *delta* virus (HDV) is a subviral satellite virus of the hepatitis B virus (HBV) and is composed of a single-stranded, circular RNA molecule of approximately 1,700 nucleotides (Chen et al., 1986). Its genome forms an unbranched, rod-shaped structure with a high proportion of canonical base pairing, contains two complementary ribozyme motifs (i.e. *delta* ribozymes), and has a single open reading frame producing two viral proteins (HDAgs; Figure 2.1; Taylor, 2006). These two proteins are identical in sequence, except that the large HDAg (HDAg-L) contains 19 additional amino acids at its C-terminus resulting from RNA editing of the termination codon of the small HDAg (HDAg-S) gene (Casey et al., 1992; Wang et al., 1992; Weiner et al., 1988; Xia et al., 1990). Each protein has a distinct function: HDAg-S (195 amino acids) is essential for HDV accumulation, while the HDAg-L (214 amino acids) is required for virion assembly and is reported to be a dominant negative inhibitor of replication (Chang et al., 1991; Kuo et al., 1989; Ryu et al., 1992).

Figure 2.1: Schematic representation of the HDV RNA genome. The structure illustrates the superimposed features of both the genomic and antigenomic HDV polarities. The genome is depicted in its circular, unbranched rod-like conformation. The location of the genomic and antigenomic ribozymes and the open reading frame for HDAg are indicated. The position of the primers used to generate the R199G RNA molecule are indicated. The numbering is in accordance with Kuo et al. (1988a).


R199GFW
δRz-
HDAg
δRz+
R199GREV
1
1679

Although HDV requires HBV for encapsidation and transmission, it relies entirely on its host for replication. HDV replicates through a symmetrical rolling circle mechanism that involves only RNA intermediates (Chen et al. 1986; Kuo et al., 1988a; Kuo et al., 1988b; Lai, 2005; Taylor, 2006). Replication of the genomic circular monomer produces linear, multimeric strands which are subsequently self-cleaved and ligated, resulting in antigenomic polarity circular monomers. This process is repeated to generate the genomic RNA, which is found at a greater intracellular abundance than the antigenomic species (Chen et al., 1986). Because HDV does not encode its own replicase, host DNA-dependent RNA polymerases (RNAPs) are considered to be involved in the replication and transcription of HDV RNA (Lai, 2005).

Many studies using cultured cells and cell extracts have suggested that RNAP II might be responsible for HDV replication based on the sensitivity of the accumulation of HDV mRNA and processed unit-length genomic HDV RNAs to low levels of α-amanitin, a mycotoxin that inhibits DNA-dependent RNA synthesis by RNAP II (de Mercoyrol et al., 1989; Filipovska and Konarska, 2000; Fu and Taylor, 1993; Macnaughton et al., 1991). This hypothesis was substantiated by experiments using cells containing an α-amanitin-resistant allele of the largest subunit of RNAP II, which partially relieved transcription inhibition by α-amanitin (Modahl et al., 2000). RNAP II is furthermore speculated to be involved in the transcription of the HDAg mRNA because, *in vivo,* this mRNA was shown to be post-transcriptionally processed with a 5'-cap and a 3'-poly(A) tail, which are typical features of transcripts generated by RNAP II (Gudima et al., 2000; Nie et al., 2004). Additionally, in *in vitro* transcription assays using nuclear extract (NE) from HeLa cells and RNA derived from the left terminal stem-loop domain of HDV antigenomic RNA, synthesis of the

complementary strand was possible (Filipovska and Konarska, 2000). Accumulation of this RNA product was highly sensitive to α-amanitin. This sensitivity was partially abrogated in experiments conducted in NE from cells containing an α-amanitin-resistant allele of the largest subunit of RNAP II, suggesting the involvement of RNAP II in this reaction. Interestingly, the transcription did not proceed by *de novo* initiation, but rather by cleavage of the RNA template followed by extension of the new 3' end, generating a chimeric template/transcript product.

Conversely, the accumulation of the antigenomic RNA species has been shown to be resistant to a higher dose of α-amanitin, suggesting the involvement of another yet unknown RNAP in the life cycle of HDV (Macnaughton et al., 2002; Modahl et al., 2000). It has been hypothesized that RNAP I or an RNAP I-like polymerase might be involved in HDV replication because antigenomic HDV RNA was shown to accumulate in the nucleolus and its synthesis is associated with the RNAP I-specific transcription factor SL1 (Li et al., 2006).

Because general host DNA-directed RNAP II transcription is inhibited at relatively low doses of α-amanitin (de Mercoyrol et al., 1989), the observed inhibition of HDV RNA accumulation in the presence of this transcription inhibitor might be an indirect rather than a direct effect. It is difficult to differentiate between specific inhibition of HDV replication and non-specific toxic effects on host cells or the need for some additional host DNA-directed RNAP II transcripts. In addition, because the enzymes involved in HDV transcription/replication are host DNA-dependent RNAPs coerced to use an RNA template, sensitivity of RNAP II, or any other host RNAP(s), to α-amanitin might be affected by the nature of the template (i.e. RNA versus DNA). Such a phenomenon was reported for actinomycin D (Meienhofer and Atherton, 1973; Pelchat and Perreault, 2004) and was

recently observed with tagetitoxin (Pelchat et al., 2001), a potent inhibitor of bacterial RNAP and nuclear DNA-dependent RNAP III (Mathews and Durbin, 1994). Like α-amanitin, tagetitoxin is believed to inhibit the elongation phase of RNA synthesis by interacting with the ternary complex that contains the RNAP, the template, and the nascent RNA (Gong et al., 2004; Mathews and Durbin, 1994). RNAP inhibitors must thus be used with caution when utilized as a means of identifying an RNAP involved in RNA-dependent transcription.

Although HDV replication has been shown to be sensitive to low concentrations of α-amanitin, suggesting RNAP II activity, no direct interaction between RNAP II and HDV RNA has been reported. The present study was undertaken (i) to confirm the association of RNAP II with HDV RNA, and (ii) to identify the regions on the HDV RNA genome responsible for this interaction. To this end, we took advantage of an antibody raised against the carboxy terminal domain (CTD) of RNAP II to clarify the interaction of RNAP II with HDV RNA by directly examining their association *in vivo* within HeLa cells. We identified the regions involved in this interaction by co-immunoprecipitating various HDV-derived RNAs with RNAP II using HeLa nuclear extract (NE) and the α-RNAP II antibody. Then, we used covariation analysis on natural variants to analyze the conserved nucleotides and structural motifs of the interaction sites of RNAP II on HDV RNA. Our analyses support a role for the human RNAP II in the life cycle of HDV by demonstrating its association with HDV RNA *in vivo* and with the terminal stem-loop domains of both polarities of HDV RNA *in vitro*.

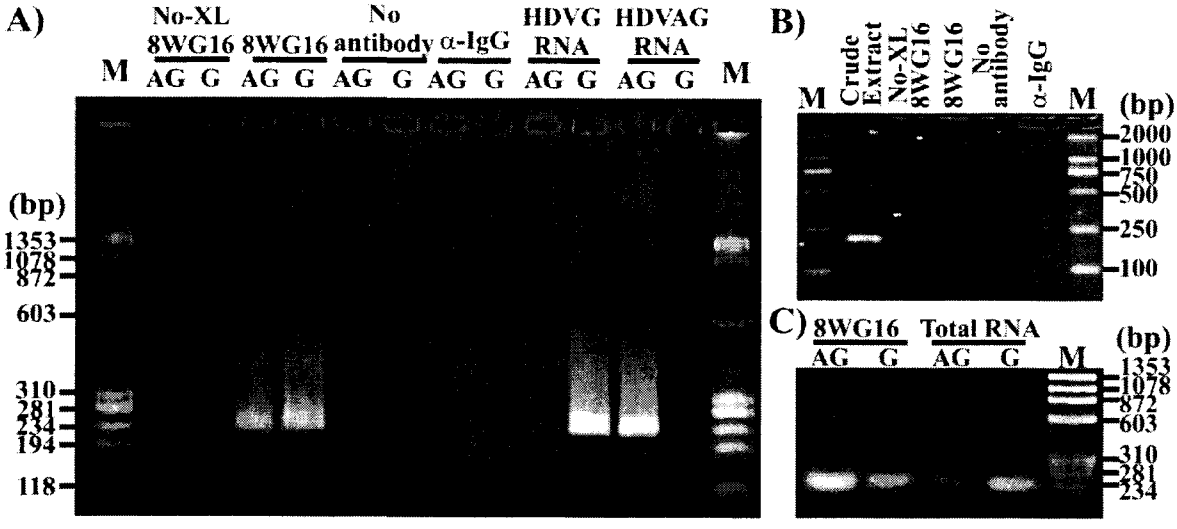
2.4 Results

RNAP II interacts with HDV RNA within HeLa cells.

To establish the direct interaction between RNAP II and HDV RNA *in vivo*, we immunoprecipitated RNAP II from HeLa cells containing both polarities of HDV RNA and tested for the presence of viral RNA in the immunoprecipitated samples. To this end, we used the monoclonal antibody (mAb) 8WG16 which was raised against wheat germ RNAP II and found to recognize epitopes conserved among the CTDs of RNAP IIs derived from a variety of different eukaryotes (Thompson et al., 1989). When this commonly used antibody is tested against HeLa NE proteins, only one band corresponding to the largest subunit of RNAP II is revealed by Western blot. To obtain cells harbouring both polarities of HDV RNA, HeLa cells were transiently transfected with a dimeric HDV transcript of genomic polarity and a eukaryotic expression vector overexpressing HDAg-S, which has been found to be essential for HDV RNA accumulation (Kuo et al., 1989).

Four days post-transfection, the cells were harvested and the RNA-protein complexes cross-linked with formaldehyde *in vivo* using the ribonucleoprotein immunoprecipitation (RIP) assay developed by Niranjanakumari et al. (2002). Following immunoprecipitation and high-stringency washes to ensure the specificity of the binding, the cross-links were reversed by heating the samples. Recovered nucleic acids were then extracted and subjected to RT-PCR. For this step, we used primers for a cDNA fragment corresponding to what we refer to as the right-terminal stem-loop domain of HDV (Figure 2.1, R199GFW and R199GREV). In addition to the HDV sequence (nucleotides 1541 to 61), the 232 bp cDNA fragment generated after RT-PCR included a promoter for the T7 RNAP and recognition sites for restriction enzymes. When we used the 8WG16 mAb to immunoprecipitate the sample, cDNA fragments corresponding to both polarities of R199 were detected (Figure 2.2,

Figure 2.2: Association of RNAP II with HDV RNA within HeLa cells. HeLa cells transfected with HDV RNA were used for the preparation of cross-linked lysates for immunoprecipitation using the 8WG16 mAb, no antibody, or the control goat anti-mouse IgG mAb. Non-cross-linked lysate immunoprecipitated using the 8WG16 mAb is also presented. A) Association of RNA with both polarities of HDV RNA. HDVG and HDVAG RNA lanes are RT-PCR product controls using *in vitro* transcribed HDV RNA of genomic and antigenomic polarity, respectively. AG and G lanes contain RT reactions performed using primers complementary to either antigenomic or genomic HDV RNA corresponding to the R199 region. B) Control RT-PCR using the β-actin cDNA to demonstrate the specificity of the co-immunoprecipitations for HDV RNAs. The co-immunoprecipitated samples are as described above. C) Semi-quantitative RT-PCR amplifications of both polarities of HDV RNA before and after co-immunoprecipitation with the the 8WG16 mAb. Lanes M contain DNA ladders: A and C, PhiX174 digested with *Hae*III; B, GelPilot Mid Range DNA ladder (Qiagen).


A)
No-XL
8WG16
8WG16
No antibody
α-IgG
HDVG RNA
HDVAG RNA
M
AG G AG G AG G AG G AG G AG G
M
(bp)
1353
1078
872
603
310
281
234
194
118
B)
M
Crude Extract
No-XL
8WG16
8WG16
No antibody
α-IgG
M
(bp)
2000
1000
750
500
250
100
C)
8WG16
Total RNA
AG G AG G
M
(bp)
1353
1078
872
603
310
281
234

8WG16 lanes). Cloning and subsequent sequence analysis confirmed the identity of the HDV-derived cDNA fragments. These cDNA fragments were detected neither in the control reactions lacking antibody nor in the goat α-mouse IgG mAb control reactions (Figure 2.2), indicating that cDNAs obtained from this assay were specific to the samples immunoprecipitated with the 8WG16 mAb. Moreover, these fragments were also absent from samples that were not subjected to the cross-linking reaction prior to co-immunoprecipitation (Figure 2.2, No-XL 8WG16 lanes). To ascertain that the association with the RNAP II complex is specific for HDV RNA, RT-PCR was also performed on an abundant cellular RNA (β-actin mRNA). As shown in Figure 2.2B, although the β-actin cDNA fragment was easily amplified from crude extract from HeLa cells, this fragment was not detected in the various immunoprecipitation reactions performed above. Overall, the presence of HDV RNA fragments in the cross-linked, immunoprecipitated samples provides evidence that formaldehyde covalently cross-links the largest subunit of RNAP II or a complex containing RNAP II to HDV RNA in HeLa cells. To corroborate this finding, we used a pair of primers corresponding to the other terminal stem-loop region of the HDV RNA genome (i.e. L213, nucleotides 687 to 900); identical results were obtained (data not shown).

The primers used in the RT reactions were designed to differentiate between both polarities of HDV RNA. Specifically, primers corresponding to sequences of the genomic (i.e. R199GFW) and antigenomic (i.e. R199GREV) strand of HDV RNA were only able to reverse transcribe from either the antigenomic or the genomic polarity of HDV RNA, respectively (Figure 2.2, HDVG and HDVAG RNA lanes). Since both polarities of HDV RNA were recovered when the cross-linked sample was immunoprecipitated with the

8WG16 mAb, we conclude that both polarities of HDV RNA associate with RNAP II within HeLa cells, which directly establishes the association of the human RNAP II with HDV RNA and strongly suggests its involvement in the synthesis of both strands of HDV RNA.

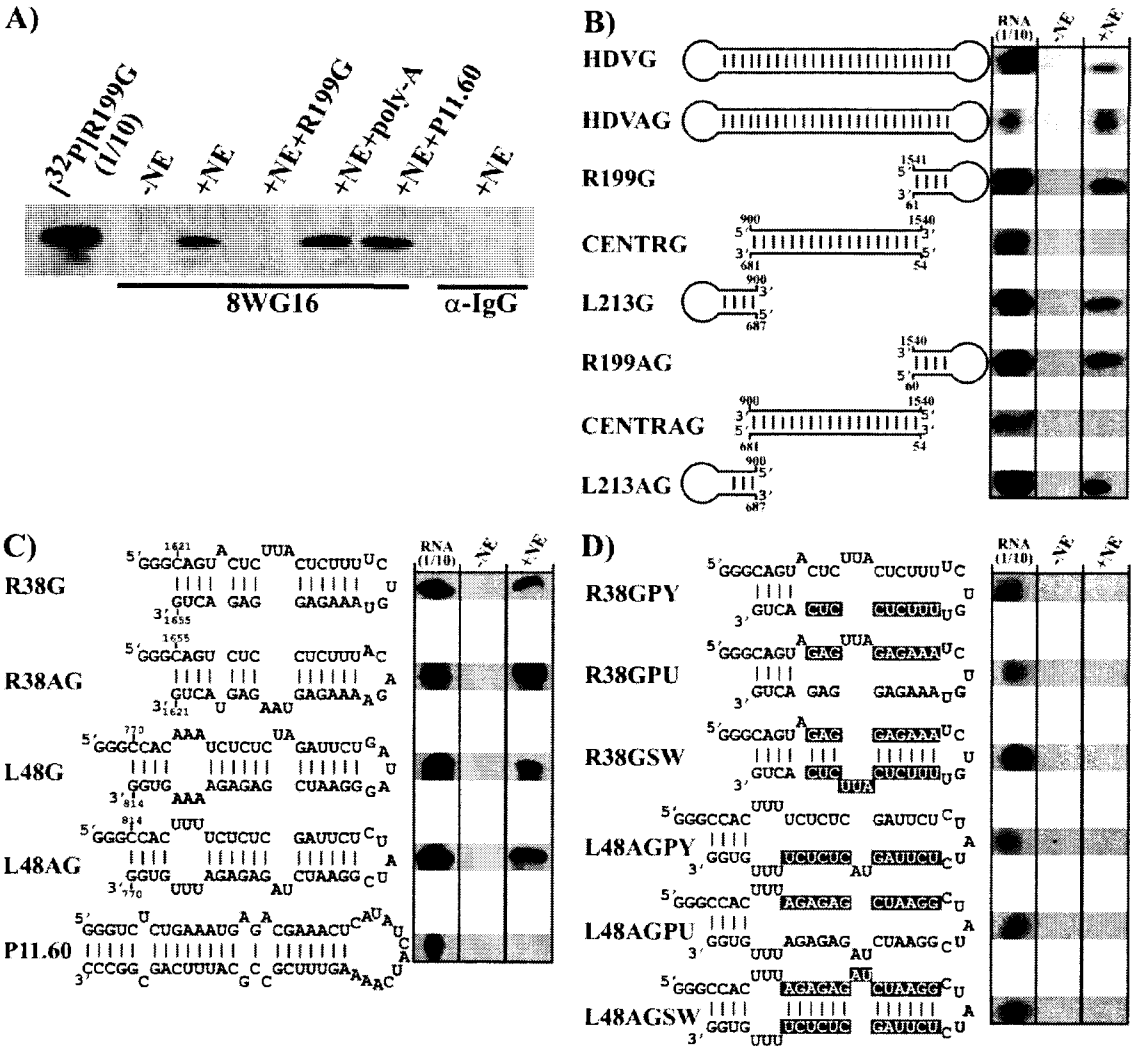
Although the RT-PCR reactions were not quantitative, the intensity of the bands corresponding to both polarities of HDV RNA co-immunoprecipitated using the 8WG16 mAb was similar. Because it has been established that genomic HDV RNA accumulates more than antigenomic HDV RNA in infected cells (Chen et al. 1986), our results suggest a more efficient binding of RNAP II to the antigenomic polarity of HDV RNA. To confirm the preferential interaction of RNAP II with this polarity, semi-quantitative RT-PCR was performed on RNA samples taken either before or after co-immunoprecipitation with the 8WG16 mAb. Because no human RNA is expected to co-immunoprecipitate with RNAP II, it was not possible to normalize the different samples using an internal RNA control. To circumvent this difficulty, amplification curves were performed and used to determine the number of PCR cycles at which all the samples were within their linear phase of amplification. Then, the intensities of the cDNA bands corresponding to both polarities of HDV RNA were compared within each sample. Prior to co-immunoprecipitation, we found a genomic:antigenomic HDV RNA ratio of 1.99:1. Although less than reported using infected liver cells, this value is in conformity with a greater intracellular abundance of genomic HDV RNA when compared to antigenomic HDV RNA (Chen et al., 1986). By performing a similar calculation for the sample co-immunoprecipitated with the 8WG16 mAb, we found that the genomic:antigenomic HDV RNA ratio shifted to 1:1.56. As a result, under our conditions, antigenomic HDV RNA was enriched 3.1-fold after co-immunoprecipitation with the 8WG16 mAb. This enrichment indicates a preferential interaction of RNAP II with antigenomic HDV RNA.

RNAP II interacts with the terminal stem-loop domains of both polarities of HDV RNA in HeLa NE.

The previous experiments provide evidence of the association of RNAP II with both the genomic and antigenomic HDV RNAs within HeLa cells. To determine the regions of the HDV RNA genome responsible for this interaction, we investigated the association of RNAP II with various HDV-derived RNAs *in vitro* using HeLa NE. As an initial step, both polarities of the HDV RNA genome were arbitrarily divided in three domains: 213 nt RNAs corresponding to what we refer to as the left terminal stem-loop (L213; nt 687 to 900), RNAs corresponding to the double-stranded central region (CENTR; nt 681 to 54 and nt 900 to 1540), and 199 nt RNAs corresponding to the right terminal stem-loop (R199; nt 1540 to 60). To guarantee that both RNA molecules forming the CENTR RNAs annealed, both strands were mixed, heated, slowly cooled, and tested for double-strand formation. Electrophoresis on a native acrylamide gel confirmed that most of the RNA molecules forming the CENTR RNAs annealed under our conditions (data not shown). The various radiolabeled HDV-derived RNAs were allowed to form RNA-protein complexes with HeLa NE proteins. The 8WG16 mAb was then added, the samples were immunoprecipitated using protein-G agarose beads, and the co-immunoprecipitated complexes were then subjected to denaturing gel electrophoresis.

A typical gel using a radiolabeled RNA corresponding to the right terminal stem-loop domain of the genomic polarity of HDV RNA (i.e. R199G) is shown in Figure 2.3A. Under our conditions, a band corresponding to the radiolabeled RNA was detected when the sample was immunoprecipitated using the 8WG16 mAb in the presence of HeLa NE. This band was not detected in the absence of NE or when the 8WG16 mAb was replaced with the α-mouse IgG mAb, indicating that the signal observed was not the result of non-specific binding of

Figure 2.3: Interaction of RNAP II with HDV-derived RNAs. A) Co-immunoprecipitation of radiolabeled R199G with RNAP II in HeLa NE. 20-fold excess non-radioactive R199G, poly-A RNA, or P11.60 (Pelchat et al., 2002) were used in competition experiments. B) Interaction of RNAP II with various HDV-derived RNAs. C) Interaction of RNAP II with the extremities of the rod-like HDV secondary structure of both polarities. D) Interaction of RNAP II with mutants affecting the secondary structure and the purine/pyrimidine polarization. The numbering is in accordance with Kuo et al. (1988a). One tenth of the RNAs used for each co-immunoprecipitation are used as markers. The co-immunoprecipitations presented in B) to D) were performed in the presence of a 20-fold excess P11.60, as a non-specific RNA competitor.


A)
[32P]R199G (1/10)
-NE
+NE
+NE+R199G
+NE+poly-A
+NE+P11.60
+NE
8WG16
α-IgG
B)
RNA (1/10)
-NE
+NE
HDVG
HDVAG
R199G
CENTRG
L213G
R199AG
CENTRAG
L213AG
C)
R38G
R38AG
L48G
L48AG
P11.60
D)
R38GPY
R38GPU
R38GSW
L48AGPY
L48AGPU
L48AGSW

R199G to the protein-G agarose beads or to the 8WG16 mAb. To ensure the specificity of the interaction, various RNA competitors were tested. When an excess of unlabeled homologous competitor was used, the intensity of the band corresponding to the radiolabeled R199G was considerably diminished. When poly(A) RNA was used as an unrelated RNA competitor, no significant decrease in the co-immunoprecipitation of the radiolabeled R199G was observed. When acting on a DNA promoter, RNAP II binds to double-stranded nucleic acid. To verify that the observed association between RNAP II and R199G was not caused by non-specific affinity for double-stranded RNA, a small RNA hairpin (Figure 2.3C, P11.60) derived from the peach latent mosaic viroid (PLMVd) that acts as an RNA promoter for the *Escherichia coli* RNAP (Pelchat et al., 2002) was used as a competitor. The interaction between RNAP II and R199G was not significantly decreased by addition of an excess amount of P11.60. Taken together, these observations demonstrate that RNAP II interacts specifically with R199G within HeLa NE.

The interaction of RNAP II with the various regions of both genomic and antigenomic HDV RNA is summarized in Figure 2.3B. Because several RNA molecules were tested, only the lanes corresponding to the amount of RNA used and co-immunoprecipitation reactions in the presence of the non-specific RNA competitor P11.60 with either the absence (NE-) or the presence (NE+) of HeLa NE are presented. Under our conditions, both polarities of the complete HDV RNA genome co-immunoprecipitated specifically using the 8WG16 mAb (Figure 2.3B, HDVG and HDVAG). When the band intensities of the co-immunoprecipitated RNAs were normalized using the amount of radiolabeled RNA used, the antigenomic HDV RNA (HDVAG) was found to be co-immunoprecipitated 1.9-fold more efficiently than the genomic HDV RNA (HDVG). These

results are in agreement with our previous observations using cross-linked lysates and indicate that RNAP II interacts more efficiently with the antigenomic polarity of HDV RNA.

Next, the various HDV-derived RNAs corresponding to the three arbitrary domains of both polarities were tested to determine the regions on the HDV RNA genome responsible for RNAP II interaction. Radioactive bands corresponding to the stem-loop domains of both polarities of HDV RNA (i.e. R199G, R199AG, L213G, and L213AG) were detected, whereas bands corresponding to the central regions of both polarities of HDV RNA (i.e. CENTRG and CENTRAG) were not. These results indicate that RNAP II interacting sites are located predominantly between nt 1540 to 60 and nt 687 to 900 on both polarities of the HDV RNA genome.

To further refine the regions interacting with RNAP II, smaller HDV-derived transcripts were synthesized and were used in the co-immunoprecipitation reactions. Specifically, R38 and L48, which correspond to 38 and 48 nt RNAs located at the tip of the right and the left hairpin domains, respectively, were tested for RNAP II interaction. Using the same co-immunoprecipitation conditions as above, we found that both polarities of R38 and L48 were able to co-immunoprecipitate using the 8WG16 mAb (Figure 2.3C), indicating that the terminal stem-loop domains of both polarities of HDV RNA can be reduced to smaller hairpins without significantly affecting their interaction with RNAP II. As an additional negative control, P11.60 was tested directly in the co-immunoprecipitation reactions. The small PLMVd-derived RNA hairpin did not co-immunoprecipitate with RNAP II when using the 8WG16 mAb (Figure 2.3C), which further supports the specificity of the co-immunoprecipitation experiments. Because RNAP II complexes are composed of many subunits and other related factors (Cramer, 2002), we cannot conclude that the binding of the RNAP II core enzyme to the various RNA molecules is direct. Also, we cannot

exclude the possibility that other RNAPs might be able to recognize these RNA fragments. However, these results provide direct evidence that a complex containing the CTD of the largest subunit of RNAP II interacts specifically with the terminal stem-loop domains of both polarities of HDV RNA in HeLa NE.

Sequence and secondary structure analysis of the terminal stem-loop domains of both polarities of HDV RNA.

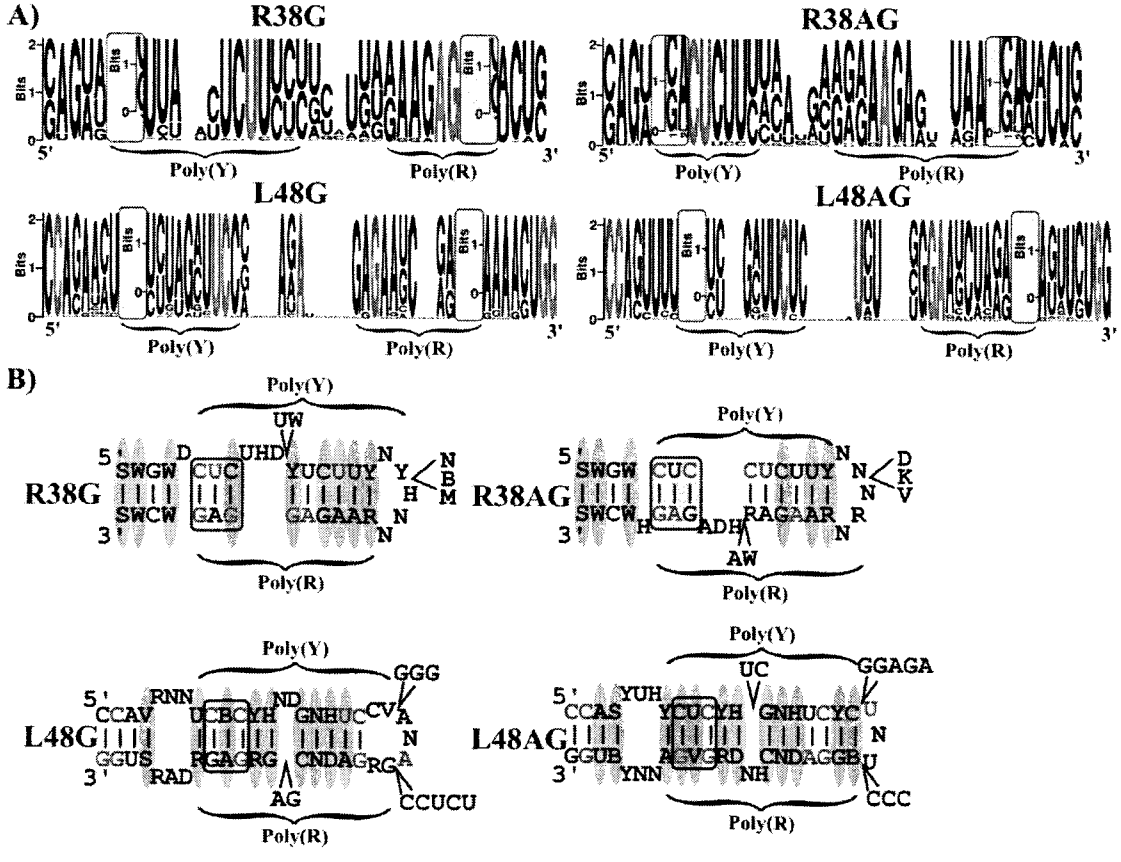
To interact with HDV RNA, RNAP II must recognize specific features (i.e. sequences and/or secondary structures) on both polarities of the terminal stem-loop domains. We have taken advantage of the large number of known nucleotide sequences of *in vivo* selected variants to identify the features of terminal stem-loop domains of HDV RNA by studying positions of covariation and nucleotide conservation.

We collected the sequences corresponding to R38 and L48 from a total of 81 variants representing the various HDV genotypes indexed in the Subviral RNA Database (Rocheleau and Pelchat, 2006). First, we aligned all the RNA sequences using the Clustal software (Thompson et al., 1994), which generated a tree showing the sequence differences between HDV genotypes. Due to the imperfect alignment of the analyzed regions among the different HDV sequences, we applied alignments to smaller datasets according to HDV genotypes because of their lower degree of heterogeneity. For each of these genotypes, the consensus features of the RNA secondary structures generated by analysis of base-pair covariation were determined. Finally, we combined alignments by insertion of gaps so that homologous motifs (i.e. secondary structures) were superimposed. Using the consensus sequence derived from the alignment, the most thermodynamically-stable secondary structure was predicted and re-adjusted based on the base-pair covariation data. Sequence conservation and the resulting

consensus RNA secondary structures for both polarities of R38 and L48 are summarized in Figure 2.4 together with the fully conserved nucleotides (red nucleotides) and the location of the base-pairs predicted by the covariation analysis (blue circles).

Despite heterogeneity among the different sequences and genotypes, base-pair covariation analysis indicates a strong selective pressure to maintain the rod-like conformation of these regions in which most of the base-pairs can be predicted. A few nucleotides within this region are conserved across all the HDV sequences examined (Figure 2.4, red nucleotides). Considering the degree of nucleotide divergence in the adjacent nucleotides, the observed conservation suggests a functional role for the conserved nucleotides. Strikingly, a short double-stranded stem (CUC/GAG) was found to be exceptionally conserved (Figure 2.4B, yellow box). In fact, most of the nucleotides comprising this short CUC/GAG stem are completely conserved in all the analyzed variants and the other positions exhibit only minor variations. Notably, we observed a strong polarization in the purine/pyrimidine content near the tip of the rod up to the end of the CUC/GAG stem in all the sequences analyzed. A region of 12 nt upstream of the terminal loop of R38G and R38AG exhibits an average of 91.0 ±1.0 % pyrimidine content which is matched by a region on the opposite strand containing almost exclusively purines (Figures 2.4B). In the case of L48G and L48AG, the sequence upstream of the loop contains 86.7 ± 5.9% pyrimidines which is matched by a region on the opposite strand having 82.0 ± 5.8% purines (Figure 2.4B).

Figure 2.4: Primary sequence conservation and proposed secondary structure of R38G/- and L48G/-. The nucleotide motif (A) and the consensus secondary structure (B) derived from the analysis of the alignment of 81 HDV natural variants are presented. Completely conserved nucleotides (red), observed base-pair covariation (blue ovals), and the conserved CUC/GAG motifs (yellow) are indicated. IUPAC 1-letter code abbreviations are used for the identity of the nucleotides on the secondary structures showing less than 95% conservation.


A)
R38G
R38AG
L48G
L48AG
Bits
5'
3'
Poly(Y)
Poly(R)
B)
R38G
R38AG
L48G
L48AG
Poly(Y)
Poly(R)

The secondary structure and the polarization in the purine/pyrimidine content near the tip of the rod are required for RNAP II interaction.

Sequence analysis of natural HDV variants suggests that the maintenance of the rod-like conformation and the polarization in the purine/pyrimidine content near the tip of the rod of both polarities of HDV RNA might be required for RNAP II interaction. We therefore introduced substitutions to R38G and L48AG aimed at destabilizing either of these conserved features. First, the involvement of the purine-rich segments located downstream of the terminal loops was investigated. R38GPY and L48AGPY are derivative RNAs in which the purine-rich segments were mutated to the pyrimidine-rich sequences located on their respective upper strands. Upon co-immunoprecipitation using the 8WG16 mAb, no significant association of RNAP II with either R38GPY or L48AGPY was detected (Figure 2.3D). Similar results were obtained when R38GPU and L48AGPU, two derivative RNAs in which the pyrimidine-rich segments located upstream of the terminal loops were mutated to purine-rich sequences, were used. Specifically, no co-immunoprecipitation of either R3G8PU or L48AGPU with RNAP II was detected (Figure 2.3D).

Because the mutations found in these four mutants also result in the disruption of their predicted rod-like structures, the absence of RNAP II interaction with these RNAs could be the result of the loss of their rod-like conformation. R38GSW and L48AGSW, which contain compensatory mutations that restore the predicted secondary structure of these hairpins, were therefore synthesized and tested for association with RNAP II. No complex formation was detected when using either of these "flip" mutants (Figure 2.3D). Taken together, these results are in accordance with the sequence analysis of natural HDV variants and indicate that the association of RNAP II with these regions is dependent on the conformation and on the polarization in the purine/pyrimidine content near the tip of the rod.

2.5 Discussion

The mechanism of RNA synthesis is highly conserved and can be roughly divided into three distinct phases: initiation, elongation, and termination. For initiation to occur, RNAPs search for and bind to promoter sequences to form a "closed complex"; thus, association of RNAP II would indicate sites of transcription (Smale and Kadonaga, 2003). Here, we have used an antibody raised against the CTD of RNAP II to clarify RNAP II interaction with HDV RNA by studying the first step of the transcription process (i.e. promoter recognition). This approach was used to directly demonstrate specific RNAP II interaction with HDV RNA, both *in vivo* within HeLa cells and *in vitro* using HeLa NE. Using various HDV-derived RNAs with HeLa NE proteins, we have shown that direct and specific RNAP II interaction occurs *in vitro* predominantly via small domains located at the extremities of both polarities of the rod-like secondary structure of the HDV RNA genome, and that RNAP II might bind more efficiently to the antigenomic polarity of HDV RNA.

It is possible that the observed interaction of RNAP II with these regions might not lead to transcription initiation. When binding to a DNA promoter, the RNAP II found within an initiation site is hypophosphorylated at its C-terminal domain (CTD) and is distinct from the elongating or terminating forms of the enzyme that are hyperphosphorylated (Hahn, 2004; Sims et al., 2004). In this study, we have used the 8WG16 mAb, which specifically recognizes the unphosphorylated RNAP II CTD (Patturanjan et al., 1998); the extremities of both polarities of HDV RNA are thus likely to be sites of pre-initiation complex formation (i.e. active promoters; Thompson et al., 1989; Smale and Kadonaga, 2003). Moreover, most of these interaction sites are in agreement with previous reports.

The 5' end of HDAg mRNA corresponds to an initiation site located within R38G (Gudima et al., 2000). It is thus likely that the right terminal stem-loop domain contains an

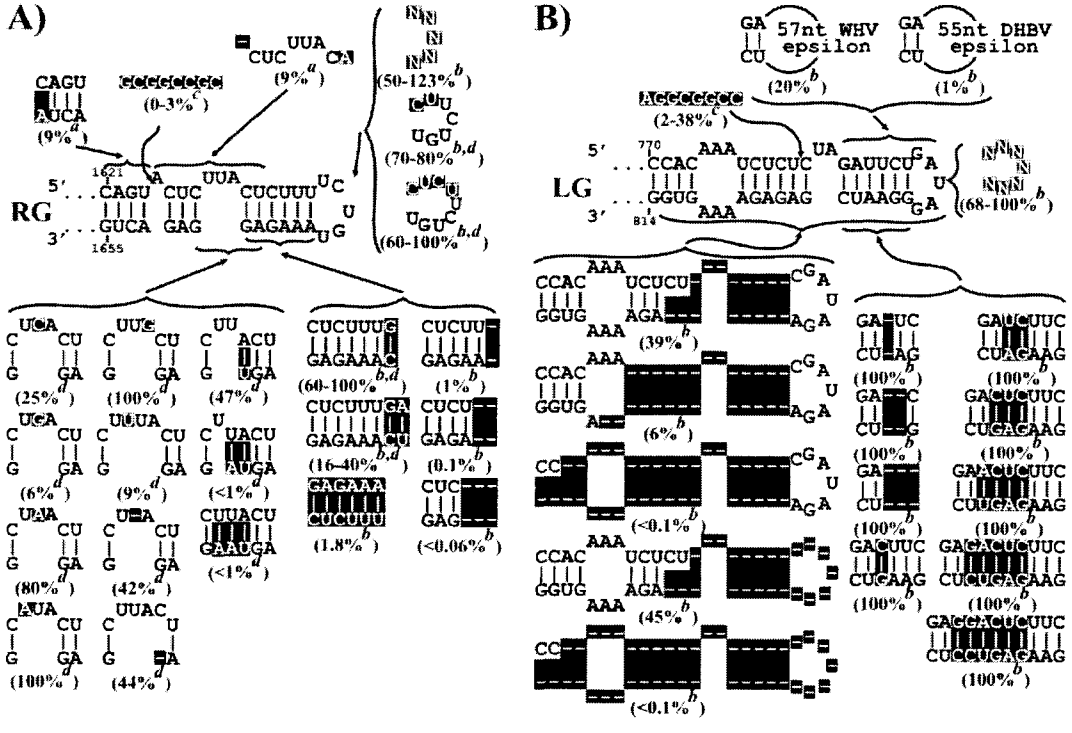
RNA promoter recognized by RNAP II. RNA fragments similar to both R199G and L213AG were reported to be able to initiate polymerization or promote 3' end nucleotide addition *in vitro* using HDV-derived RNAs and HeLa NE (Beard et al., 1996; Filipovska and Konarska, 2000). Interestingly, transcription from the RNA template similar to L213AG did not proceed by *de novo* initiation, but rather by cleavage of the RNA template, followed by extension of the new 3' end, generating a chimeric template/transcript product (Filipovska and Konarska, 2000). An RNA fragment similar to L213G was also able to generate a specific product in this system, despite the authors' claim that the signal detected did not correspond to their typical chimeric template/transcript product (Filipovska and Konarska, 2000). Here, we have shown that the right terminal stem-loop region of antigenomic HDV RNA (i.e. R199AG) can also interact with RNAP II and we have demonstrated that these four RNA regions can be reduced significantly in length without affecting RNAP II interaction. In summary, our results suggest the involvement of RNAP II in the synthesis of both genomic and antigenomic HDV RNAs. In addition, because RNAP II interacts with the extremities of both polarities of the HDV RNA genome, our data suggest the existence of two transcription/replication start sites for both genomic and antigenomic HDV RNAs. Moreover, it is possible that other RNAP(s) are involved in HDV transcription/replication, given that only the 8WG16 mAb was tested.

The RNA fragments interacting with RNAP II were further investigated by taking advantage of a large number of natural HDV variants. Analysis of base-pair covariation indicated that although changes in the primary sequence of these segments might be tolerated, their respective secondary structures are most likely critical for HDV replication/accumulation. The conserved features of the two extremities of the rod-shaped HDV RNA are very similar. Our analysis revealed a highly-ordered RNA secondary

structure that is comprised of regions of double-stranded RNA with a conserved CUC/GAG triple base-pair motif and one strand of pyrimidines upstream of a 3-8 nt terminal loop followed by a complementary strand of purines. Because these features are found near all the terminal loops of HDV RNA which were shown to interact with RNAP II, it is tempting to suggest the involvement of these features in RNAP II binding. Accordingly, mutagenesis of R38G and L48AG, affecting the conformation and the polarization in the purine/pyrimidine content near the tip of the rod, indicated that the interaction of RNAP II with these regions is dependent on these features.

Previous studies have reported extensive mutagenesis on the extremities of the rod-like structure of HDV RNA and have observed their effects on HDV RNA accumulation *in vivo* (Beard et al., 1996; Gudima et al., 1999; Gudima et al., 2006; Wang et al., 1997; Wu et al., 1997). Figure 2.5 presents a summary of these published results. For the right terminal extremities, variation of the terminal loop sequence and size had no effect on HDV RNA accumulation (Gudima et al., 1999; Wu et al., 1997). In contrast, disruption of the overall rod-like conformation of this region or even small changes in the terminal stem had major effects on accumulation of HDV RNA (Beard et al., 1996; Gudima et al., 1999; Wu et al., 1997). In addition, HDV RNA accumulation was found to be sensitive to the size of the 3 nt external bulge, which includes the putative mRNA initiation site (Gudima et al., 1999). Our analysis indicates that this bulge can be on either side of the stem-loop structures and RNAP II interaction still occur (compare R38R/R38AG). However, correct positioning of this non-paired region might be critical for the initiation stage of HDAg mRNA synthesis. Noteworthy, a "flip" mutant in the terminal stem was found to greatly affect HDV RNA accumulation (Wu et al., 1997) in complete agreement with our results. The association of

Figure 2.5: Summary of published mutagenesis results on the extremities of the rod-like structure of HDV RNA. The mutations and the resulting consequences on HDV RNA accumulation were taken from *a*) Beard et al., 1996, *b*) Wu et al., 1997, *c*) Wang et al., 1997, and *d*) Gudima et al., 1999. Reported mutagenesis on either the right- or the left-terminal extremities of the rod-like secondary structure of the HDV RNA are presented in A) or B), respectively.


A)
RG
CAGU
AUCA
(9%a)
(0-3%c)
(9%a)
(50-123%b)
(70-80%b,d)
(60-100%b,d)
(25%d)
(100%d)
(47%d)
(6%d)
(9%d)
(<1%d)
(80%d)
(42%d)
(<1%a)
(100%d)
(44%d)
(60-100%b,d)
(1%b)
(16-40%b,d)
(0.1%b)
(1.8%b)
(<0.06%b)
B)
57nt WHV epsilon
55nt DHBV epsilon
(20%b)
(1%b)
(2-38%c)
LG
(68-100%b)
(39%b)
(6%b)
(<0.1%b)
(45%b)
(<0.1%b)
(100%b)

RNAP II with this region and initiation of transcription are thus dependent on the conformation and on the polarization in the purine/pyrimidine content near the tip of the rod.

Conversely, the left-terminal region was found to be highly tolerant to nucleotide changes within the loop and to insertions as well as deletions in the adjacent stem region (Wu et al., 1997). Only the removal of large segments or the insertion of various non-HDV sequences in the conserved CUC/GAG motifs resulted in loss of accumulation of HDV RNA (Gudima et al., 2006; Wang et al., 1997). Interestingly, most genomes containing large inserts lost their inserted sequences suggesting the involvement of this region in HDV RNA replication (Gudima et al., 2006). The divergence of these reports with our results might be explained by the likely involvement of another RNAP, such as RNAP I. Moreover, because RNAP II was found to be associated with the stem-loop region of both polarities of HDV RNA, suggesting at least two replication start sites for each polarity of HDV RNA, complementation of replication impaired mutants of the left-terminal stem-loop structure might occur. Since the right extremity contains sequences required for HDAg mRNA transcription, such complementation might not occur, causing the right terminal stem-loop region to be more sensitive to mutation.

We cannot exclude the possibility that the conserved features of the HDV extremities might also be linked to activities unrelated to RNAP II binding, such as RNA stabilization, RNA transportation, or RNA encapsidation. For example, these polypyrimidine/polypurine tracts could serve as binding sites for other RNA-binding proteins. Typical representatives of these types of proteins are the polypyrimidine tract-binding proteins (PTBs) and the heterogeneous nuclear ribonucleoproteins (hnRNPs), which, in uninfected cells, are involved in a variety of processes, such as pre-mRNA splicing in the nucleus (Garcia-Blanco et al., 1989; Swanson and Dreyfuss, 1988). However, we recently reported the finding that the

polypyrimidine tract-binding protein-associated splicing factor (PSF) is also able to bind specifically to both polarities of HDV through the terminal stem-loop domains of the rod-like structure (Greco-Stewart et al., 2006). Because PSF is known to bind to the CTD of RNAP II (Emili et al., 2002) and PSF binding sites correlate with the locations on the HDV RNA genome found to be associated with RNAP II, it is tempting to suggest that by interacting with both an HDV RNA promoter and RNAP II, PSF or another polypyrimidine/polypurine tract-binding protein might provide docking sites or assist RNAP(s) in the formation of pre-initiation complexes. The co-immunoprecipitation assay we used in this study will be a perfect tool in determining the content of the RNAP II complex interacting with HDV RNA and the contribution of additional *trans*-acting host factors in RNA promoter recognition by RNAP II. For instance, although HDAg-S was demonstrated to be essential for HDV accumulation (Kuo et al., 1989), our results indicate that this protein is not required for the interaction of RNAP II with HDV RNAs in HeLa NE since it was absent from all co-immunoprecipitation assays. This observation correlates with previous results suggesting that HDAg-S could regulate the elongation phase of transcription by binding to RNAP II directly and by displacing the negative elongation factor (Yamaguchi et al., 2001).

Interestingly, the features and the locations of RNAP II interaction sites on HDV RNA share structural similarities with DNA-dependent RNAP binding and/or initiation sites reported for other related subviral RNAs. *Escherichia coli* RNAP can bind to and transcribe from a stem-loop domain derived from the PLMVd genome (Pelchat et al., 2002; Pelchat and Perreault, 2004). Any structural alterations that disrupt the overall rod-like conformation of the RNA template were shown to interfere with *de novo* transcription by *E. coli* RNAP (Pelchat et al., 2002; Pelchat and Perreault, 2004). Replication of the avocado sunblotch viroid genome is proposed to initiate within large terminal loops (Navarro and Flores, 2000).

Wheat germ RNAP II was shown to bind to the large terminal loops of potato spindle tuber viroid genome (Goodman et al., 1984), and it was recently reported that potato NE was able to initiate transcription of this viroid within a terminal stem-loop region of the RNA genome (Kolonko et al., 2006).

It is now known that the observed structural similarities between bacterial RNAP and eukaryotic RNAP II involve not only similarity in overall structural organization, but also detailed similarity in folding topologies of subunits (Cramer, 2002). In view of the similarities between the RNAPs and the similarities between the various subviral RNA domains reported to bind DNA-dependent RNAP, it is tempting to suggest a common mechanism of RNAP recognition acting on subviral RNA species. It seems clear that a defined hairpin domain is a common feature of RNA promoters for these DNA-dependent RNAPs. Such terminal stem-loop secondary structures with RNA features specific for each RNAP could thus represent "universal" RNA binding sites and/or promoter elements for these DNA-dependent RNAPs.

2.6 Materials and methods

Synthesis of HDV-derived RNAs.

Polymerase chain reaction (PCR) with Vent polymerase (New England Biolabs; NEB) was used to amplify DNA fragments for the synthesis of both genomic and antigenomic cDNA fragments from a derivative of pBluescriptKS$^+$ (Stratagene) containing dimeric HDV genomic cDNA (Kuo et al., 1988a) flanked by T7 and T3 promoter recognition sites. P11.60 was synthesized as previously described (Pelchat and Perreault, 2004). All R38 and L48 derivatives were transcribed from double-stranded DNA oligonucleotides possessing a T7 RNA promoter. DNA products were used for *in vitro* run off transcription using T7 RNA

polymerase (NEB). HDV dimeric RNA genomes were similarly synthesized using linearized plasmid containing a dimer of HDV cDNA as a template for transcription. After incubation of the transcription reactions at 37°C for 2 hours, the reactions were treated with DNase I (Promega) for an additional 30 minutes. The RNA products were electrophoresed on denaturing 5% or 10% polyacrylamide gels in 1X TBE buffer (100 mM Tris-borate, pH 8.3, 1 mM EDTA) and 7 M urea. Following electrophoresis, UV shadowing was used to detect the RNA bands, which were subsequently excised, eluted overnight at 4°C in elution buffer (500 mM ammonium acetate, 0.1% SDS, 20 mM EDTA), precipitated with ethanol, and resuspended in H_2O. Debris and residual urea were removed by passage through Sephadex G-50 columns (Amersham Pharmacia Biotech) and the RNAs were precipitated with ethanol, resuspended in H_2O, quantified by spectrophotometry at 260 nm, and stored at -20°C. CENTR molecules were generated using two RNA strands folded together by heat denaturation at 94 °C and slow renaturation at room temperature in 40 mM Tris-HCl, pH 8.0/10 mM $MgCl_2$.

Immunoprecipitation of HDV-derived RNAs.

The Protein-G Immunoprecipitation Kit (Sigma-Aldrich) was employed to co-immunoprecipitate HDV RNAs in accordance with the manufacturer's protocol. Briefly, 5 pmol of RNA was 5'-end radiolabeled with [γ-^{32}P]ATP as previously described (Pelchat et al., 2002) and incubated with 26 µg of HeLaScribe Nuclear Extract (Promega) in RIP buffer (50 mM Tris–Cl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl; Niranjanakumari et al., 2002) at 37°C for 30 min. A 20-fold excess of poly(A) RNA or P11.60 was added to some reactions as non-specific

competitors. A 20-fold excess of cold homologous RNA was used as a specific competitor. Approximately 5 μg of either mouse monoclonal antibody (8WG16; Upstate) directed against the Rpb1 C-terminal domain (CTD) of RNAP II or goat α-mouse IgG (Sigma-Aldrich) was then added to the reaction mixtures and they were incubated for 2 hours at 4°C on a rotator. Pre-washed protein-G beads were then added to the mixtures and incubated overnight as above. Following incubation, reactions were centrifuged, washed 5 times in RIP buffer, once with 0.1X RIP buffer, and eluted according to manufacturer's protocols in acrylamide loading buffer (Sambrook et al., 1989). Samples were electrophoresed on denaturing 10% polyacrylamide gels and the bands were detected by autoradiography or phosphorimager scanning (Molecular Dynamics).

Ribonucleoprotein immunoprecipitation (RIP) assay.

The RIP assay was performed as previously described (Niranjanakumari et al., 2002) with some modification. HeLa cells were transfected with 6 μg of pcDNA3-AgS and 10 μg of dimeric HDV RNA genome of positive polarity using Lipofectamine TM 2000 Reagent (Invitrogen) in accordance with manufacturer's protocol. pcDNA3-AgS is a derivative of pcDNA3 (Stratagene) encoding the small *delta* antigen under the control of a CMV promoter (Lazinski and Taylor, 1993). Formaldehyde (1%) was used to cross-link nucleic acids to cellular proteins for 5 min at 37°C. About 10^6 cells were harvested by centrifugation, washed twice with ice-cold PBS, the supernatant decanted, and pellets stored at -80°C. Cells were resuspended in 2 ml RIP buffer, homogenized by sonication, and centrifuged at 16,000 g for 30 min at 4°C. Immunoprecipitation was performed using the Protein-G Immunoprecipitation Kit as above using 500 μl of cell lysate. Samples were washed 5 times

with high-stringency RIP wash buffer containing 1 M urea (Niranjanakumari et al., 2002), heated to 70°C for 45 minutes in a solution of 50 mM Tris-HCl, pH 7.0, 5 mM EDTA, 10 mM dithiotriol (DTT), and 1% SDS to reverse the cross-links, and eluted by centrifugation. The nucleic acids were extracted by phenol/chloroform, ethanol precipitated, resuspended in H_2O, and treated with DNAse I (Promega) for 45 minutes at 37°C. After a second RNA extraction with phenol/chloroform and precipitation with ethanol, the RNA was resuspended in H_2O. The RNA was heated for 5 min at 65°C with 1 pmol of one of the primers used to generate the R199G cDNA (R199GFW: 5'-GGAATTCTAATACGACTCACTATAGGG^{1541}ACTGCTCGAGGATCTCTTCTCTCCC^{156} ^{5}-3' or R199GREV: 5'-GGAATTC^{60}ACATCCCCTCTCGGGTGC^{43}-3'; underlined nucleotides indicate the T7 promoter) in 5µl of annealing buffer (50mM Tris-HCl, pH 8.3, 10mM $MgCl_2$, 80mM KCl solution), and then the mixture quickly snap frozen by immersion in −80°C ethanol/dry ice for 2 min. Synthesis was performed by adding 4 units avian myeloblastosis virus reverse transcriptase (Amersham Pharmacia Biotech) in 5µl of ice cold annealing buffer containing 8mM DTT and 4mM dNTP to the primer annealed templates, and then incubating at 42°C for 30 min. The cDNA was extracted by phenol/chloroform, ethanol precipitated, and resuspended in H_2O. PCR amplification using the same primers was composed of 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, and extension at 76°C for 1 min. The primer B-ActinREV (5'-GGGGTACTTCAGGGTGAGGATGCCTCTCTT-3') for the RT reaction and both B-ActinREV and B-ActinFW (5'- ATGGATGATGATATCGCCGCGCTCGTCGTC-3') were used to perform RT-PCR to generate a small 200 nt cDNA fragment corresponding to the 5'-end of the β-actin mRNA. Bands were resolved by 2% agarose gel electrophoresis and

visualized under UV light in the presence of ethidium bromide. Semi-quantitative RT-PCR was performed as above by ending the PCR amplification during the linear phase of amplification. Amplification curves were carried out by real-time RT-PCR of HDV RNA performed with an ABI Prism 7300 DNA analyzer (Applied Biosystems) in the presence of SYBR Green (Invitrogen).

Computer analysis of putative HDV promoter sequences.

All the HDV sequences were taken from the Subviral RNA Database (http://subviral.med.uottawa.ca/; Rocheleau and Pelchat, 2006). Multiple alignments of HDV sequences were obtained using ClustalW (Thompson et al., 1994). The alignments were corrected manually to maximize sequence identities and secondary structural features. Covariation analyses were performed using BioEdit v7.0.5 (Brown, 1991). Thermodynamic parameters of RNA promoters were established for each secondary structure using the RNAeval Software from the Vienna RNA Structure Software Suite (Hofacker, 2003). Logos were generated using WebLogo (Crooks et al., 2004).

2.7 Acknowledgments

Technical assistance in cell maintenance and transfection was provided by E.G. Brown and M. Gallant. V.G.-S. is supported by a Canadian Graduate Scholarship awarded by the Natural Science and Engineering Research Council of Canada. A.A. is supported by an award from the Libyan government. This work was funded by a grant from The Canadian Institutes of Health Research awarded to M.P.

2.8 References

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Brown, J.W., 1991. Phylogenetic comparative analysis on Macintosh computers. Comput. Appl. Biosci. 7, 391-393.

Casey, J.L., Bergmann, K.F., Brown, T.L., Gerin, J.L., 1992. Structural requirements for RNA editing in hepatitis *delta* virus: evidence for a uridine-to-cytidine editing mechanism. Proc. Natl. Acad. Sci. 89, 7149-7153.

Chang, F.L., Chen, P.J., Tu, S.J., Wang, C.J., Chen, D.S., 1991. The large form of hepatitis *delta* antigen is crucial for the assembly of hepatitis *delta* virus. Proc. Natl. Acad. Sci. 88, 8490-8494.

Chen, P.J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., Taylor, J., 1986. Structure and replication of the genome of hepatitis *delta* virus. Proc. Natl. Acad. Sci. 83, 8774-8778.

Cramer, P., 2002. Multisubunit RNA polymerases. Curr. Opin. Struct. Biol. 12, 89-97.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188-1190.

de Mercoyrol, L., Job, C., Job, D., 1989. Studies on the inhibition by alpha-amanitin of single-step addition reactions and productive RNA synthesis catalyzed by wheat-germ RNA polymerase II. Biochem. J. 258, 165-169.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Filipovska, J., Konarska, M., 2000. Specific HDV RNA-templated transcription by pol II *in vitro*. RNA 6, 41-54.

Fu, T., Taylor, J., 1993. The RNAs of hepatitis *delta* virus are copied by RNA polymerase II in nuclear homogenates. J. Virol. 67, 6965-6972.

Garcia-Blanco, M.A., Jamison, S.F., Sharp, P.A., 1989. Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. Genes Dev. 3, 1874-1886.

Gong, X.Q., Nedialkov, Y.A., Burton, Z.F., 2004. Alpha-amanitin blocks translocation by human RNA polymerase II. J. Biol. Chem. 279, 27422-27427.

Goodman, T.C., Nagel, L., Rappold, W., Klotz, G., Riesner, D., 1984. Viroid replication: equilibrium association constant and comparative activity measurements for the viroid polymerase interaction. Nucleic Acids Res. 12, 6231-6246.

Greco-Stewart, V.S., St-Laurent Thibault, C., Pelchat, M., 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology 356, 35-44.

Gudima, S., Dingle, K., Wu, T.T., Moraleda, G., Taylor, J., 1999. Characterization of the 5' ends for polyadenylated RNAs synthesized during the replication of hepatitis *delta* virus. J. Virol. 73, 6533-6539.

Gudima, S., Wu, S.Y., Chiang, C.M., Moraleda, G., Taylor, J., 2000. Origin of hepatitis *delta* virus mRNA. J. Virol. 74, 7204-7210.

Gudima, S.O., Chang, J., Taylor, J.M., 2006. Restoration *in vivo* of defective hepatitis *delta* virus RNA genomes. RNA. 12, 1061-1073.

Hahn, S., 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat. Struct. Mol. Biol. 11, 394-403.

Hofacker, I.L., 2003. Vienna RNA secondary structure server. Nucleic Acids Res. 31, 3429-3431.

Kolonko, N., Bannach, O., Aschermann, K., Hu, K.H., Moors, M., Schmitz, M., Steger, G., Riesner, D., 2006. Transcription of potato spindle tuber viroid by RNA polymerase II starts in the left terminal loop. Virology 347, 392-404.

Kuo, M.Y., Goldberg, J., Coates, L., Mason, W., Gerin, J., Taylor, J.M., 1988a. Molecular cloning of hepatitis *delta* virus RNA from an infected woodchuck liver: sequence, structure, and applications. J. Virol. 62, 1855-1861.

Kuo, M.Y., Sharmeen, L., Dinter-Gottlieb, G., Taylor, J., 1988b. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis *delta* virus. J. Virol. 62, 4439-4444.

Kuo, M.Y., Chao, M., Taylor, J., 1989. Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of the *delta* antigen. J. Virol. 63, 1945-1950.

Lai, M.M., 2005. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis *delta* virus. J. Virol. 79, 7951-7958.

Lazinski, D.W., Taylor, J.M., 1993. Relating structure to function in the hepatitis *delta* virus antigen. J. Virol. 67, 2672-2680.

Li, Y.J., Macnaughton, T., Gao, L., Lai, M.M., 2006. RNA-templated replication of hepatitis *delta* virus: genomic and antigenomic RNAs associate with different nuclear bodies. J. Virol. 80, 6478-6486.

Macnaughton, T.B., Gowans, E.J., MacNamara, S.P., Burrell, C.J., 1991. Hepatitis *delta* antigen is necessary for access for hepatitis *delta* virus RNA to the cell transcriptional machinery but is not part of the transcriptional complex. Virology 184, 387-390.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M.M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

Mathews, D.B., Durbin, R.D., 1994. Mechanistic aspects of tagetitoxin inhibition of RNA polymerase from *Escherichia coli*. Biochemistry 33, 11987-11992.

Meienhofer, J., Atherton, E., 1973. Structure-activity relationships in the actinomycins. Adv. Appl. Microbiol. 16, 203-300.

Modahl, L.E., Macnaughton, T.B., Zhu, N., Johnson, D.L., Lai, M.M., 2000. RNA-dependent replication and transcription of hepatitis *delta* virus RNA involve distinct cellular RNA polymerases. Mol. Cell. Biol. 20, 6030-6039.

Navarro, J.A., Flores, R., 2000. Characterization of the initiation sites of both polarity strands of a viroid RNA reveals a motif conserved in sequence and structure. EMBO J. 19, 2662-2670.

Nie, X., Chang, J., Taylor, J.M., 2004. Alternative processing of hepatitis *delta* virus antigenomic RNA transcripts. J. Virol. 78, 4517-4524.

Niranjanakumari, S., Lasda, E., Brazas, E., Garcia-Blanco, M.A., 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods 26, 182-190.

Patturajan, M., Schulte, R.J., Sefton, B.M., Berezney, R., Vincent, M., Bensaude, O., Warren, S.L., Corden, J.L., 1998. Growth-related changes in phosphorylation of yeast RNA polymerase II. J. Biol. Chem. 273, 4689-4694.

Pelchat, M., Côté, F., Perreault, J.P., 2001. Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol. 146, 1753-1763.

Pelchat, M., Grenier, C., Perreault, J.P., 2002. Characterization of a viroid-derived RNA promoter for the DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 41, 6561-6571.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636-642.

Rocheleau, L., Pelchat, M., 2006. The Subviral RNA Database: a toolbox for viroids, the hepatitis delta virus and satellite RNAs research. BMC Microbiol. 6, 24.

Ryu, W.S., Bayer, M., Taylor J., 1992. Assembly of hepatitis *delta* virus particles. J. Virol. 66, 2310-2315.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Press, New York.

Sims, R.J. 3rd, Mandal, S.S., Reinberg, D., 2004. Recent highlights of RNA-polymerase-II-mediated transcription. Curr. Opin. Cell. Biol. 16, 263-271.

Smale, S.T., Kadonaga, J.T., 2003. The RNA polymerase II core promoter. Annu. Rev. Biochem. 72, 449-479.

Swanson, M.S., Dreyfuss, G., 1988. RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. EMBO J. 7, 3519-3529.

Taylor, J.M., 2006. Hepatitis *delta* virus. Virology. 344, 71-76.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Thompson, N.E., Steinberg, T.H., Aronson, D.B., Burgess, R.R., 1989. Inhibition of *in vivo* and *in vitro* transcription by monoclonal antibodies prepared against wheat germ RNA

polymerase II that react with the heptapeptide repeat of eukaryotic RNA polymerase II. J. Biol. Chem. 264, 11511-11520.

Wang, H.W., Wu, H.L., Chen, D.S., Chen, P.J., 1997. Identification of the functional regions required for hepatitis D virus replication and transcription by linker-scanning mutagenesis of viral genome. Virology 239, 119-131.

Wang, J.G., Cullen, J., Lemon, S.M., 1992. Immunoblot analysis demonstrates that the large and small forms of hepatitis *delta* virus antigen have different C-terminal amino acid sequences. J. Gen. Virol. 73, 183-188.

Weiner, A.J., Choo, Q.L., Wang, K.S., Govindarajan, S., Redeker, A.G., Gerin, J.L., Houghton, M. 1988. A single antigenomic open reading frame of the hepatitis *delta* virus encodes the epitope(s) of both hepatitis *delta* antigen polypeptides p24 *delta* and p27 *delta*. J. Virol. 62, 594-599.

Wu, T.T., Netter, H.J., Lazinski, D.W., Taylor, J.M., 1997. Effects of nucleotide changes on the ability of hepatitis *delta* virus to transcribe, process, and accumulate unit-length, circular RNA. J. Virol. 71, 5408-5414.

Xia, Y.P., Chang, M.F., Wei, D., Govindarajan, S., Lai, M.M., 1990. Heterogeneity of hepatitis *delta* antigen. Virology 178, 331-336.

Yamaguchi, Y., Filipovska, J., Yano, K., Furuya, A., Inukai, N., Narita, T., Wada, T., Sugimoto, S., Konarska, M.M., Handa, H., 2001. Stimulation of RNA polymerase II elongation by hepatitis *delta* antigen. Science 293, 124-217.

CHAPTER 3:

THE HEPATITIS *DELTA* VIRUS RNA GENOME INTERACTS WITH THE HUMAN RNA POLYMERASES I AND III

Valerie S. Greco-Stewart, Erica Schissel, and Martin Pelchat

3.1 Statement of contribution

The manuscript entitled "The hepatitis *delta* virus RNA genome interacts with the human RNA polymerases I and III" was published in Virology in 2009 (386, 12-15). I performed all laboratory experiments with the exception of two co-immunoprecipitation experiments (R38G and L48AG, Figure 3.2B) which were performed by E. Schissel. I wrote this manuscript and editing, revisions, and additional insights were provided by M. Pelchat.

3.2 Abstract

The hepatitis *delta* virus (HDV) relies on human transcriptional machinery for its replication and transcription. Although the involvement of RNA polymerase II in HDV RNA biosynthesis is established, the contribution of additional polymerases remains uncertain. Here, we demonstrate the interaction of both RNA polymerase I and III with HDV RNA, both *in vitro* and in human cells. Binding of these polymerases occurs near the terminal stem-loop domains of both polarities of the HDV RNA genome. Based on interactions of HDV RNA with numerous host polymerases, our results suggest a higher level of complexity of HDV biology than previously envisioned.

Keywords: Hepatitis *delta* virus; RNA polymerase; replication; transcription; TATA-binding protein; RNA-binding proteins.

3.3 Introduction

The hepatitis *delta* virus (HDV) is composed of a single-stranded, circular RNA molecule of approximately 1,700 nucleotides that folds into a rod-like structure and contains a single open reading frame (Taylor, 2006). Because HDV does not encode an RNA-dependent RNA polymerase, it relies on host polymerase(s) for its replication and transcription. It is acknowledged that RNA polymerase II (RNAP II) is involved in HDV replication based on the sensitivity of the process to α-amanitin and on the post-transcriptional processing of HDV mRNA (i.e. HDAg mRNA) with a 5' methylguanine cap and a poly(A) tail (Fu and Taylor, 1993; Gudima et al., 1999; Lai, 2005; Macnaughton et al., 1991; Taylor, 2006). Recently, we confirmed the interaction of RNAP II with HDV RNA both *in vitro* and *in vivo* and demonstrated the binding of this polymerase to the terminal stem-loop domains of both polarities of the HDV genome (Greco-Stewart et al., 2007). The involvement of RNAP I in HDV biology has been suggested based on the resistance of antigenomic RNA production to higher doses of α-amanitin, on the co-localization of RNAP I with antigenomic HDV RNA, on the interaction of HDAg-S with components of the RNAP I transcription factor SL1, and on the inhibition of antigenomic RNA synthesis *in vitro* by α-SL1 antibody (Li et al., 2006; Macnaughton et al., 2002; Modahl et al., 2000). However, no direct interaction between RNAP I and HDV RNA has been reported, and a role for RNAP III in HDV replication has never been investigated. In this study, we report the association of both RNAP I and RNAP III components with HDV RNA, both *in vitro* and *in vivo*. Our results indicate the

recruitment of the two RNAPs by both polarities of HDV RNA and the possible involvement of the TATA-binding protein (TBP) in these processes. Our data suggest a mechanism of HDV replication involving all three human RNAPs.

3.4 Results

HDV RNA interacts with components of RNAP I and RNAP III in cells replicating HDV. Because RNAPs must bind their templates to transcribe, we investigated whether components of RNAP I and/or RNAP III associate with the HDV RNA genome in cells replicating HDV RNA. To establish these interactions, we took advantage of a ribonucleoprotein immunoprecipitation assay (RIPA), which has been successfully used to indicate HDAg, PSF, and RNAP II interaction with HDV RNA (Niranjanakumari et al., 2002; Greco-Stewart et al., 2006; 2007). HeLa cells were co-transfected with a dimeric HDV RNA genome and a plasmid expressing HDAg-S. Four days post-transfection, the ribonucleoproteins were cross-linked with formaldehyde and immunoprecipitated using antibodies against either TAF I p95/110 (TAF1) or RNAP III polypeptide K (POLR3K). TAF1 is a component of the transcription factor SL1 which directs the assembly of initiation complexes at RNAP I promoters. POLR3K is one of the core subunits of RNAP III. Following immunoprecipitation, the cross-links were reversed by heating the samples and the recovered nucleic acids were subjected to RT-PCR using primers to generate a 232 bp cDNA fragment corresponding to the R199 region (see Figure 3.2B; Greco-Stewart et al., 2007).

HDV RNA associated with both TAF1 and POLR3K in cross-linked HeLa cells. The interactions were specific since HDV RNA failed to co-immunoprecipitate in the absence of formaldehyde treatment, without antibody, or in the presence of a non-related antibody (i.e.

α-IgG; Figure 3.1). In addition, the primers used in the RT reactions allowed differentiation between the polarities of HDV RNA, as demonstrated when *in vitro* transcribed HDV RNAs were subjected to RT-PCR (Figure 3.1, HDVG RNA and HDVAG RNA lanes; Greco-Stewart et al., 2007). Since both polarities of HDV RNA were recovered when the cross-linked samples were immunoprecipitated with these antibodies, we conclude that both TAF1 and POLR3K associate with both polarities of HDV RNA within HeLa cells.

RNAP I and RNAP III components bind close to the terminal stem-loop domains of HDV.

To determine the regions of the HDV genome to which subunits of RNAP I and RNAP III bind, various portions of the HDV genome were transcribed *in vitro*, radiolabeled, and subjected to co-immunoprecipitation using α-TAF1 and α-POLR3K antibodies (Figure 3.2). Figure 3.2A shows a representative experiment using a segment of the right terminal stem-loop domain of genomic polarity (i.e. R199G). R199G associated with TAF1 in the presence of HeLa NE proteins. In the absence of antibody or in the presence of a non-related antibody (i.e. α-IgG), R199G failed to associate with TAF1. TAF1 binding was not hindered by the addition of a 50-fold excess of either P11.60 or poly(A) RNA, although a 50-fold excess of unlabeled homologous competitor prevented association. P11.60 is a small RNA species that folds into a hairpin (Pelchat and Perreault, 2004), thus providing an excellent competitive RNA to test for non-specific affinity of the proteins to double-stranded RNA or stem-loop secondary structures.

Co-immunoprecipitations of the terminal stem-loop domains (i.e. R199G, R199AG, L213G, and L213AG) and the central regions (i.e. CENTRG and CENTRAG) of both polarities of HDV RNA were performed using both antibodies. We observed interaction of

Figure 3.1: Interaction of RNAP I and RNAP III components with HDV RNA in HeLa cells. HeLa cells replicating HDV RNA were used for the preparation of cross-linked lysates for co-immunoprecipitation using the either α-TAF1 (A), α-POLR3K (B), α-IgG, or no antibody. Experiments using non-cross-linked lysates are also presented. G and AG lanes contain RT reactions performed using primers complementary to either genomic or antigenomic HDV RNA corresponding to the R199 region. HDVG and HDVAG RNA lanes are controls using *in vitro* transcribed HDV RNA of genomic and antigenomic polarity, respectively. Lanes M contain GelPilot Mid Range DNA ladder (Qiagen).

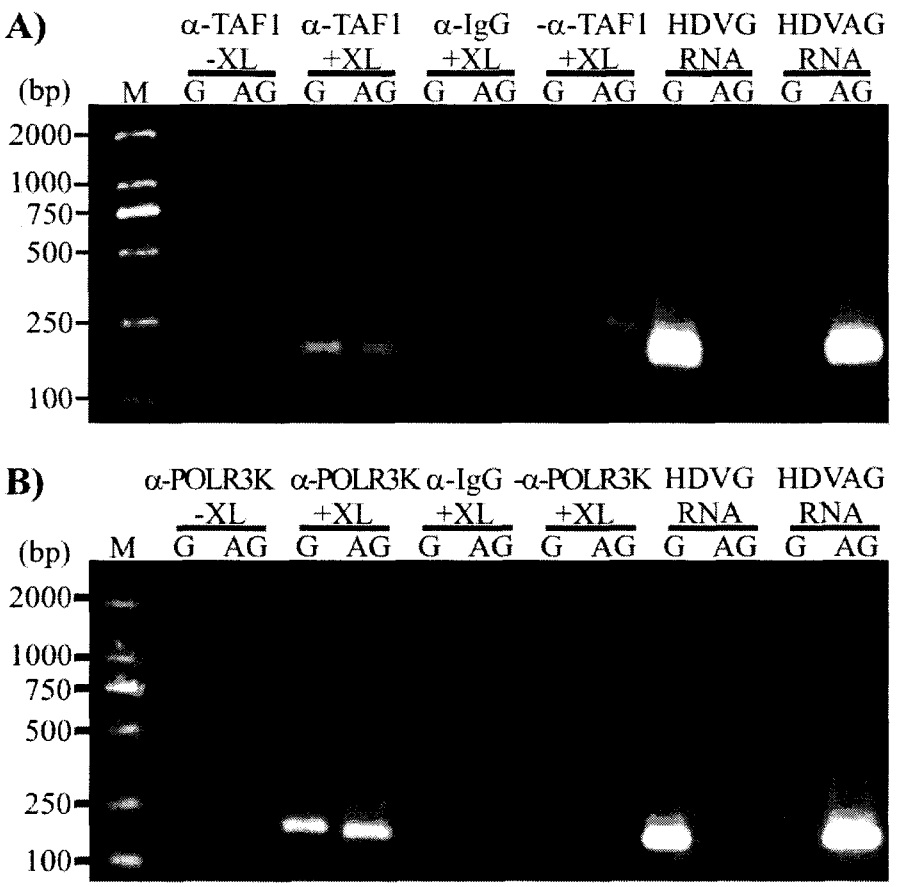
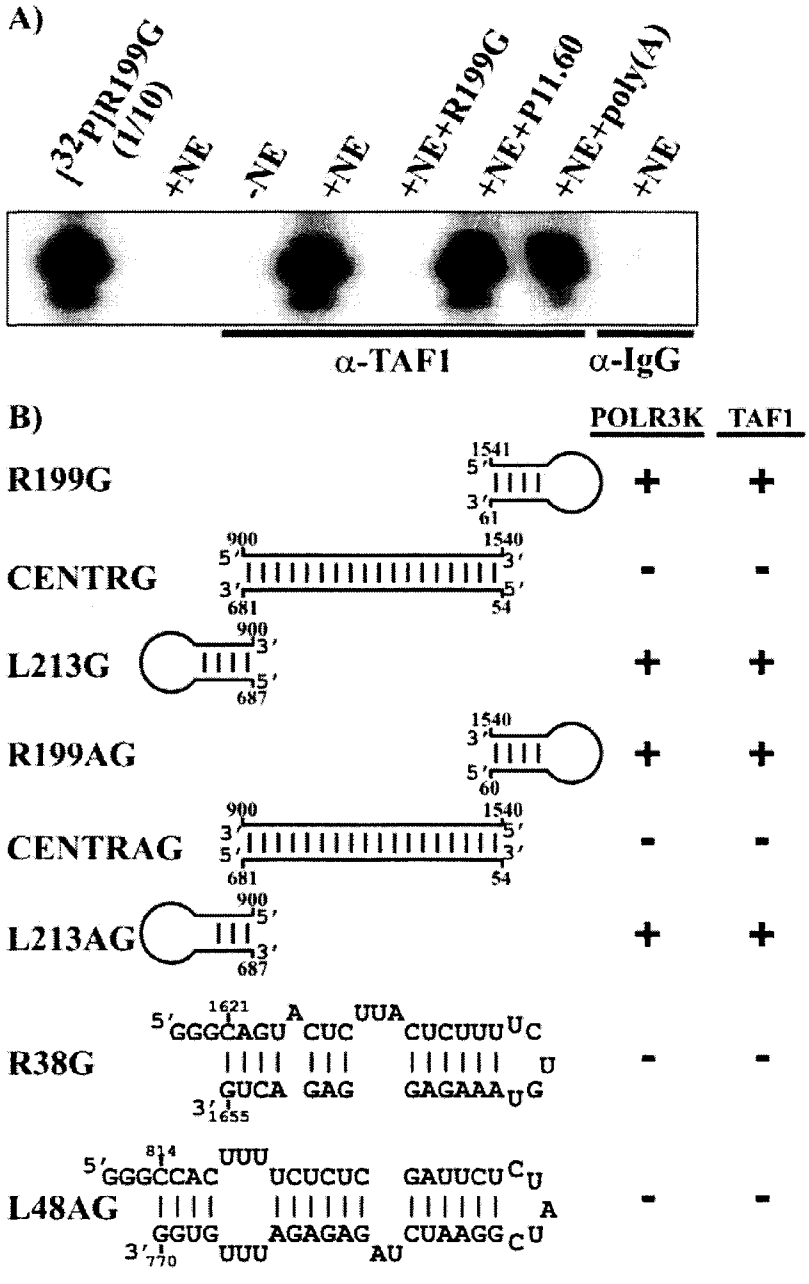

A)
α-TAF1 -XL
α-TAF1 +XL
α-IgG +XL
-α-TAF1 +XL
HDVG RNA
HDVAG RNA
(bp)
M
G AG G AG G AG G AG G AG G AG
2000
1000
750
500
250
100
B)
α-POLR3K -XL
α-POLR3K +XL
α-IgG +XL
-α-POLR3K +XL
HDVG RNA
HDVAG RNA
(bp)
M
G AG G AG G AG G AG G AG G AG
2000
1000
750
500
250
100

Figure 3.2: Association of TAF1 and POLR3K with HDV-derived RNAs. A) Co-immunoprecipitations of radiolabeled R199G with TAF1 in HeLa NE. 50-fold excess non-radioactive R199G, poly(A) RNA, or P11.60 (Pelchat and Perreault, 2004) were used as competitors. B) Compilation of the interaction of various HDV-derived RNAs with both TAF1 and POLR3K.


A)
[32P]R199G (1/10)
+NE
-NE
+NE
+NE+R199G
+NE+P11.60
+NE+poly(A)
+NE
α-TAF1
α-IgG
B)
POLR3K
TAF1
R199G
CENTRG
L213G
R199AG
CENTRAG
L213AG
R38G
L48AG

TAF1 and POLR3K with fragments containing the terminal stem-loop domains of the HDV genome of both polarities (Figure 3.2B) in a manner similar to what was observed for RNAP II interaction (Greco-Stewart et al., 2007). These data allow for the possibility that not only RNAP II, but also RNAP I and RNAP III, might be involved in HDV biology through specific association with the extremities of both polarities of the HDV RNA genome.

To further refine the regions interacting with both RNAP I and RNAP III and to determine if the RNA features bound by these RNAPs are identical to those used by RNAP II, two small HDV-derived transcripts reported to be able to associate with RNAP II were synthesized and were used in the co-immunoprecipitation reactions (Greco-Stewart et al., 2007). Specifically, R38G and L48AG, which correspond to 38 and 48 nt RNAs located at the tip of the right genomic and the left antigenomic hairpin domains, respectively, were tested. These two small RNA molecules were not able to co-immunoprecipitate using either α-TAF1 or α-POLR3K antibodies (Figure 3.2B), suggesting either that those proteins do not bind to the tip of the rod-like structure of HDV RNA or that additional features present on the larger RNA species are required.

RNAP I and RNAP III binding to HDV RNA are independent and involve the TATA-binding protein.

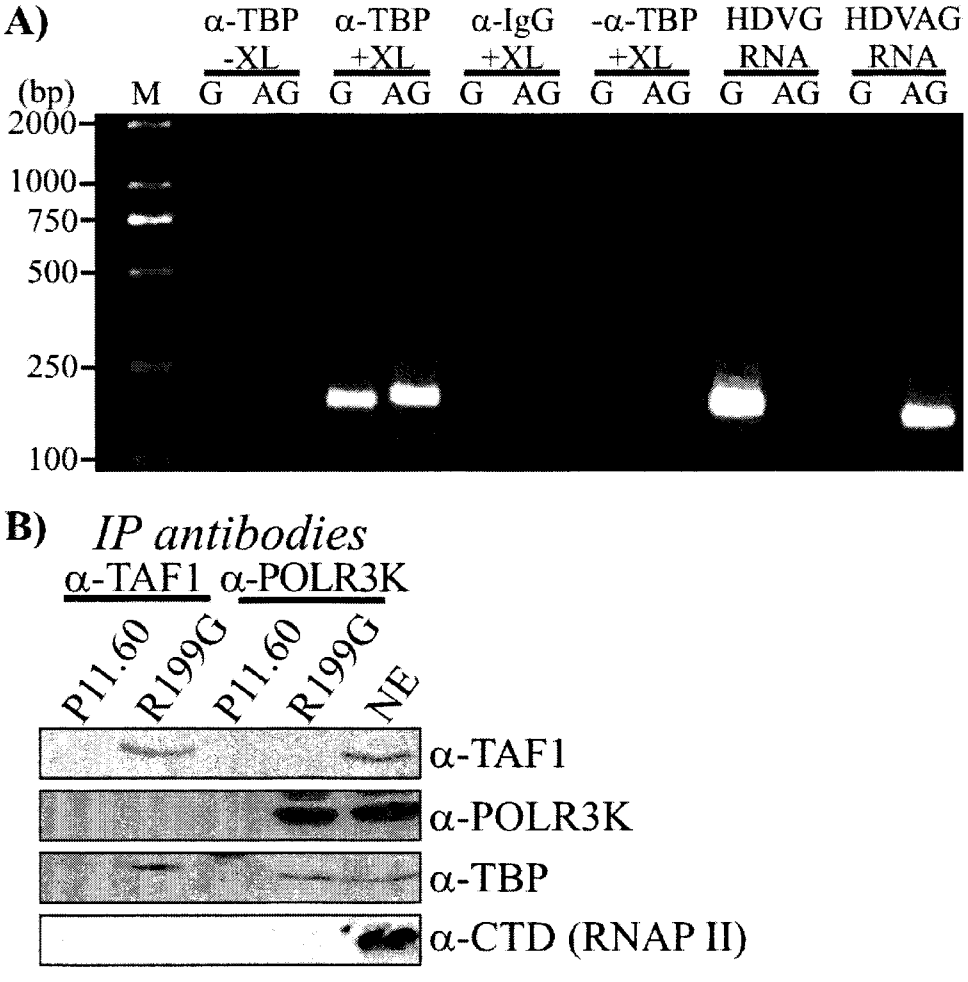
The binding of the three RNAPs to the same regions of HDV RNA suggests a common mechanism of RNA recognition which might involve a shared subunit. Since TBP is common to three RNAPs during DNA promoter recognition (Goodrich and Tjian, 1994) and this protein was suggested to be involved in HDV RNA promoter recognition by RNAP II (Abrahem and Pelchat, 2008), we speculated that TBP might be present during the binding of

HDV RNA to both RNAP I and RNAP III. RIPA was initially performed to determine the association of TBP with HDV RNA in HeLa cells. As shown in Figure 3.3A, TBP associated with both polarities of HDV RNA *in vivo*. HDV RNA binding to TBP was specific since the RIPA was only successful in cross-linked HeLa cells and failed in the absence of α-TBP antibody or in the presence of a non-related antibody. However, this assay does not determine the polymerase involved.

To verify the presence of TBP in both RNAP I and RNAP III complexes independently of one another and of RNAP II, we performed RNA affinity chromatography followed by immunoprecipitation and Western analysis. For this purpose, we selected R199G, which binds to all three RNAPs (Figure 3.2B and Greco-Stewart et al., 2007), and added an RNA tag with an affinity for streptavidin (Srisawat and Engelke, 2002). We used a tagged P11.60 as a negative control since P11.60 does not bind human RNAPs (this work and Greco-Stewart et al., 2007). The tagged RNAs were immobilized on streptavidin-agarose beads and incubated with HeLa NE. Following washes, the RNP complexes were eluted by competition with molecular excess of *d*-biotin, immunoprecipitated with either α-TAF1 or α-POLR3K antibody, and the presence of the tagged RNAs was confirmed by PAGE (data not shown). The samples were then subjected to SDS-PAGE and Western blotting was performed to determine the composition of these complexes using specific antibodies.

TAF1 but not POLR3K was detected in complexes immunoprecipitated using the α-TAF1 antibody (Figure 3.3B). Conversely, complexes immunoprecipitated using the α-POLR3K antibody contained POLR3K but not TAF1. However, TBP was common to both complexes. Neither protein was detected when R199G-S1 was substituted with P11.60-S1, indicating the specificity of these protein-RNA interactions (Figure 3.3B). To further

Figure 3.3: Presence of TBP in both RNAP I and RNAP III complexes on HDV RNA. A) Interaction of TBP with HDV RNA within HeLa cells, as described in Fig. 1 using α-TBP. B) Mutual exclusivity of RNAP complexes and the involvement of TBP. S1-R199G and S1-P11.60 RNA were immobilized on streptavidin-agarose beads and incubated with HeLa NE. Following washes, the RNP complexes were eluted by competition with a molecular excess of *d*-biotin, and immunoprecipitated with either α-TAF1 or α-POLR3K antibodies. The proteins were detected using specific antibodies.


A)
α-TBP -XL
α-TBP +XL
α-IgG +XL
-α-TBP +XL
HDVG RNA
HDVAG RNA
(bp)
M
G AG G AG G AG G AG G AG G AG
2000
1000
750
500
250
100
B)
IP antibodies
α-TAF1
α-POLR3K
P11.60
R199G
P11.60
R199G
NE
α-TAF1
α-POLR3K
α-TBP
α-CTD (RNAP II)

demonstrate that RNAP II was not involved in the formation of these complexes, the TAF1 and POLR3K complexes were each probed with an antibody against RNAP II carboxy-terminal domain (Figure 3.3B). RNAP II was not present in either of these complexes, indicating that HDV RNA binding to both TAF1 and POLR3K is independent of RNAP II binding. These results demonstrate the binding of TAF1 and POLR3K complexes to HDV RNA, either independently or through a common subunit, and suggest the involvement of TBP in both RNAP I and RNAP III binding.

3.5 Discussion

Using antibodies directed against components of both RNAP I and RNAP III, we have further implicated RNAP I in HDV replication and strongly suggested a role for RNAP III in this process. Combined with the established role of RNAP II, the association of both RNAP I and RNAP III with HDV RNA suggests the involvement of the three human RNAPs in HDV biology. Interestingly, the three RNAPs bind near or at the terminal stem-loop domains of HDV (this work and Greco-Stewart et al., 2007). TBP is common to the three RNAPs; it is found in the SL1 complex of RNAP I, the TFIID complex of RNAP II, the TFIIIB complex of RNAP III, and the $SNAP_c$ complex used in transcription of small nuclear RNAs (snRNAs) by both RNAPs II and III (Goodrich and Tjian, 1994). TBP can recognize both TATA and TATA-less promoters dependent on accessory transcription factors and its affinity for DNA templates is mediated by electrostatic interactions with the phosphate backbone (Bewley et al., 1998), which could account for its ability to bind HDV RNA despite the absence of known DNA promoter features. Because TBP was found in the RNAP complexes, it is tempting to suggest that this shared polymerase subunit might mediate binding of all three RNAPs to an HDV template. In addition, we recently reported that TBP binding occurs to

the tip of the rod-like structure of HDV (i.e. binding sites are within R38G and L48AG; Abrahem and Pelchat, 2008). Because additional features present on the larger RNA species might be required for both RNAP I and RNAP III association, it is likely that accessory host factors might bind to these features and provide a link between the RNAPs and the HDV-derived RNA promoters at the initiation phase of transcription.

RNAP I was reported to be implicated in the synthesis of antigenomic, but not genomic, HDV RNA (Li et al., 2006; Macnaughton et al., 2002; Modahl et al., 2000). In addition, a role for RNAP III in HDV replication has never been reported. It is therefore possible that some of the associations observed in this study might not lead to transcription initiation. Because there is currently no suitable assay to directly study the transcription of RNAPs on HDV RNA templates, the roles of these RNAPs in the HDV life cycle will be difficult to address. The development of a sensitive and reproducible transcription system using HDV RNAs will be crucial to enable straightforward testing of the relevance of our findings. Nevertheless, the three human RNAPs are capable of independent binding to the terminal stem-loop domains of HDV and each RNAP complex contains TBP. Taken together, this study provides a foundation to further investigate HDV biology by suggesting a role for the three human RNAPs where each polymerase complex might have a specific function during HDV replication and transcription.

3.6 Materials and methods

RNA synthesis.

HDV-derived RNAs and P11.60 were synthesized by run-off transcription using T7 RNA polymerase, as previously described (Greco-Stewart et al., 2007; Pelchat and Perreault,

2004). When needed, S1 aptamer was encoded in one of the primers to generate a streptavidin tag at the 3' end of the RNAs (Srisawat and Engelke, 2002).

Ribonucleoprotein immunoprecipitation assay (RIPA).

HeLa cells were transfected with 6 μg of pCDNA3-AgS and 10 μg of dimeric HDV genomic RNA (Greco-Stewart et al., 2007). ~10^6 cells were treated with 1% formaldehyde, disrupted by sonication, and co-immunoprecipitated with the Protein-G Immunoprecipitation Kit (Sigma-Aldrich) using α-TAF1 (Santa Cruz), α-POLR3K (U.S. Biologicals), α-TBP (Upstate), or mouse α-IgG (Sigma-Aldrich) antibodies in RIPA buffer (50 mM Tris-HCl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl) as previously described (Greco-Stewart et al., 2007; Niranjanakumari et al., 2002). Samples were then eluted, treated with DNAse I, phenol/chloroform extracted, ethanol precipitated, and reverse transcribed using strand-specific primers (Greco-Stewart et al., 2007). Samples were amplified by 35 cycles of PCR using the same primers. Products were resolved on 2% agarose gels and visualized by SYBR green staining. These assays were performed at least twice.

Co-Immunoprecipitation analysis.

The protein-G Immunoprecipitation Kit was used with ~50 μg of HeLa Nuclear Extract (NE; Promega), 5 μg of antibody, and ~1 pmol of 5'-end radiolabeled RNA, as previously described (Greco-Stewart et al., 2007). Competitors included 50-fold molar excess of poly(A) RNA, P11.60, or unlabeled homologous RNA. Samples were incubated for 16 h at 4°C with agitation, washed five times with RIPA buffer, and eluted with acrylamide loading

dye (Sambrook et al., 1989). RNAs were separated on denaturing polyacrylamide gels (PAGE) and results visualized by phosphorimager scanning (Molecular Dynamics). These assays were performed at least twice.

Affinity purification of ribonucleoprotein complexes.

30 µl of pre-washed streptavidin agarose beads (Invitrogen) were incubated with ~200 µg of HeLa NE and 1 nmol of RNA (either R199G-S1 or P11.60-S1) in 50 mM HEPES (pH 7.4), 10 mM $MgCl_2$, and 100 mM NaCl for 1 h at 4°C, followed by a 3 hour wash in the same buffer. Samples were eluted by the addition of 5 mM *d*-biotin (Sigma-Aldrich) and then subjected to co-immunoprecipitation as above. SDS loading dye was added to elute the samples which were subjected to SDS-PAGE and Western blotting using the One Hour Complete Western Kit (Genscript) using the antibodies described above, or one directed against RNAP II carboxy-terminal domain (Upstate).

3.7 Acknowledgements

V.S.G.-S. was supported by a Graduate Fellowship Award from Canadian Blood Services. This work was funded by a Canadian Institutes of Health Research grant to M.P.

3.8 References

Abrahem, A., Pelchat, M., 2008. Formation of an RNA polymerase II preinitiation complex on an RNA promoter derived from the hepatitis *delta* virus RNA genome. Nucleic Acids Res. 36, 5201-5211.

Bewley, C.A., Gronenborn, A.M., Clore, G.M., 1998. Minor groove-binding architectural proteins: structure, function, and DNA recognition. Annu. Rev. Biophys. Biomol. Struct. 27, 105-131.

Fu, T.B., Taylor, J., 1993. The RNAs of hepatitis *delta* virus are copied by RNA polymerase II in nuclear homogenates. J. Virol. 67, 6965-6972.

Goodrich, J.A., Tjian, R., 1994. TBP-TAF complexes: selectivity factors for eukaryotic transcription. Curr. Opin. Cell Biol. 6, 403-409.

Greco-Stewart, V.S., Thibault, C.S., Pelchat, M., 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology 356, 35-44.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology 357, 68-78.

Gudima, S., Dingle, K., Wu, T.T., Moraleda, G., Taylor, J., 1999. Characterization of the 5' ends for polyadenylated RNAs synthesized during the replication of hepatitis *delta* virus. J. Virol. 73, 6533-6539.

Lai, M.M., 2005. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis *delta* virus. J. Virol. 79, 7951-7958.

Li, Y.J., Macnaughton, T., Gao, L., Lai, M.M., 2006. RNA-templated replication of hepatitis *delta* virus: genomic and antigenomic RNAs associate with different nuclear bodies. J. Virol. 80, 6478-6486.

Macnaughton, T.B., Gowans, E.J., McNamara, S.P., Burrell, C.J., 1991. Hepatitis *delta* antigen is necessary for access of hepatitis *delta* virus RNA to the cell transcriptional machinery but is not part of the transcriptional complex. Virology 184, 387-390.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M.M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

Modahl, L.E., Macnaughton, T.B., Zhu, N., Johnson, D.L., Lai, M.M., 2000. RNA-Dependent replication and transcription of hepatitis *delta* virus RNA involve distinct cellular RNA polymerases. Mol. Cell. Biol. 20, 6030-6039.

Niranjanakumari, S., Lasda, E., Brazas, R., Garcia-Blanco, M.A., 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods 26, 182-190.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636-642.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Press, New York.

Srisawat, C., Engelke, D.R., 2002. RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. Methods 26, 156-161.

Taylor, J.M., 2006. Hepatitis *delta* virus. Virology 344, 71-76.

CHAPTER 4:

BINDING OF THE POLYPYRIMIDINE TRACT-BINDING PROTEIN-ASSOCIATED SPLICING FACTOR (PSF) TO THE HEPATITIS *DELTA* VIRUS RNA

Valerie S. Greco-Stewart, Catherine St-Laurent Thibault and Martin Pelchat

4.1 Statement of contribution

The manuscript entitled "Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA" was published in Virology in 2006 (356, 35-44). I contributed all experimental data presented in this publication with the exception of Figure 4.2B which was initially performed by C. St-Laurent Thibault. I performed the gel presented in the paper in order to include more comprehensive controls at the request of the manuscript reviewers. Figure 4.1 was composed by M. Pelchat. I wrote this manuscript and editing, revisions, and additional insights were provided by M. Pelchat.

4.2 Abstract

The hepatitis *delta* virus (HDV) has a very limited protein coding capacity and must rely on host proteins for its replication. A ribonucleoprotein complex was detected following UV cross-linking between HeLa nuclear proteins and an RNA corresponding to the right terminal stem-loop domain of HDV genomic RNA. Mass spectrometric analysis of the complex revealed the polypyrimidine tract-binding protein-associated splicing factor (PSF) as a novel HDV RNA-interacting protein. Co-immunoprecipitation demonstrated the interaction between HDV RNA and PSF both *in vitro* in HeLa nuclear extract and *in vivo* within HeLa

cells containing both polarities of the HDV genome. Analysis of the binding of various HDV-derived RNAs to purified, recombinant PSF further confirmed the specificity of the interaction and revealed that PSF directly binds to the terminal stem-loop domains of both polarities of HDV RNA. Our findings provide evidence of the involvement of a host mRNA processing protein in the HDV life cycle.

Keywords: Hepatitis *delta* virus; RNA binding protein; RNA replication; splicing factor; polypyrimidine tract-binding protein; mass spectrometry; UV cross-linking; co-immunoprecipitation; PSF; RNA virus.

4.3 Introduction

There is increasing evidence that cellular factors are substantially involved in viral RNA replication and transcription (Lai, 1998; Lai, 2005; Taylor, 2006). These cellular proteins are often associated with RNA-processing pathways or translation machinery, suggesting that RNA viruses might actively subvert normal cellular RNA processing or translation proteins for viral RNA synthesis. This is particularly crucial for the hepatitis *delta* virus (HDV) which has an extremely small genome and possesses a very limited protein coding capacity (Taylor, 2006). HDV must thus rely heavily on host components for most steps of its replication and transcription.

HDV has a small (~1,700 nucleotides, nt) single-stranded, circular RNA genome and requires the hepatitis B virus (HBV) envelope proteins for encapsidation and dissemination (Chen et al., 1986; Rizzetto et al., 1977; Wang et al., 1986). Its genome folds into an unbranched, rod-like structure that includes two self-cleaving motifs (*delta* motifs; Kuo et al., 1988b) and a single open reading frame (ORF) encoding two viral proteins (HDAgs;

Wang et al., 1986; Weiner et al., 1987). These two proteins are identical in sequence, except that the large HDAg (HDAg-L) contains 19 additional amino acids at its C-terminus resulting from RNA editing of the termination codon of the small HDAg (HDAg-S) gene (Luo et al., 1990; Wong and Lazinski, 2002). Although they are identical over most of their lengths, each protein has a distinct function. HDAg-S (195 amino acids) is essential for HDV replication (Kuo et al., 1989), while the HDAg-L (214 amino acids) is necessary for virion assembly and is reported to be a dominant negative inhibitor of replication (Chang et al., 1991).

In the currently accepted model, HDV replicates in mammalian cells by a symmetrical, rolling circle mechanism (Chen et al., 1986; Taylor, 2006). Replication of the infectious circular monomer (which is assigned genomic polarity by convention and accumulates at a greater intracellular abundance than the antigenomic species) produces linear, multimeric strands which are subsequently self-cleaved and ligated, yielding antigenomic polarity circular monomers. Using the latter RNAs as templates, the same three steps are repeated to generate the genomic progeny. Nuclear RNA polymerase II (RNAP II) is considered to be involved in the replication and transcription of HDV.

HDAg mRNA has a distinct 5'-end and is post-transcriptionally processed with a 5'-cap and a 3'-poly(A) tail (Gudima et al., 2000), which are typical features of transcripts generated by RNAP II. Low doses of α-amanitin, a mycotoxin that specifically inhibits DNA-dependent RNA transcription by RNAP II, inhibit the accumulation of HDV mRNA and processed unit-length genomic RNA species in infected cells (Fu and Taylor, 1993; Macnaughton et al., 1991). Additionally, in *in vitro* transcription assays using nuclear extract (NE) from HeLa cells and RNA derived from the left terminal stem-loop domain of HDV

antigenomic RNA, synthesis of the complementary strand was possible (Filipovska and Konarska, 2000). Accumulation of this RNA product was highly sensitive to α-amanitin. This sensitivity was partially abrogated in experiments conducted in NE from cells containing an α-amanitin-resistant allele of the largest subunit of RNAP II, suggesting the involvement of RNAP II in this reaction. However, the transcription did not proceed by *de novo* initiation, but rather by cleavage of the RNA template followed by extension of the new 3' end, generating a chimeric template/transcript product. Conversely, because the accumulation of the antigenomic species is resistant to higher doses of α-amanitin, it is suggested that the synthesis of the antigenome could be carried out by a polymerase other than RNAP II (Macnaughton and Lai, 2002a; Macnaughton et al., 2002; Modahl et al., 2000).

To date, few proteins are reported to be associated with HDV. HDAg associates with HDV genomic RNA to form a ribonucleoprotein (RNP) complex in the HDV-containing HBV virion and in transfected cells (Chang et al., 1988; Lin et al., 1990; Ryu et al., 1993). In addition, HDAg can function as an RNA chaperone and modulate the ribozyme activity of HDV RNA (Huang and Wu, 1998). Host proteins that were identified to be associated with HDV biology can be divided in two categories: HDAg- or HDV RNA-interacting proteins. The host proteins reported to interact with HDAg are the cellular double-stranded RNA-dependent protein kinase (PKR), nucleolin, and a cellular homolog of HDAg named *delta*-interacting protein A (DIPA; Brazas and Ganem, 1996; Chen et al., 2002; Lee et al., 1998). The host proteins observed to interact directly with HDV RNA are PKR, the negative elongation factor (NELF), glyceraldehyde 3-phosphate dehydrogenase (GAPDH), and the small form of adenosine deaminase acting on RNA (ADAR-1; Circle et al., 1997; Robertson

et al., 1996; Wong and Lazinski, 2002; Yamaguchi et al., 2002). ADAR-1 is reported to carry out the post-transcriptional RNA editing of the HDV genome and GAPDH is believed to be involved in the regulation of HDV ribozyme activity (Lin et al., 2000). However, the physiological significance of the interaction of most of these proteins remains enigmatic and the identification of additional host factors involved in viral replication is required to gain a better understanding of HDV replication.

To identify additional host proteins involved in HDV replication, we analyzed the protein content of a ribonucleoprotein complex produced by UV cross-linking an HDV-derived RNA previously reported to act as an RNA promoter *in vitro* (Beard et al., 1996) to proteins included in NE from HeLa cells. Mass spectrometry of the complex led to the identification of the polypyrimidine tract-binding protein (PTB)-associated splicing factor (PSF) as an HDV-binding protein. Using co-immunoprecipitation with HeLa NE proteins and electrophoretic mobility shift assays (EMSAs) of a series of HDV-derived RNAs with purified, recombinant hexahistidine-tagged PSF, we have found that PSF binds directly to both terminal stem-loop domains of both polarities of HDV RNA. In addition, the interaction between HDV RNA and PSF was demonstrated *in vivo* using HeLa cells containing both polarities of HDV RNA. Together, these results suggest a role for host mRNA processing proteins in HDV replication and/or propagation and the possible involvement of PSF in the life cycle of HDV.

4.4 Results

Formation of a specific complex between the right terminal stem-loop domain of genomic HDV RNA and HeLa NE proteins.

In order to identify host proteins that interact with HDV RNA, we used a 208 nt RNA fragment containing 199 nt derived from the right terminal stem-loop domain of the genomic polarity of HDV (R199G; nucleotides 1541 to 61) as bait. We selected this fragment because it includes the reported initiation site for HDAg mRNA transcription (i.e. position 1630; Gudima et al., 2000) and a similar HDV-derived RNA fragment was previously shown to be able to initiate transcription *in vitro* using HeLa NE proteins (Beard et al., 1996). In addition, the stem-loop structure conformation of this domain has been confirmed *in vitro* under approximately physiological conditions (Beard et al., 1996). Figure 4.1 shows the location, the sequence, and the putative secondary structure of the radiolabeled synthetic RNA used as bait in our study.

We performed UV cross-linking assays of internally radiolabeled R199G with proteins contained in HeLa NE. To obtain a better cross-linking yield, photoreactive 4-thio-uridine was randomly incorporated into R199G during its *in vitro* synthesis. In addition, an excess amount of total yeast tRNA was added to each sample to ensure the specificity of the interaction. Following UV irradiation and treatment with RNase A to remove excess probe, the mixtures were fractionated by SDS-PAGE. UV cross-linking of NE proteins to the radiolabeled bait generated a single major band migrating as a species of approximately 250 kDa (Figure 4.2). This complex was only detected in the presence of NE, indicating that formation of the radiolabeled complex was dependent on the presence of the NE proteins. This complex was also detected in the presence of heparin, poly(A) RNA, and P11.60 RNA as non-specific competitors. P11.60 is a small RNA fragment derived from the terminal

Figure 4.1: Location and secondary structure of the HDV-derived RNA used as bait. A) Schematic representation of the genome of HDV. The structure illustrates the superimposition of the features of both the genomic and antigenomic HDV polarities. The genome is depicted in its circular, unbranched rod-like conformation. The numbering is in accordance with Kuo *et al.* (1988a). The location of the genomic and antigenomic ribozymes, open reading frame for HDAgs, and various RNA fragments used are illustrated. B) Secondary structure of the R199G RNA fragment used as bait to identify host proteins that interact with HDV RNA.

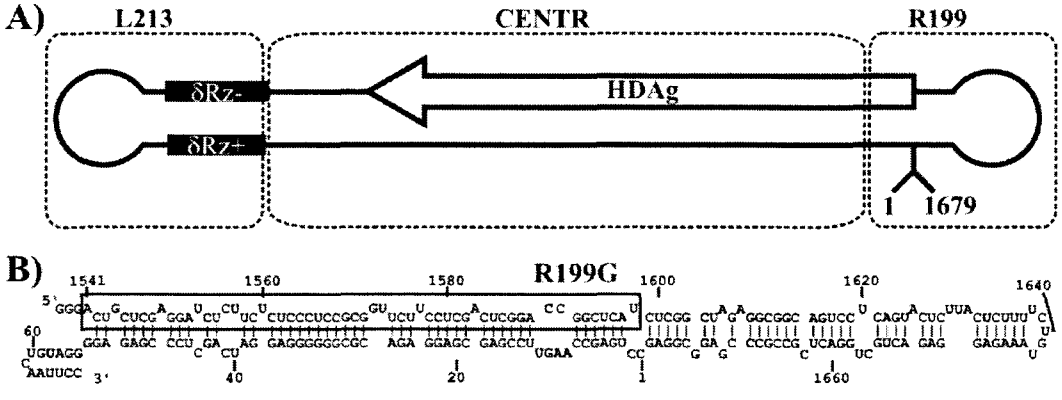
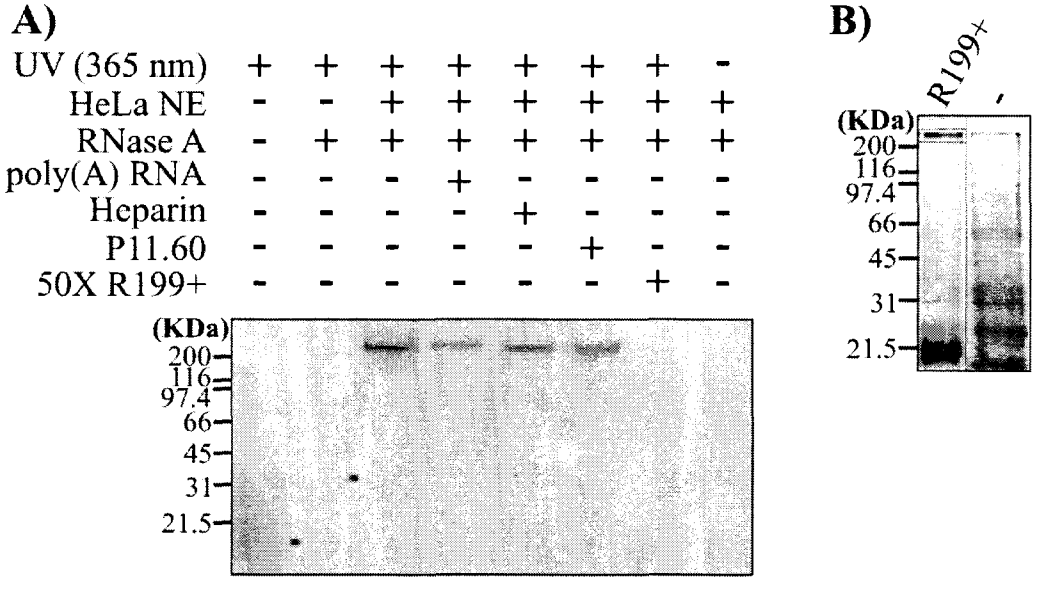

A)
L213
CENTR
R199
δRz-
δRz+
HDAg
1
1679
B)
R199G
1541
1560
1580
1600
1620
1640
1660
40
20
1

Figure 4.2: Detection and specificity of the cross-linked complex between R199G and HeLa NE proteins. A) Internally radiolabeled R199G containing 4-thio-uridine was UV cross-linked at 365 nm with HeLa NE proteins, treated with RNase A, and resolved by SDS-PAGE. 1 pmol of radiolabeled R199G and 100 μg/ml of total yeast tRNA was used for these experiments. Depending on the assays, 10 μg HeLa NE, heparin (100 μg/ml), poly(A) RNA (100 μg/ml), 20 pmol of P11.60 (Pelchat and Perreault, 2004), or 50 pmol of non-labeled R199G were used. Indicated protein markers were visualized by Coomassie blue staining. B) Methylene blue-stained gel of the cross-linking reactions performed with or without 50 pmol of R199G.


A)
UV (365 nm) + + + + + + + -
HeLa NE - - + + + + + +
RNase A - + + + + + + +
poly(A) RNA - - - + - - - -
Heparin - - - - + - - -
P11.60 - - - - - + - -
50X R199+ - - - - - - + -
(KDa)
200
116
97.4
66
45
31
21.5
B)
R199+
-
(KDa)
200
116
97.4
66
45
31
21.5

stem-loop domain of the peach latent mosaic viroid genome which has been demonstrated to have promoter activity in the presence of the *Escherichia coli* RNAP (Pelchat and Perreault, 2004), providing an excellent competitive RNA species for this experiment. As an additional control to test the specificity of the interaction, unlabeled R199G was included in the reaction. Inclusion of a 50-fold molar excess of unlabeled, homologous competitor RNA (i.e. R199G) inhibited formation of the radiolabeled RNA-protein complexes. Together, these observations indicate that at least one protein present in the NE from uninfected HeLa cells is able to bind specifically to the right terminal stem-loop region of genomic HDV RNA.

Identification of the polypyrimidine tract-binding protein-associated splicing factor (PSF) as a protein cross-linked to the HDV RNA fragment.

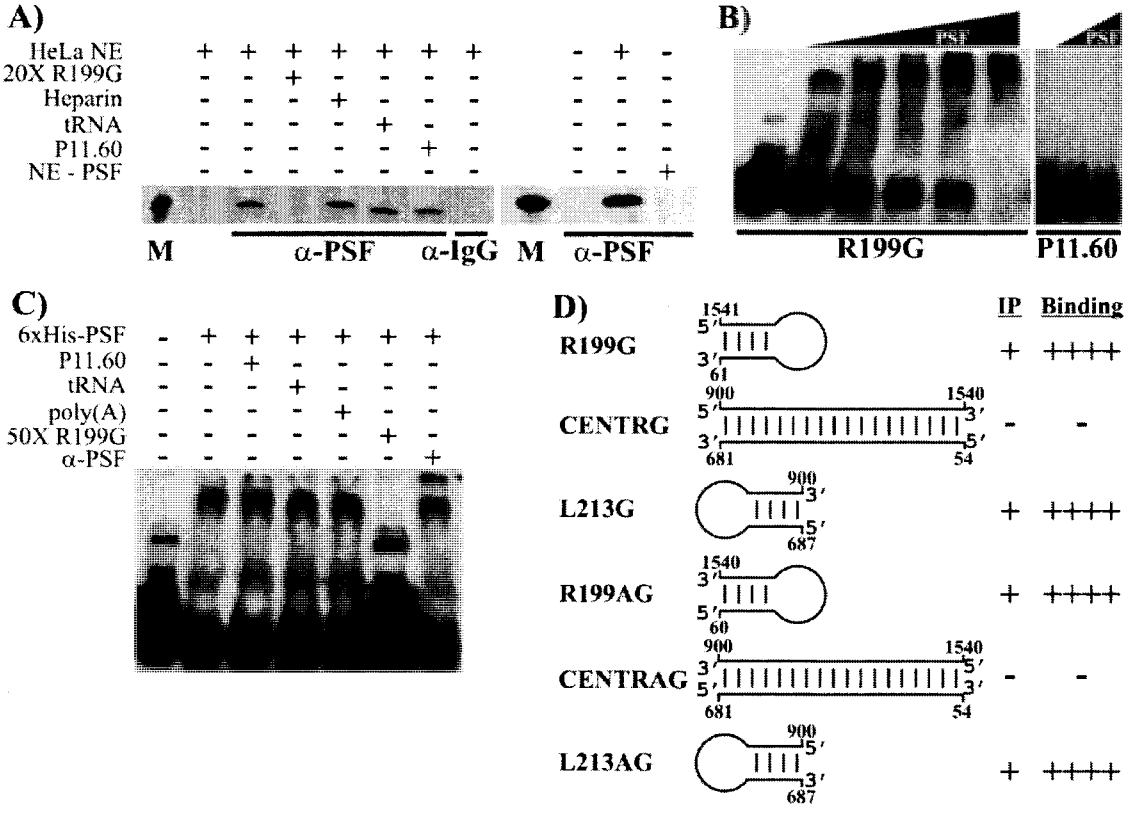
The band corresponding to the UV cross-linked complex between non-radiolabeled R199G and HeLa NE proteins was excised from methylene blue-stained gels (Figure 4.2B), subjected to in-gel trypsin digestion, and the resulting peptides were eluted and analyzed by mass spectrometry (LC-MS/MS). At the same time, UV irradiation was performed on a sample without R199G and subjected to the same treatment. This sample was used as a negative control in the identification of the proteins interacting with the HDV-derived RNA. As expected, when the peptide amino acid sequence information was used to search human protein sequence databases, many proteins with high molecular masses were detected. However, three peptides whose protonated masses corresponded to YGEPGEVFINK, LFVGNLPADITEDEFK, and NLSPYVSNELLEEAFSQFGPIER, matching the 76 kDa human polypyrimidine tract-binding protein (PTB)-associated splicing factor (PSF) were repeatedly detected in the sample containing R199G. This multifunctional protein, also known as splicing factor proline/glutamine-rich (SFPQ), possesses a DNA-binding domain,

two RNA recognition motifs and two nuclear localization sequences (Shav-Tal and Zipori, 2002; Song et al., 2005). In addition, PSF forms heterotetramers with $p54^{nrb}$ and interacts with many other proteins (Patton et al., 1993; Peng et al., 2002; Shav-Tal and Zipori, 2002). Notably, PSF is involved in the post-transcriptional regulation of HIV-1 where it is speculated to repress *env* transcription, mediate nuclear retention, and promote shuttling to the cytosol when complexed with $p54^{nrb}$ (Zolotukhin et al., 2003). Nevertheless, this result suggests that PSF is a nuclear protein that can interact with the right terminal stem-loop region of genomic HDV RNA.

The terminal stem-loop domain of both polarities of HDV RNA co-immunoprecipitate with PSF within HeLa NE.

To demonstrate the association of R199G with PSF, co-immunoprecipitation experiments with α-PSF antibody were performed on samples containing radiolabeled R199G in HeLa NE. A typical gel is shown in Figure 4.3A. A band corresponding to radiolabeled R199G was detected when PSF was immunoprecipitated with the α-PSF antibody. The interaction of PSF with the HDV probe was considerably abolished by an excess of unlabeled homologous competitor, but not by unrelated competitors including heparin, total yeast tRNA, or P11.60 RNA, clearly indicating the specific interaction of HDV RNA with PSF. This radioactive band was not observed when the experiment was performed in the absence of the α-PSF antibody, when using HeLa NE depleted of PSF by α-PSF antibody prior to incubation (i.e. NE - PSF), or when using an unrelated antibody for the immunoprecipitation (i.e. α-IgG). These observations unambiguously demonstrate that a complex including PSF interacts specifically with R199G within HeLa NE.

Figure 4.3: Interaction of various HDV-derived RNAs with PSF. A) Co-immunoprecipitation of radiolabeled R199G with PSF in HeLa NE. 1 pmol of radiolabeled R199G was used for these experiments. Depending on the assays, 10 μg HeLa NE, heparin (100 μg/ml), total yeast tRNA (100 μg/ml), 20 pmol of P11.60, 50 pmol of non-labeled R199G, or 25 μg of immunodepleted extract for PSF were used. Lanes M contain radiolabeled R199G. B) Titration of a constant amount of radiolabeled R199G or P11.60 with increasing amounts of hexahistidine-tagged PSF. 0, 0.4, 0.8, 1.2, 1.6, and 2.0 μM 6xHis-PSF or 0, 0.8 and 1.2 μM 6xHis-PSF were used for R199G or P11.60, respectively. C) EMSA of R199G bound to 4 pmol of hexahistidine-tagged PSF in the presence of either 20 pmol of P11.60, total yeast tRNA (100 μg/ml), poly(A) RNA (100 μg/ml), 50 pmol of non-labeled R199G, or anti-PSF antibody. D) Interaction of PSF with various HDV-derived RNAs. The polarity of each HDV-derived RNA is indicated. For the RNAs subjected to EMSA analysis, the resulting binding efficiency was quantified densitometrically and compared to R199G (-: <5%, +: 6%-33%, ++: 34%-56%, +++: 57%-79%, ++++: >80%).


A)
HeLa NE
20X R199G
Heparin
tRNA
P11.60
NE - PSF
M
α-PSF
α-IgG
M
α-PSF
B)
PSF
R199G
P11.60
C)
6xHis-PSF
P11.60
tRNA
poly(A)
50X R199G
α-PSF
D)
IP
Binding
R199G
CENTRG
L213G
R199AG
CENTRAG
L213AG

In order to establish whether or not PSF interacts with other regions of HDV RNA, radiolabeled RNA transcripts representing various regions of both genomic and antigenomic HDV RNA were synthesized *in vitro* and used in co-immunoprecipitation experiments as above. The association of these RNA molecules with PSF is summarized in Figure 4.3D. Results indicate that RNA regions consisting of the stem-loop domains of both polarities of HDV RNA co-immunoprecipitate with PSF (i.e. R199G, R199AG, L213G, and L213AG), whereas the central regions of both HDV polarities fail to co-immunoprecipitate specifically with PSF (i.e. CENTRG and CENTRAG). To guarantee that both RNA molecules forming the CENTR RNAs annealed, both strands were mixed, heated, slowly cooled, and tested for double-strand formation. Electrophoresis on a native agarose gel confirmed that most of the RNA molecules forming the CENTR RNAs annealed under our conditions (data not shown). These results indicate that PSF associates specifically with HDV RNA at locations included within the terminal stem-loop domains of both polarities of HDV RNA within HeLa NE.

PSF directly binds to the terminal stem-loop domains of both polarities of HDV RNA.

To confirm the binding of PSF to the HDV-derived RNAs, PSF tagged at its N-terminal end with a hexahistidine peptide was expressed in *Escherichia coli* and purified by immobilized Ni^{2+} affinity chromatography. Purified, recombinant PSF was then used in EMSA experiments. Radiolabeled R199G or the P11.60 negative control were allowed to bind with increasing amounts of PSF for 5 min at 4°C, and then the samples were subjected to electrophoresis through a non-denaturing 5% polyacrylamide gel. A typical gel showing a titration of a constant amount of either radiolabeled R199G or P11.60 with increasing amounts of protein is shown in Figure 4.3B. Purified, recombinant PSF binds directly to R199G, as illustrated by the retardation of the migration of the complex when compared to

free RNA. Conversely, recombinant PSF does not bind to the unrelated RNA molecule (i.e. P11.60). The binding affinity of PSF for R199G was then determined using this gel. Our results suggest a high dissociation constant (K_D) for the interaction between R199G and PSF (i.e. 0.372 +/- 0.031 μM). However, we found that some of the RNA consistently degraded during the assays, and that this degradation was dependent on protein concentration.

Since a high K_D is often associated with a low binding specificity, the binding of R199G to PSF was then studied by competition with either P11.60, yeast total tRNA, poly(A) RNA, or with unlabeled homologous competitors (Figure 4.3C). Binding of recombinant PSF to radiolabeled R199G was abrogated by the addition of unlabeled, homologous competitor RNA, but not by the unrelated competitors. In addition, monoclonal PSF antibody was able to supershift the radiolabeled complex, attesting to the binding of PSF.

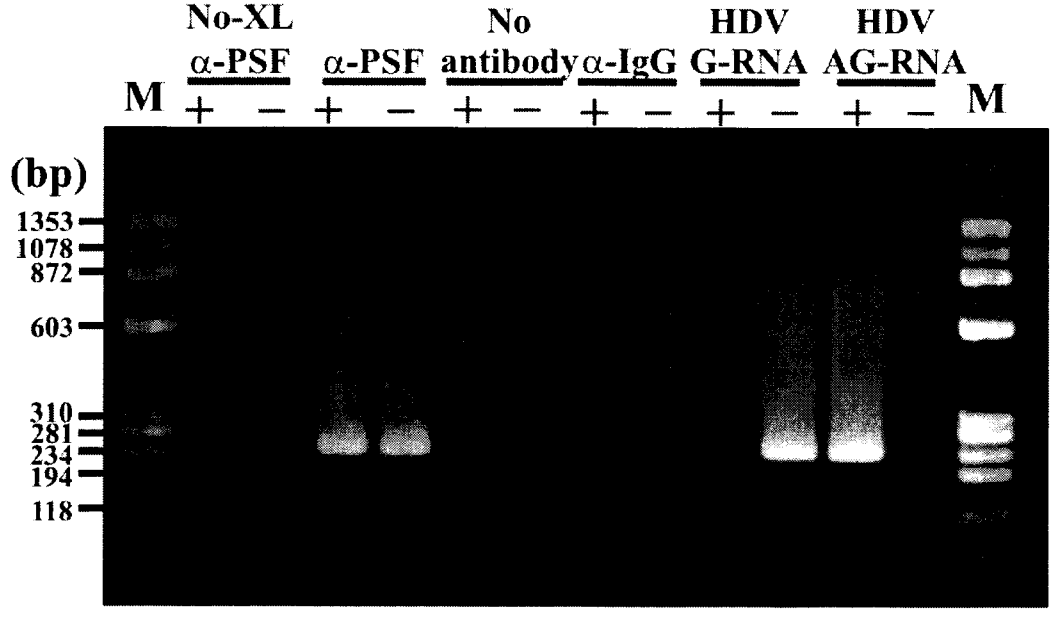
Similar EMSA experiments were then conducted with radiolabeled RNAs representing the various regions of the HDV genome used in the co-immunoprecipitation experiments (Figure 4.3D). In correlation with the results obtained using co-immunoprecipitation in HeLa NE, PSF was found to bind specifically to the RNA regions consisting of the stem-loop domains of both polarities of HDV RNA (i.e. R199G, R199AG, L213G, and L213AG), but not to the central regions (i.e. CENTRG and CENTRAG). Although we cannot exclude that additional factors missing in these experiments, such as $p54^{nrb}$ or HDAg, could facilitate the binding of PSF to the HDV-derived RNAs, our results clearly indicate that the recombinant PSF binds directly and specifically to the terminal stem-loop domains of both polarities of HDV RNA.

PSF interacts with HDV RNA in HeLa cells containing both polarities of HDV RNA.

In vivo, PSF associates with PTBs (Patton et al., 1993) and with the small nuclear U1A ribonucleoprotein (Lutz et al., 1998). It was also reported to specifically bind U5 snRNA (Peng et al., 2002) and to be associated with a variety of nuclear complexes. The previous experiments clearly indicate specific binding of PSF to the terminal stem-loop domains of both polarities of HDV RNA *in vitro*. Because PSF and HDV RNA might not interact in a cellular environment where HDV RNA has to compete for PSF binding and where their respective localizations might differ, we decided to investigate the validity of this interaction *in vivo*. HeLa cells were transfected with a dimeric HDV transcript of genomic polarity. Because HDAg-S is required to initiate HDV RNA replication (Kuo et al., 1989), the cells were co-transfected with a plasmid transiently expressing HDAg-S from a CMV promoter. Four days post-transfection, cells were harvested and the RNA-protein complexes cross-linked *in vivo* using a ribonucleoprotein immunoprecipitation (RIP) assay that employs formaldehyde as a reversible cross-linking agent (Niranjanakumari et al., 2002).

Following cross-linking with formaldehyde and co-immunoprecipitation using α-PSF antibody, the cross-links were reversed by heating the samples. The RNA was isolated and analyzed by RT-PCR exploiting both primers used to produce the R199G cDNA (Figure 4.4). A band corresponding to the R199G cDNA (i.e. 232 bp) was detected when the cross-linked extract was subjected to co-immunoprecipitation with the α-PSF antibody. Cloning and subsequent sequence analysis confirmed the identity of the HDV-derived cDNA fragments. However, R199G cDNA was not detected when the experiment was performed in the absence of cross-linking, without α-PSF antibody, or when an unrelated antibody (i.e. α-IgG) was used. These results indicate that the band detected when the α-PSF antibody was used in the co-immunoprecipitation experiment was not resultant of RT-PCR products

Figure 4.4: Association of HDV RNA with PSF in HeLa cells. HeLa cells transfected with HDV RNA were used for the preparation of cross-linked lysates for co-immunoprecipitation using anti-PSF antibody, no antibody, or goat anti-mouse IgG antibody. After cross-linking, the isolated RNA samples were reverse transcribed using either the sense (+) or antisense (-) primer used to produce both polarities of the R199G region, respectively. Following PCR with the two primers, the R199G cDNA was resolved on an agarose gel. Lanes M contain a DNA ladder (PhiX174 digested with *Hae*III); HDV G-RNA and HDV AG-RNA lanes are RT-PCR product controls prepared from either synthesized genomic or antigenomic HDV RNA.


No-XL
α-PSF
α-PSF
No
antibody
α-IgG
HDV
G-RNA
HDV
AG-RNA
M
+ − + − + − + − + − + −
M
(bp)
1353
1078
872
603
310
281
234
194
118

obtained from non-specific binding of HDV RNA to either the α-PSF antibody or the protein-G agarose beads, and that it was dependent on the presence of the α-PSF antibody. The presence of these RNA fragments in the cross-linked, co-immunoprecipitated samples provides evidence that formaldehyde covalently cross-linked HDV RNA to PSF, implicating their interaction *in vivo*.

The RT-PCR reactions were also designed to differentiate between RNA molecules derived from both polarities of HDV. For instance, the sense and antisense primers used to generate the R199G cDNA were only able to reverse transcribe from either the antigenomic or the genomic polarity of HDV RNA, respectively (Figure 4.4, HDV RNA lanes). Because the RT-PCR reactions were not quantitative, no information about the binding efficiency of PSF to the HDV RNAs was revealed. Since RNA fragments corresponding to both polarities of HDV were detected when the cross-linked extract was co-immunoprecipitated with the α-PSF antibody, we conclude that PSF associates with both polarities of HDV RNAs within HeLa cells.

4.5 Discussion

Previous work from several laboratories has demonstrated that viruses often alter host nuclear organization for their own replication by mechanisms such as RNA retention and sequestration to subnuclear domains for transcription and RNA processing. A number of cellular proteins that associate with both HDAgs and HDV RNA and that are proposed to participate in viral replication have been identified. However, the association of many of these proteins with HDV RNA has not been clearly confirmed both *in vitro* and *in vivo*. In the current investigation, we have used UV cross-linking to identify a cellular protein that binds to the right terminal stem-loop region of HDV RNA. We anticipated two properties

required for a host protein that would be capable of interaction with HDV RNAs: nuclear localization and RNA-binding activity. Accordingly, mass spectrometric analysis revealed PSF, a sequence-specific nuclear RNA-binding protein that appears to be predominantly involved in RNA splice site selection, as a likely HDV-interacting candidate. This interaction was confirmed *in vitro* by co-immunoprecipitation of HDV RNAs with PSF within HeLa NE and by the direct binding of HDV-derived RNA fragments to recombinant PSF. Furthermore, our experiments demonstrated that PSF binds specifically to both polarities of HDV RNA through the regions located at the extremities of their rod-like structures. In addition, we have confirmed the biological relevance of the PSF interaction with HDV RNA within HeLa cells using a method based on that of Niranjanakumari *et al.* (2002), which has previously been successfully performed to confirm specific interaction between HDAg and HDV RNA. This method, which uses reversible cross-linking with formaldehyde, is also commonly used in the chromatin immunoprecipitation (ChIP) assay which has proven to be a powerful method of delineating many DNA-protein interactions (Jackson, 1978; Orlando, 2000; Orlando et al., 1997; Shang et al., 2000).

Curiously, UV cross-linking of R199G to HeLa NE did not yield a radiolabeled band corresponding to the molecular weight of PSF. Because PSF is known to multimerize and to interact with many human proteins, one possibility is that a high molecular weight complex including PSF could have been produced during the UV cross-linking experiments. Alternatively, many HeLa NE proteins could have been able to cross-link with R199G, which would account for the detection of a single high molecular weight complex. R199G includes the reported initiation site for the HDAg mRNA transcription (Gudima et al., 2000) and a similar HDV-derived RNA fragment was previously shown to be able to initiate transcription *in vitro* (Beard et al., 1996). It is therefore possible that undetected host RNAP

or related host factors cross-linked to R199G to generate the high molecular weight complex observed.

Typically, PTB proteins are involved in pre-mRNA splicing in the nucleus. PSF is a multifunctional 76 kDa protein that is comprised of an N-terminal RGG motif, a glycine-rich domain which might be involved in protein-protein interactions, a proline/glutamine-rich domain, a DNA-binding domain, two canonical RNA recognition motifs (RRMs), and a C-terminal region possessing two nuclear localization signals (NLSs), and a high proportion of charged residues (Patton et al., 1993; Shav-Tal and Zipori, 2002; Song et al., 2005). *In vivo,* PSF forms a heterotetramer with $p54^{nrb}$ and participates in both early and late steps of splicing. PSF associates with PTB (Patton et al., 1993) and with the small nuclear ribonucleoprotein U1A (Lutz et al., 1998), whereas both $p54^{nrb}$ and PSF specifically bind U5 snRNA *in vitro* (Peng et al., 2002). The mechanism by which PSF influences pre-mRNA splice site selection is not known, although there is some evidence that it could recognize a specific RNA structure within certain splice acceptor sites. In addition to splicing, PSF was reported to be involved in a variety of nuclear processes (for a review, see Shav-Tal and Zipori, 2002), including transcription, topoisomerase activity, nuclear retention of defectively-edited RNAs, and DNA recombination (Akhmedov and Lopez, 2000; Mathur et al., 2001; Sewer et al., 2002; Song et al., 2005; Straub et al., 1998; Urban et al., 2000; Zhang and Carmichael, 2001). Because HDV RNA interacts with PSF, and PSF binds to many host factors and is involved in a variety of processes, it is likely that during infection HDV RNA-PSF interaction might disrupt one or several of these cellular activities. This hypothesis is consistent with a recent report showing that binding of non-coding mouse VL30 retroelement RNA to PSF induces multiple genes repressed by PSF by releasing PSF from the promoter of suppressed genes (Song et al., 2005). It is also possible that HDAgs might modulate such

interactions and processes, even if they are not required for the binding of PSF to HDV RNA *in vitro*. Investigation of the consequences of HDV RNA replication on the various PSF-mediated processes might thus lead to a better comprehension of HDV pathogenesis.

The specific biochemical function of PSF in HDV replication is not clear. Although precise activities of PSF in the nucleus are not fully understood, it is tempting to hypothesize at least two roles for the interaction between HDV RNAs and PSF. First, PSF interaction might be involved in HDV localization and/or retention within the nucleus. Like many proteins involved in RNA processing (Misteli and Spector, 1998), PSF was demonstrated to localize to splicing factor-enriched nuclear subdomains known as speckles (Dye and Patton, 2001). Additionally, a complex including PSF and $p54^{nrb}$ was reported to mediate RNA retention in the nucleus by directly binding to inosine-rich determinants in hyperedited RNAs (Zhang and Carmichael, 2001), suggesting that RNA might be retained via anchoring to insoluble nuclear components. Because HDV is directly recognized by the components of this complex, it could share this proposed mechanism. This hypothesis is consistent with previous work suggesting the transient co-localization of HDAgs with the splicing factor SC35 in nuclear speckles early after infection (Bichko and Taylor, 1996). However, later during infection, HDAgs are no longer associated with SC35, and it was reported that a subpopulation of HDAg-L and of antigenomic HDV RNA spatially associates with ND10 in PML bodies (Bell et al., 2000, Bichko and Taylor, 1996). Although we cannot exclude the involvement of other host proteins or the role of HDAgs in subnuclear localization, it is tempting to interpret the localization and retention of HDV ribonucleoproteins to nuclear speckles, as observed during early infection, as a result of their binding to PSF.

Furthermore, PSF might be involved with HDV RNA transcription and/or replication. In virus-infected cells, PTB and PTB-associated proteins participate in the transcription and

replication of viral RNA (reviewed in Lai, 1998). For example, adenoviruses have been shown to recruit replication and splicing factors to the sites of active viral replication and transcription (Jimenez-Garcia and Spector, 1993; Pombo et al., 1994). Particularly, PSF was shown to be one of the key factors mediating the post-transcriptional regulation of HIV-1 (Zolotukhin et al., 2003). PSF and $p54^{nrb}$ have also been implicated in transcriptional regulation based on the finding that both proteins avidly bind to the carboxy-terminal domain (CTD) of the largest subunit of RNAP II (Emili et al., 2002). *In vitro*, PSF binds to both polarities of HDV through the terminal stem-loop domains of the rod-like structure. Both the R199G and L213AG RNA segments were reported to be able to initiate transcription *in vitro* using HeLa NE proteins (Beard et al., 1996, Filipovska and Konarska, 2000). A cDNA fragment corresponding to R199AG was also able to initiate transcription *in vitro* in this system (Macnaughton et al., 1993). PSF could therefore provide a direct physical link between the CTD of RNAP II and the putative HDV-derived RNA promoters at the initiation phase of transcription by interacting with both the RNAP II CTD and RNA simultaneously. However, direct evidence of the involvement of PSF in the transcription/replication reaction will require further investigation.

In summary, we have shown that PSF binds specifically and preferentially to the two terminal stem-loop domains of both polarities of HDV RNA. Identification of PSF as a specific binding partner of HDV RNA provides new opportunities for isolating and characterizing HDV RNA-protein complexes. Specifically, further examination of the biochemical activities of PSF and identification of other cellular proteins that interact with this protein, such as $p54^{nrb}$, in the context of viral RNA replication and propagation should provide important information about the mechanisms of HDV replication and their effects on host cell metabolism.

4.6 Materials and methods

Synthesis of RNAs.

The various HDV-derived RNAs were synthesized by *in vitro* run-off transcription using T7 RNA polymerase (New England Biolabs; NEB) from DNA templates generated by PCR amplification of sequences of pHDVd2 using Vent DNA Polymerase (NEB). pHDVd2 is a derivative of pBluescriptKS$^+$ (Stratagene) containing a dimeric *Eco*RI-flanked HDV cDNA insert (Kuo et al., 1988a). P11.60 RNA was made as previously described (Pelchat and Perreault, 2004). HDV dimeric genomes were similarly synthesized from *Hin*DIII-linearized pHDVd2. Double-stranded regions corresponding to the central domain of HDV were generated by annealing two strands of RNA that were made by PCR of the central region of the HDV genome followed by run-off transcription. Transcription products were subjected to DNAse I (Promega) digestion for 30 minutes at 37°C and the RNAs were fractionated by denaturing polyacrylamide gel electrophoresis (PAGE) containing 1X TBE buffer (100 mM Tris-borate, pH 8.3, 1 mM EDTA) and 7 M urea. RNA bands were visualized by UV shadowing, excised, eluted overnight at 4°C in elution buffer (500 mM ammonium acetate, 0.1% SDS), precipitated in ethanol, resuspended in 100 µl H_2O, passed through Sephadex G-50 columns (Amersham Pharmacia Biotech), precipitated in ethanol, suspended in 20 µl H_2O, and stored at -20°C. RNA was quantified spectrophotometrically at 260 nm. When 5'-radiolabeled RNA was required, 10 pmol of RNA was dephosphorylated using calf intestinal phosphatase (CIP; NEB), radiolabeled at the 5'-end using T4 polynucleotide kinase (NEB) according to the manufacturer's protocol in the presence of [γ-^{32}P]ATP (GE Health Sciences) and purified by phenol/chloroform extraction, passage through a Sephadex G-50 column, and ethanol precipitation. CENTR molecules were generated using two RNA strands folded

together by heat denaturation at 94°C and slow renaturation at room temperature in 40 mM Tris-HCl, pH 8.0/10 mM $MgCl_2$.

UV cross-linking of RNA-protein complexes.

Cross-linking of RNA to proteins was performed as previously described (Pelchat and Perreault, 2004). Briefly, R199G was transcribed in the presence of [α-^{32}P]GTP (GE Health Sciences) and 4-thio-uridine triphosphate (TriLink BioTech) to increase cross-linking efficiency, cleaned by phenol/chloroform extraction, purified using a Sephadex-G50 column, and ethanol precipitated. Formation of the complexes was performed using 1 pmol radiolabeled 4-thio-uridine-containing R199G, 100 µg/ml total yeast tRNA, and 10 µg of HeLaScribe NE (Promega) in HeLaScribe 1X Transcription Buffer (20 mM HEPES [pH 7.9], 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol) in a total reaction volume of 10 µl. Some reactions were also performed in the presence of 100 µg/ml heparin (Sigma-Aldrich), 100 µg/ml poly(A) RNA, 20 pmol P11.60 RNA (Pelchat and Perreault, 2004), or 50 pmol of non-labeled R199G. Cross-linking was performed by 20 minute exposure of the reaction mixture to 365 nm light at 4°C using a UV cross-linker (Hoefer) at a distance of ~ 4 cm. Mixtures were then treated with RNase A (QIAgen), diluted 1:10 in SDS loading dye and resolved by 10% SDS-PAGE (Sambrook et al. 1989). Bands were detected with a phosphorimager (Molecular Dynamics) after overnight exposure to phosphorimager screens at 4°C.

Mass spectrometry and database analysis.

Complexes containing HeLa NE proteins cross-linked to either non-radioactive R199G containing 4-thio-uridine with 100 μg/ml total yeast tRNA or 100 μg/ml total yeast tRNA alone were resolved by 10% SDS-PAGE and the gel was stained with methylene blue. The resulting bands were excised and three samples were sent individually to WEMB Biochem Inc. (Toronto, Canada) for protein analysis by LC-MS/MS. The proteins were in-gel digested with trypsin and the resulting peptides analyzed by LC-MS/MS with a ThermoFinnigin Deca XP ion trap mass spectrometer. Bioworks 3.1 version software was used to process the MS/MS data, and NCBI nrdatabase was used for determination of peptide sequence identity.

Co-immunoprecipitation reactions.

Co-immunoprecipitation experiments were performed using the Protein-G Immunoprecipitation Kit (Sigma-Aldrich) according to manufacturer's recommended protocols with some modifications. 1 pmol 5'-radiolabeled RNA was incubated with 25 μg HeLaScribe NE in a 97.5 μl reaction volume of 1X IP buffer (50 mM Tris–Cl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl) at 4°C for 30 minutes. Total yeast tRNA (100 μg/ml), heparin (100 μg/ml), 20 pmol P11.60, or 20 pmol of non-radioactive HDV-derived RNA was added to some mixtures. Additionally, 50 μg of NE was immunodepleted of PSF using 5 μg α-PSF antibody (mouse monoclonal IgG SFPQ antibody [B92]; Abcam) and the Protein-G Immunoprecipitation Kit according to manufacturer's protocols; immunodepletion was confirmed by Western blot analysis (Sambrook et al., 1989) and the presence of residual NE proteins determined by SDS-PAGE followed by Coomassie blue staining. 25 μg of the immunodepleted extract was

used for co-immunoprecipitation. 5 μg of either α-PSF antibody or unspecific antibody (goat monoclonal α-mouse IgG; Sigma-Aldrich) was added to the reaction mixtures and incubated for 2 hours at 4°C with agitation prior to the addition of 30 μl pre-washed protein-G beads. Completed reactions were incubated at 4°C overnight with agitation. Samples were centrifuged at 16,000 g, washed 4 times in 1X IP buffer, once in 0.1X IP buffer, and eluted according to manufacturer's protocols. Eluates were electrophoresed by denaturing PAGE and bands were detected by phosphorimager scanning.

Purification of PSF.

Plasmid pET15b-PSF (Patton et al., 1993) containing PSF cDNA with a hexahistidine tag at its N-terminus was transformed into competent *Escherichia coli* BL21 (Novagen) for PSF overproduction (Sambrook et al., 1989). Protein was purified by immobilized metal affinity chromatography using the Ni-NTA Spin Kit (QIAgen) according to the manufacturer's protocol for purification under native conditions. Protein purity was estimated by Coomassie blue staining of SDS-PAGE gels (Sambrook et al., 1989), concentration was determined using Quick Start Bradford Protein Assay Kit 1 (Bio-Rad Laboratories), and identity confirmed by Western Blot analysis (Sambrook et al., 1989). The samples were stored at -20 °C in 20 mM HEPES [pH 7.9], 100 mM KCl, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF and 20% glycerol.

EMSA analysis.

Formation of the complexes was performed using 1 pmol of 5'-radiolabeled RNA with various concentration of purified, recombinant hexahistidine-tagged PSF in HeLaScribe 1X

Transcription Buffer. Following incubation for 5 minutes at 4°C, agarose loading dye (Sambrook et al., 1989) was added and the complexes were resolved by 5% PAGE in 1X TBE under non-denaturing conditions. Bands were visualized using phosphorimager scanning. Some reactions were also performed in the presence 100 μg/ml of total yeast tRNA (Sigma-Aldrich), 100 μg/ml heparin (Sigma-Aldrich), 20 pmol P11.60, or 50 pmol of non-radioactive HDV-derived RNA. For supershift experiments, 2.5 μg of α-PSF antibody was added to the mixture.

HeLa in vivo RNA cross-linking experiments.

Experiments were conducted using the ribonucleoprotein immunoprecipitation (RIP) assay as described by Niranjanakumari *et al.* (Niranjanakumari et al., 2002) with some modifications. Briefly, HeLa cells were grown to 60% confluence and transfected with 6 μg of pcDNA3-AgS and 10 μg HDV dimeric RNA genome of positive polarity using Lipofectamine TM 2000 Reagent (Invitrogen) according to the manufacturer's protocol. pcDNA3-AgS is a derivative of the pcDNA3 expression vector (Stratagene) encoding the small *delta* antigen under the control of a CMV promoter (Lazinski and Taylor, 1993). Cross-linking was performed by the addition of 1% formaldehyde to ~10^6 cells followed by 5 minute incubation. Cells were pelleted, resuspended in 2 ml 1X IP buffer, homogenized by sonication, and centrifuged at 16,000 g for 20 minutes at 4°C. The soluble fraction of the cell lysate was pre-cleared by the addition of total yeast tRNA (100 μg/mL), incubation with protein-G beads for one hour at 4°C with agitation, and centrifugation at 1,300 g. RNA fragmentation was omitted from the protocol. Co-immunoprecipitation was performed using the Protein-G Immunoprecipitation Kit (Sigma-Aldrich) using 500 μl of cell lysate and 5 μg

of α-PSF antibody. Mixtures were then incubated for 3 hours at 4°C with agitation, 50 µl of pre-washed protein-G beads were added, and samples were further incubated overnight. Samples were washed 5 times with high-stringency wash buffer (50 mM Tris–Cl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl, 1 M urea; Niranjanakumari et al., 2002), heated to 70°C for 45 minutes in buffer containing 50 mM Tris-HCl, pH 7.0, 5 mM EDTA, 10 mM dithiotriol (DTT), and 1% SDS to reverse cross-links, and eluted from the column by centrifugation. Proteins were removed by phenol/chloroform extraction and the RNA was isolated by ethanol precipitation. The resuspended nucleic acids were treated with DNAse I (Promega) for 45 minutes at 37°C. Samples were again phenol/chloroform extracted, ethanol precipitated, and subjected to reverse transcription using each primer used to generate the R199G cDNA (5'-GGAATTCTAATACGACTCACTATAGGGACTGCTCGAGGATCTCTTCTCTCCC-3' and 5'-GGAATTCACATCCCCTCTCGGGTGC-3'; underlined nucleotides indicate the T7 promoter) as described (Pelchat et al., 2001). PCR was then performed using this primer pair, products resolved by 2% agarose gel electrophoresis, and the bands were visualized under UV light in the presence of ethidium bromide. Bands were excised, cloned and submitted for sequencing to the University of Ottawa Core DNA Sequencing and Synthesis Facility (Ottawa).

4.7 Acknowledgments

We are grateful to J. Patton (Vanderbilt University, TN, United States) for providing the hexahistidine-tagged, PSF-harboring plasmid. We thank J.P. Perreault (University of Sherbrooke, Canada) for pHDVd2, the HDAg-S harboring plasmid, and useful comments.

We also thank M. Gallant and E.G. Brown (University of Ottawa, Canada) for technical assistance in cell transfection and cell maintenance. V.G.-S. is supported by a Canadian Graduate Scholarship from the Natural Science and Engineering Research Council of Canada (NSERC). This research was supported by a grant from the Canadian Institutes of Health Research (CIHR) awarded to M.P.

4.8 References

Akhmedov, A.T., Lopez, B.S., 2000. Human 100-kDa homologous DNA-pairing protein is the splicing factor PSF and promotes DNA strand invasion. Nucleic Acids Res. 28, 3022-3030.

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Bell, P., Brazas, R., Ganem, D., Maul, G.G., 2000. Hepatitis *delta* virus replication generates complexes of large hepatitis *delta* antigen and antigenomic RNA that affiliate with and alter nuclear domain 10. J. Virol. 74, 5329-5336.

Bichko, V.V., Taylor, J.M., 1996. Redistribution of the *delta* antigens in cells replicating the genome of hepatitis *delta* virus. J. Virol. 70, 8064-8070.

Brazas, R., Ganem, D., 1996. A cellular homolog of hepatitis *delta* antigen: implications for viral replication and evolution. Science 274, 90-94.

Chang, F.L., Chen, P.J., Tu, S.J., Wang, C.J., Chen, D.S., 1991. The large form of hepatitis *delta* antigen is crucial for the assembly of hepatitis *delta* virus. Proc. Natl. Acad. Sci. 88, 8490-8494.

Chang, M.F., Baker, S.C., Soe, L.H., Kamahora, T., Keck, J.G., Makino, S., Govindarajan, S., Lai, M.M., 1988. Human hepatitis *delta* antigen is a nuclear phosphoprotein with RNA-binding activity. J. Virol. 62, 403-410.

Chen, C.W., Tsay, Y.G., Wu, H.L., Lee, C.H., Chen, D.S., Chen, P.J., 2002. The double-stranded RNA-activated kinase, PKR, can phosphorylate hepatitis D virus small *delta* antigen at functional serine and threonine residues. J. Biol. Chem. 277, 33058-33067.

Chen, P.J. , Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., Taylor, J., 1986. Structure and replication of the genome of hepatitis *delta* virus. Proc. Natl. Acad. Sci. 83, 8774-8778.

Circle, D.A., Neel, O.D., Robertson, H.D., Clarke, P.A., Mathews, M.B., 1997. Surprising specificity of PKR binding to *delta* agent genomic RNA. RNA 3, 438-448.

Dye, B.T., Patton, J.G., 2001. An RNA recognition motif (RRM) is required for the localization of PTB-associated splicing factor (PSF) to subnuclear speckles. Exp. Cell Res. 263, 131-144.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Filipovska, J., Konarska, M., 2000. Specific HDV RNA-templated transcription by pol II *in vitro*. RNA 6, 41-54.

Fu, T., Taylor, J., 1993. The RNAs of hepatitis *delta* virus are copied by RNA polymerase II in nuclear homogenates. J. Virol. 67, 6965-6972.

Gudima, S., Wu, S.Y., Chiang, C.M., Moraleda, G., Taylor, J., 2000. Origin of hepatitis *delta* virus mRNA. J. Virol. 74, 7204-7210.

Huang Z.S., Wu, H.N., 1998. Identification and characterization of the RNA chaperone activity of hepatitis *delta* antigen peptides. J. Biol. Chem. 273, 26455-26461.

Jackson, V., 1978. Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. Cell 15, 945-954.

Jimenez-Garcia, L.F., Spector, D.L., 1993. *In vivo* evidence that transcription and splicing are coordinated by a recruiting mechanism. Cell 73, 47-59.

Kuo, M.Y., Chao, M., Taylor, J., 1989. Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of the *delta* antigen. J. Virol. 63, 1945-1950.

Kuo, M.Y., Goldberg, J., Coates, L., Mason, W., Gerin, J., Taylor J.M., 1988a. Molecular cloning of hepatitis *delta* virus RNA from an infected woodchuck liver: sequence, structure, and applications. J. Virol. 62, 1855-1861.

Kuo, M.Y., Sharmeen, L., Dinter-Gottlieb, G., Taylor, J., 1988b. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis *delta* virus. J. Virol. 62, 4439-4444.

Lai, M.M., 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology 244, 1-12.

Lai, M.M., 2005. RNA replication without RNA-dependent RNA polymerase: surprises from hepatitis *delta* virus. J. Virol. 79, 7951-7958.

Lazinski, D.W., Taylor, J.M., 1993. Relating structure to function in the hepatitis *delta* virus antigen. J. Virol. 67, 2672-2680.

Lee, C.H., Chang, S.C., Chen, C.J., Chang, M.F., 1998. The nucleolin binding activity of hepatitis *delta* antigen is associated with nucleolus targeting. J. Biol. Chem. 273, 7650-7656.

Lin, J.H., Chang, M.F., Baker, S.C., Govindarajan, S., Lai, M.M., 1990. Characterization of hepatitis *delta* antigen: specific binding to hepatitis *delta* virus RNA. J. Virol. 64, 4051-4058.

Lin, S.S., Chang, S.C., Wang, Y.H., Sun, C.Y., Chang, M.F., 2000. Specific interaction between the hepatitis *delta* virus RNA and glyceraldehyde 3-phosphate dehydrogenase: an enhancement on ribozyme catalysis. Virology 271, 46-57.

Luo, G.X., Chao, M., Hsieh, S.Y., Sureau, C., Nishikura, K., Taylor, J., 1990. A specific base transition occurs on replicating hepatitis *delta* virus RNA. J. Virol. 64, 1021-1027.

Lutz, C.S., Cooke, C., O'Connor, J.P., Kobayashi, R., Alwine, J.C., 1998. The snRNP-free U1A (SF-A) complex(es): identification of the largest subunit as PSF, the polypyrimidine-tract binding protein-associated splicing factor. RNA 4, 1493-1499.

Macnaughton, T.B., Gowans, E.J., MacNamara, S.P., Burrell, C.J., 1991. Hepatitis *delta* antigen is necessary for access for hepatitis *delta* virus RNA to the cell transcriptional machinery but is not part of the transcriptional complex. Virology 184, 387-390.

Macnaughton, T.B., Beard, M.R., Chao, M., Gowans, E.J., Lai, M.M., 1993. Endogenous promoters can direct the transcription of hepatitis *delta* virus RNA from a recircularized cDNA template. Virology 196, 629-636.

Macnaughton, T.B., Lai, M.M., 2002. Genomic but not antigenomic hepatitis *delta* virus RNA is preferentially exported from the nucleus immediately after synthesis and processing. J. Virol. 76, 3928-3935.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M.M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

Mathur, M., Tucker, P.W., Samuels, H.H., 2001. PSF is a novel corepressor that mediates its effect through Sin3A and the DNA binding domain of nuclear hormone receptors. Mol. Cell. Biol. 21, 2298-2311.

Misteli, T., Spector, D.L., 1998. The cellular organization of gene expression. Curr. Opin. Cell Biol. 10, 323-331.

Modahl, L.E., Macnaughton, T.B., Zhu, N., Johnson, D.L., Lai, M.M., 2000. RNA-dependent replication and transcription of hepatitis *delta* virus RNA involve distinct cellular RNA polymerases. Mol. Cell. Biol. 20, 6030-6039.

Niranjanakumari, S., Lasda, E., Brazas, E., Garcia-Blanco, M.A., 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods 26, 182-190.

Orlando, V., Strutt, H., Paro, R., 1997. Analysis of chromatin structure by *in vivo* formaldehyde crosslinking. Methods 11, 205-214.

Orlando, V., 2000. Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. Trends Biochem. Sci. 25, 99-104.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P., Nadal-Ginard, B., 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev. 7, 393-406.

Pelchat, M., Côté, F., Perreault, J.P., 2001. Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol. 146, 1753-1763.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636-642.

Peng, R., Dye, B.T., Perez, I., Barnard, D.C., Thompson, A.B., Patton, J.G., 2002. PSF and $p54^{nrb}$ bind a conserved stem in U5 snRNA. RNA 8, 1334-1347.

Pombo, A., Ferreira, J., Bridge, E., Carmo-Fonseca, M., 1994. Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells. EMBO J. 13, 5075-5085.

Rizzetto, M., Canese, M.G., Arico, S., Crivelli, O., Trepo, C., Bonino, F., Verme, G., 1977. Immunofluorescence detection of new antigen-antibody system (*delta*/anti-*delta*) associated to hepatitis B virus in liver and in serum of HBsAg carriers. Gut 18, 997-1003.

Robertson, H.D., Manche, L., Mathews, M.B., 1996. Paradoxical interactions between human *delta* hepatitis agent RNA and the cellular protein kinase PKR. J. Virol. 70. 5611-5617.

Ryu, W.S., Netter, H.J., Bayer, M., Taylor, J., 1993. Ribonucleoprotein complexes of hepatitis *delta* virus. J. Virol. 67, 3281-3287.

Sambrook, J., Fritsch, E.F., Maniatis, T. (Eds.), 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

Sewer, M.B., Nguyen, V.Q., Huang, C.J., Tucker, P.W., Kagawa, N., Waterman, M.R., 2002. Transcriptional activation of human CYP17 in H295R adrenocortical cells depends on complex formation among p54nrbNonO, protein-associated splicing factor, and SF-1, a complex that also participates in repression of transcription. Endocrinology 143, 1280-1290.

Shang, Y., Hu, X., DiRenzo, J., Lazar, M.A., Brown, M., 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell 103, 843-852.

Shav-Tal, Y., Zipori, D., 2002. PSF and p54nrbNonO--multi-functional nuclear proteins. FEBS Lett. 531, 109-114.

Song, X., Sun, Y., Garen, A., 2005. Roles of PSF protein and VL30 RNA in reversible gene regulation. Proc. Natl. Acad. Sci. 102, 12189-12193.

Straub, T., Grue, P., Uhse, A., Lisby, M., Knudsen, B.R., Tange, T.O., Westergaard, O., Boege, F., 1998. The RNA-splicing factor PSF/p54^{nrb} controls DNA-topoisomerase I activity by a direct interaction. J. Biol. Chem. 273, 26261-26264.

Taylor, J.M., 2006. Hepatitis *delta* virus. Virology 344, 71-76.

Urban, R.J., Bodenburg, Y., Kurosky, A., Wood, T.G., Gasic, S., 2000. Polypyrimidine tract-binding protein-associated splicing factor is a negative regulator of transcriptional activity of the porcine p450scc insulin-like growth factor response element. Mol. Endocrinol. 14, 774-782.

Wang, K.S., Choo, Q.L., Weiner, A.J., Ou, J.H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L., Houghton, M., 1986. Structure, sequence and expression of the hepatitis *delta* viral genome. Nature 328, 508-514.

Weiner, A.J., Choo, Q.L., Wang, K.S., Govindarajan, S., Redeker, A.G., Gerin, J.L., Houghton, M., 1987. A single antigenomic open reading frame of the hepatitis *delta* virus encodes the epitope(s) of both hepatitis *delta* antigen polypeptides p24 *delta* and p27 *delta*. J. Virol. 62, 594-599.

Wong, S.K., Lazinski, D.W., 2002. Replicating hepatitis *delta* virus RNA is edited in the nucleus by the small form of ADAR1. Proc. Natl. Acad. Sci. 99, 15118-15123.

Yamaguchi, Y., Inukai, N., Narita, T., Wada, T., Handa, H., 2002. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. Mol. Cell. Biol. 22, 2918-2927.

Zhang, Z., Carmichael, G.G., 2001. The fate of dsRNA in the nucleus: a $p54^{nrb}$-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. Cell 106, 465-475.

Zolotukhin, A.S., Michalowski, D., Bear, J., Smulevitch, S.V., Traish, A.M., Peng, R., Patton, J., Shatsky, I.N., Felber, B.K., 2003. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Mol. Cell. Biol. 23, 6618-6630.

CHAPTER 5:

SUPPLEMENTARY DATA: PSF STIMULATES RNAP II BINDING TO HDV RNA *IN VITRO*

The data described in the following section provide additional, unpublished insight into the PSF-HDV RNA interaction and propose a physiological function for PSF in HDV RNA synthesis. These data will be used as a second-author contribution to an upcoming publication from the Pelchat laboratory. Experiments were performed as described in Chapters 2 through 4 with additional methodologies presented below. The results obtained and brief discussion is presented below and further implications of these data are discussed in Chapter 6.

5.1 Results

Determination of the dissociation constant (K_D) of the PSF- HDV RNA interaction.

In order to determine whether or not the binding of HDV RNA to PSF could occur under physiological conditions, the dissociation constant (K_D) for this interaction was determined as stated briefly in Chapter 4. To elaborate, K_D was determined by EMSA analysis using purified, recombinant PSF and radiolabeled R199G RNA, a species that was previously shown to interact with PSF (Chapter 4). Increasing amounts of PSF were incubated with a trace amount of radiolabeled R199G, EMSA performed, and results visualized by phosphorimager scanning (a representative gel is shown in Figure 4.3B). Densitometric analysis of the autoradiogram was used to determine the percentage of RNA bound by PSF at each PSF concentration and the data were plotted to generate a binding curve for this

interaction (Figure 5.1; graphic interpretation performed by M. Pelchat). The K_D for the PSF-R199G interaction was determined to be 0.372 $\pm$ 0.031 μM, a value consistent with K_D values observed for RNA-protein interactions under physiological conditions.

Interaction of PSF with HDV is localized to the extreme terminal stem-loop domains of both polarities of HDV RNA.

In Chapter 4, I demonstrated that binding of PSF to HDV RNA was localized to the terminal stem-loop domains of both polarities of HDV RNA. To further delineate the precise regions of HDV bound by PSF, small RNAs corresponding to 38 nt of the right terminal domain of genomic polarity (R38G) and 48 nt of the left terminal domain of antigenomic polarity (L48G) were synthesized and subjected to EMSA analysis with purified PSF (Figure 5.2). Both R38G and L48AG bound strongly to PSF, although the binding was slightly diminished as compared to binding of the longer R199G and L213AG molecules to PSF as observed by the presence of more free RNA on the EMSA gels. To further characterize the HDV RNA requirements for PSF binding, mutant variants of R38G and L48AG were constructed in which some of the base pairing was abrogated (R38Gmut1 and L48AGmut1; Figure 5.2). These mutants with disruptions in their secondary structures failed to interact with PSF. Mutant variants of R38G and L48AG species containing compensatory mutations that restored base pairing but retained sequential differences (R38mut2 and L48mut2; Figure 5.2) also largely failed to interact with PSF by EMSA. These data are reminiscent of those obtained by co-immunoprecipitation for the HDV RNA-RNAP II interaction; both extreme tips of both polarities of HDV RNA interacted with RNAP II, and mutation that altered the secondary structure by abrogation of base pairing or strand inversion prevented RNAP II

Figure 5.1: The dissociation constant (K_D) of the PSF-R199G HDV RNA interaction as determined by densitometric analysis. EMSA was performed using various concentrations of PSF and a trace of radiolabeled R199G; samples were resolved by native gel electrophoresis and autoradiography followed by densitometric analysis. The K_D was determined to be 0.372 ± 0.031 μM.

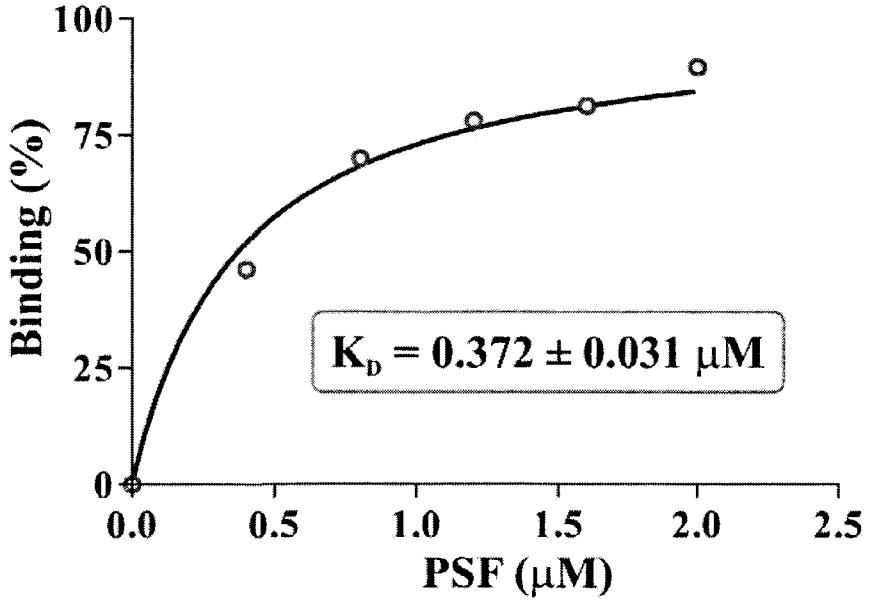
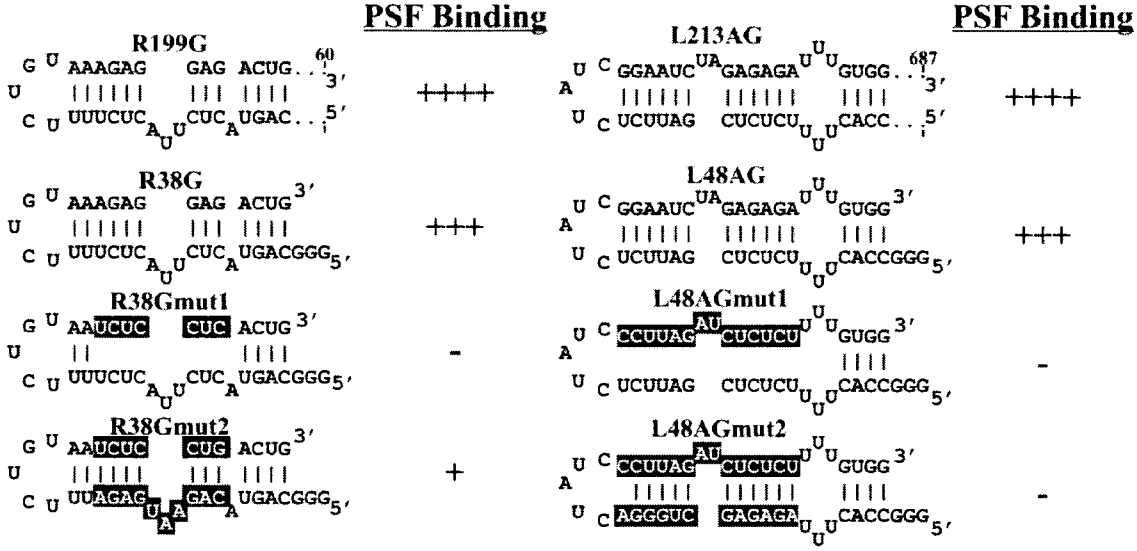

100
75
50
25
0
Binding (%)
0.0
0.5
1.0
1.5
2.0
2.5
PSF (μM)
K_D = 0.372 ± 0.031 μM

Figure 5.2: PSF binding to HDV RNA is localized to the extreme terminal stem-loop domains of both polarities of HDV RNA and is dependent on structure and sequence. Binding was determined by EMSA analysis; the resulting binding efficiency was quantified densitometrically and compared to that of R199G (-: <5%, +: 6%-33%, ++: 34%-56%, +++: 57%-79%, ++++: >80%).


PSF Binding
R199G
++++
R38G
+++
R38Gmut1
-
R38Gmut2
+
PSF Binding
L213AG
++++
L48AG
+++
L48AGmut1
-
L48AGmut2
-

binding to HDV RNA. These observations suggest that similar RNA requirements might be necessary for both PSF and RNAP II binding to HDV RNA, and that binding occurs at the same regions of HDV RNA for both PSF and RNAP II.

Binding of PSF to HDV RNA and RNAP II mediates the RNAP II-HDV RNA interaction.

While I have definitively established interaction between PSF and HDV RNA, the physiological relevance of this interaction has not been elucidated (Greco-Stewart et al., 2006). It has previously been demonstrated that PSF interacts with both the hypo- and hyperphosphorylated forms of the CTD of RNAP II in HeLa cell extract (Emili et al., 2002). PSF might thus act in recruiting HDV RNA to this polymerase for transcription and replication by providing a direct physical link between HDV RNA and RNAP II. To test whether or not PSF-HDV RNA and PSF-RNAP II interactions are necessary for RNAP II binding of HDV RNA, co-immunoprecipitation experiments were performed in HeLa nuclear extracts using the α-RNAP II antibody as described in Chapter 2. Although it would be tempting to use NE that had been immunodepleted of PSF to determine if RNAP II could still bind to HDV RNA in the absence of PSF, this approach was not used as immunodepletion of PSF might result in concurrent removal of RNAP II due to the aforementioned RNAP II-PSF interaction. Since PSF also interacts with other nuclear proteins including $p54^{nrb}$, immunodepletion of PSF might remove additional components of this complex, subsequently interfering with HDV RNA binding. Alternative approaches to disrupting the PSF-HDV RNA interaction *in vitro* were thus explored in order to perform these co-immunoprecipitation experiments.

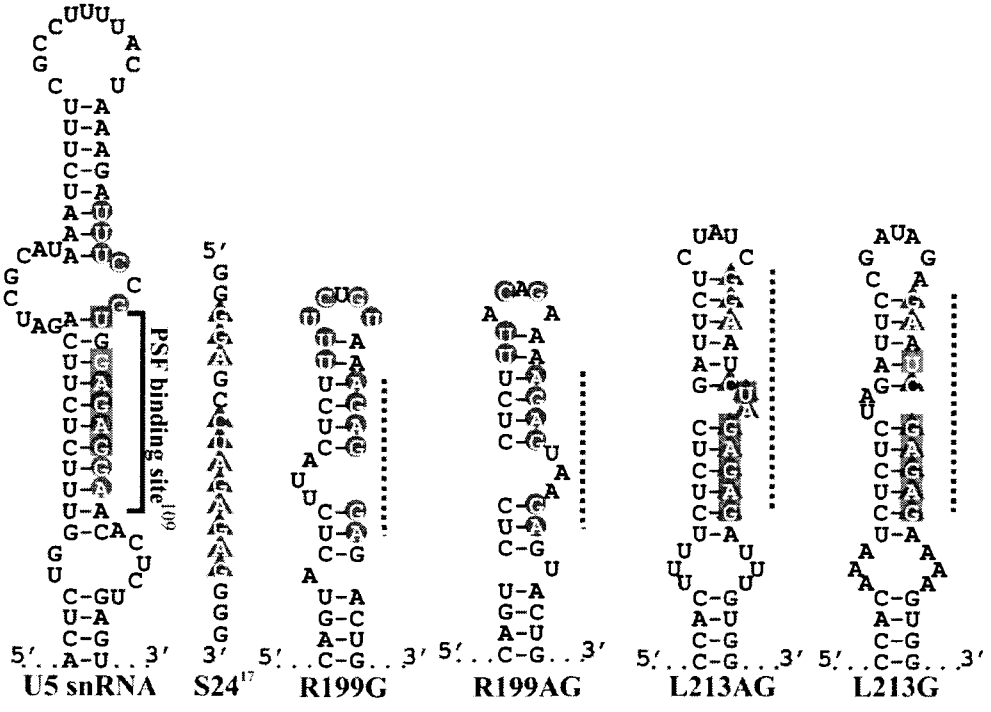
The Patton laboratory has described several PSF-binding RNA sequences that they obtained using SELEX (**s**ystematic **e**volution of **l**igands by **e**xponential **e**nrichment) and

showed the similarity that existed between these species and U5 snRNA, a nuclear RNA that is capable of PSF interaction (Peng et al., 2002). Bioinformatic analyses of U5 RNA, the SELEX-derived RNAs, and the extreme terminal stem-loop domains of HDV RNA of both polarities show sequence conservation not only between the SELEX-derived RNAs and U5 snRNA but also with both polarities of both terminal stem-loop domains of HDV RNA (Figure 5.3A). These SELEX-derived RNAs might thus provide a useful tool in studies involving the PSF-HDV interaction by competing for PSF binding. Since there are no known PSF-specific inhibitors, these SELEX-derived RNA species could be useful for *in vitro* experiments to prevent PSF binding to HDV RNA through competitive inhibition.

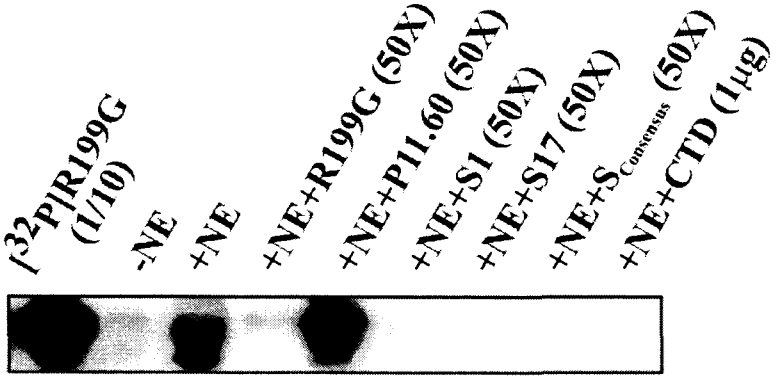
As an alternative to immunodepletion, the SELEX-derived RNA species that bind PSF were used to competitively inhibit HDV RNA binding to PSF in co-immunoprecipitation reactions. The competitive RNAs chosen included two SELEX-derived PSF-binding sequences (S1 and S17) and a trimeric repeat of the SELEX-derived PSF-binding consensus sequence ($S_{Consensus}$; Peng et al., 2002). A 50-fold molar excess of these competitor RNAs were added to some co-immunoprecipitation mixtures to examine the effects on RNAP II binding to HDV RNA. A sample co-immunoprecipitation using radiolabeled R199G and α-RNAP II antibody is shown in Figure 5.3B. As observed in previous experiments, R199G is co-immunoprecipitated in the presence of NE and in NE containing P11.60 hairpin RNA. In the absence of NE or when an excess of unlabeled, homologous competitor is added, R199G fails to co-immunoprecipitate with RNAP II. However, when S1, S17, or $S_{Consensus}$ are added to the co-immunoprecipitation experiments, R199G fails to co-immunoprecipitate with the RNAP II complex. These data suggest that the PSF-HDV RNA interaction is necessary for RNAP II binding to HDV RNA *in vitro*.

Figure 5.3: PSF facilitates binding of RNAP II to HDV RNA *in vitro*. A) HDV might use molecular mimicry of U5 snRNA to bind PSF. S24 is a SELEX-derived RNA sequence that has been demonstrated to bind PSF (Peng et al., 2002). B) PSF enables binding of RNAP II to R199G as determined by co-immunoprecipitation analysis. Experiments were performed using an α-RNAP II antibody in the presence of PSF-binding competitors such as three different SELEX-derived sequences (Peng et al., 2002) or purified RNAP II carboxy terminal domain (CTD).

A)


PSF binding site
U5 snRNA
S24
R199G
R199AG
L213AG
L213G

B)


[32P]R199G (1/10)
-NE
+NE
+NE+R199G (50X)
+NE+P11.60 (50X)
+NE+S1 (50X)
+NE+S17 (50X)
+NE+SConsensus (50X)
+NE+CTD (1μg)

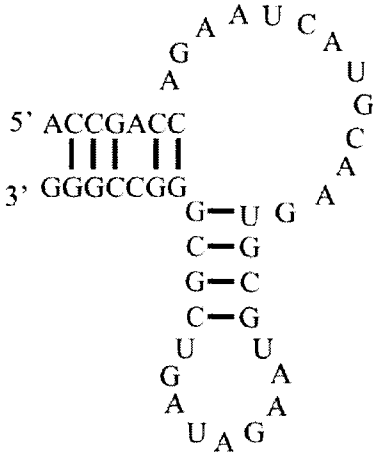
To further confirm that PSF is necessary for RNAP II recruitment to HDV RNA, I wanted to determine if the PSF-RNAP II interaction is essential for RNAP II binding to HDV RNA. Purified RNAP II CTD peptide was added to one of the co-immunoprecipitation experiments to compete with RNAP II for PSF binding (Figure 5.3B). An excess of CTD inhibits R199G co-immunoprecipitation with the α-RNAP II antibody, suggesting that PSF must bind the CTD of RNAP II to stimulate RNAP II binding to HDV RNA. Together, these data suggest that PSF interaction with both HDV RNA and the CTD of RNAP II are necessary for the RNAP II-HDV RNA interaction *in vitro*; PSF might thus facilitate or enhance the binding of RNAP II to HDV RNA in HDV-infected cells, acting like a transcription factor for HDV RNA synthesis.

PSF and RNAP II are present simultaneously on HDV RNA.

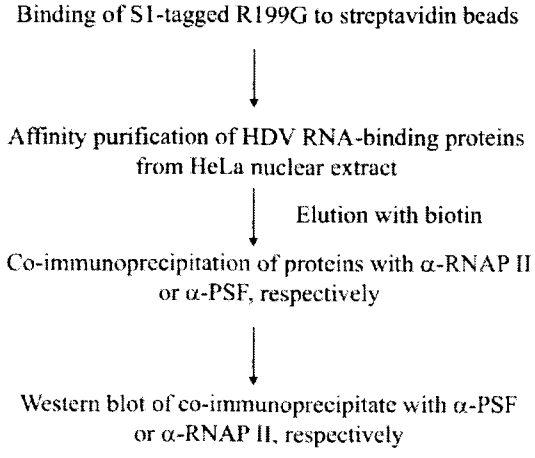
Since the experiments described above implicate PSF as a factor that might promote the binding of RNAP II to HDV RNA, I wanted to examine whether or not PSF and RNAP II are present simultaneously on HDV RNA. To this end, an RNA affinity chromatography approach was used to determine the composition of the complex formed on the R199G molecule as described in Chapter 3. Briefly, R199G containing the S1 aptamer (R199G-S1; not to be confused with the S1 SELEX-derived PSF-binding RNA) was transcribed *in vitro* to use as the bait molecule for RNA affinity chromatography. This aptamer was developed by Srisawat and Engelke (2002) and binds streptavidin with high affinity (Figure 5.4A). Embedding an S1 aptamer tag on an RNA species thus enables retention of S1-tagged RNA on affinity columns containing streptavidin-agarose due to S1 interaction with streptavidin; *d*-biotin also binds streptavidin and can be used to competitively displace S1-tagged RNA from the streptavidin-agarose beads for elution of S1-tagged RNA-containing complexes. *In*

Figure 5.4: PSF and RNAP II bind R199G HDV RNA simultaneously. A) The streptavidin-binding S1 aptamer used for affinity chromatography experiments (Srisawat and Engelke, 2002). B) Procedural overview of the tandem affinity chromatography-co-immunoprecipitation experiment used to resolve components of RNP complexes. C) Complexes co-immunoprecipitated with α-RNAP II contain PSF. Co-immunoprecipitated complexes were subjected to Western blot analysis to identify the presence of PSF. D) Reciprocally, complexes co-immunoprecipitated with α-PSF contain RNAP II. In both (C) and (D) P11.60 was unable to form complexes containing PSF and RNAP II and was unable to competatively abrogate R199G binding of these proteins.

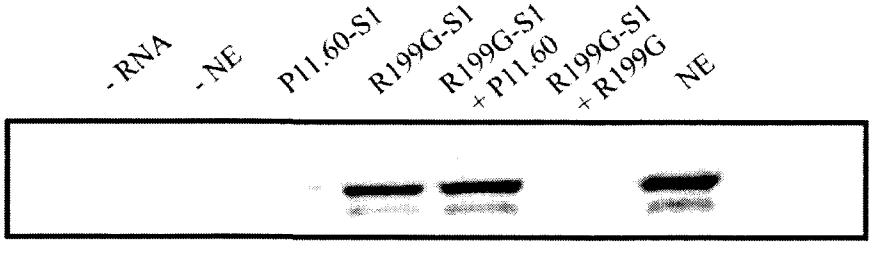
A)


5' ACCGACC
3' GGGCCGG

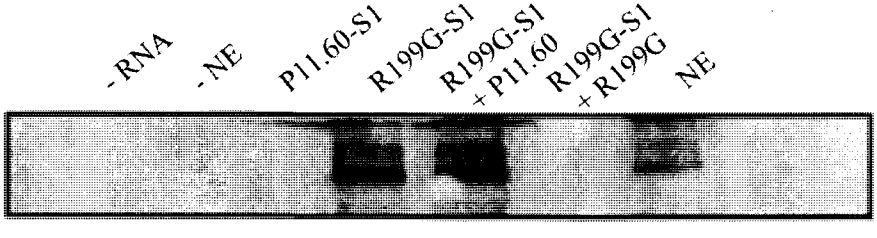
B)


Binding of S1-tagged R199G to streptavidin beads
Affinity purification of HDV RNA-binding proteins from HeLa nuclear extract
Elution with biotin
Co-immunoprecipitation of proteins with α-RNAP II or α-PSF, respectively
Western blot of co-immunoprecipitate with α-PSF or α-RNAP II, respectively

C)


-RNA
-NE
P11.60-S1
R199G-S1
R199G-S1 + P11.60
R199G-S1 + R199G
NE

D)


-RNA
-NE
P11.60-S1
R199G-S1
R199G-S1 + P11.60
R199G-S1 + R199G
NE

vitro-transcribed R199G-S1 RNA was incubated with HeLa NE in a column containing streptavidin agarose beads to capture and retain the complexes formed on R199G-S1. Following several washes, RNA-protein complexes were eluted by competition with *d*-biotin and the recovered eluate was subjected to co-immunoprecipitation with antibody directed against RNAP II. To detect the presence of PSF in the complexes, the co-immunoprecipitate was subjected to Western blot analysis using an α-PSF antibody. The reciprocal experiment was performed using the α-PSF antibody for co-immunoprecipitation followed by Western blot analysis with the α-RNAP II antibody. A procedural overview is depicted in Figure 5.4B.

When R199G-S1 RNA is incubated in HeLa NE, co-immunoprecipitated with α-RNAP II, and blotted using α-PSF antibody, RNP complexes containing PSF are recovered (Figure 5.4C). In the absence of either R199G-S1 RNA or NE, complexes containing PSF fail to form. These data indicated that complexes containing both RNAP II and PSF form on R199G *in vitro*. Addition of P11.60 hairpin to the chromatography mixture containing R199G-S1 does not prevent complex formation, and P11.60-S1 RNA is unable to form complexes with either of these two proteins. However, incubation with a 10-fold molar excess of R199G lacking the S1 aptamer tag competitively inhibits complex formation. Complimentary data is obtained when co-immunoprecipitation is performed using the α-PSF antibody followed by Western blotting with α-RNAP II antibody (Figure 5.4D), further substantiating the ability of these proteins to bind R199G simultaneously. These data show that a complex containing both PSF and RNAP II forms on R199G and that this complex is specific to HDV-derived R199G RNA insofar as complexes containing a structurally-similar P11.60 hairpin fail to form PSF- and RNAP II-containing complexes.

5.2 Discussion

In Chapter 4, I showed the novel interaction of HDV RNA with PSF and demonstrated that binding was localized to both terminal stem-loop domains of both polarities of HDV RNA. In this set of experiments, I further refine the regions of HDV RNA bound by PSF to the extreme terminal tips of these domains; interestingly, these are the same regions that interact with RNAP II (Chapter 2; Greco-Stewart et al., 2007). Given these observations, and the fact that PSF interacts with the CTD of RNAP II, it is tempting to speculate that PSF might act as a transcription factor for HDV by recruiting RNAP II to a non-DNA template (Emili et al., 2002; Greco-Stewart et al., 2006). PSF could thus provide a direct physical link between HDV RNA and RNAP II, promoting the non-canonical use of this RNA polymerase on an RNA template.

The investigations in the present chapter were designed to further characterize the PSF-HDV RNA interaction and to address the physiological role of PSF in HDV biology. Specifically, I wanted to determine whether or not PSF binding to HDV RNA and the CTD of RNAP II is essential for RNAP II recruitment to a segment of HDV RNA reported to have putative promoter activity and to determine whether or not both proteins bind this RNA species simultaneously (Beard et al., 1996). Using densitometric quantification of the PSF-HDV RNA complexes resolved by EMSA, a K_D value of approximately 300 nM was obtained for the PSF-HDV RNA interaction. Although binding was 100-fold weaker than binding of some RNA species to their protein partners, it is within the range of K_D values seen in numerous physiologically-relevant RNA-protein interactions, particularly those between viral RNAs and host proteins. Examples of reaction kinetics for RNA-protein interactions including K_D can be found in databases such as the Kinetic Data of Bio-molecular Interactions database (Ji et al., 2003; Kumar et al., 2009). A wide range of binding

affinities from 1 nM to almost 1 μM have been reported for various cellular RNA-protein interactions, thus the PSF-HDV RNA interaction could occur in a physiological context. The hyperbolic nature of the K_D curve implies the 1:1 stoichiometric binding of HDV to PSF. Given this data, it would seem that one or both of the RRMs of PSF might be responsible for binding a single HDV RNA molecule, and that binding can occur independently of PSF homodimerization or heterotetramerization with $p54^{nrb}$. However, these data should be interpreted with some caution; some degradation of R199G was observed when purified PSF was incubated for extended periods of time with R199G. PSF has been associated with the snRNP-free UA1 complex that is involved with 3' end cleavage and degradation during mRNA maturation, thus PSF might directly cause the degradation of RNA, although it was later speculated that $p54^{nrb}$ is likely the component of this RNP complex responsible for these activities (Lutz et al., 1998; Liang and Lutz, 2006). A more likely reason for this observed RNA degradation is contamination of the PSF protein preparation with bacterial proteins with endo- or exonucleolytic properties. Trace contaminants were routinely visualized by Coomassie staining of SDS-PAGE gels containing purified PSF; attempts to recover a large amount impurity-free PSF were unsuccessful due to significant growth impairment of *E. coli* cells expressing PSF and to the rapid conversion of PSF into a smaller, 68 kDa species during protein purification and manipulation as previously reported (Gozani et al., 1994; Patton et al., 1993). Furthermore, PSF is known to associate with exonucleases such as XRN2 *in vivo,* leading to speculation that similar nucleases might associate with PSF during bacterial overexpression and could concurrently be recovered with PSF during protein purification, leading to RNA degradation during EMSA analysis (Kaneko et al., 2007). To minimize RNA degradation, EMSA experiments were performed as soon as the HDV RNA

was introduced to the PSF-containing reaction mixtures, though it is still possible that some RNA degradation occurred, reducing the accuracy of K_D determination. In the future, an ameliorated protocol for PSF purification to reduce contamination might improve the accuracy of these experiments to more closely determine the K_D of the PSF-RNA interaction. Improved technologies such as surface plasmon resonance might also be useful in determining a more accurate K_D value.

Given the latter data presented in this chapter, it can be speculated that PSF mediates the HDV RNA-RNAP II interaction which subsequently facilitates HDV replication and transcription. In Chapter 6, these implications will be further discussed and a hypothetical model of HDV-directed RNA synthesis by RNAP II proposed.

5.3 Materials and methods

RNA synthesis.

RNA was synthesized as described previously in Chapters 2.6, 3.6, and 4.6. Briefly, RNA was synthesized by PCR amplification of regions of the HDV genome using Vent polymerase (New England Biolabs) from a plasmid containing dimeric cDNA of the HDV genome (HDVd2; Kuo et al., 1988a) using primers containing the T7 polymerase initiation site. Small RNA species (i.e. those comprising the extreme terminal domains of HDV RNA, their corresponding mutants, P11.60, and the SELEX-derived RNAs) were synthesized by primer extension; DNA primers containing the T7 polymerase initiation site and the desired sequence were subjected to five cycles of PCR amplification followed by *in vitro* transcription as described in Chapter 2.6. The primers for the SELEX-derived small RNAs described by Peng et al. (2002) possessed a T7 promoter at the 5' end flanked by one following sequences: S1 (5'-AGGTGGTCTGGAGCAAGCCT-3'), S17 (5'-

TAAGAGGTAGAATGAGAAAAG-3'), or $S_{Consensus}$ (5'-TGGAGAGGAAC TGGAGAGGAACTGGAGAGGAAC-3').

In the case of RNA species containing the S1 aptamer, the sequence of the aptamer was embedded in the primers to generate the streptavidin binding motif at the 3' end of the RNA following *in vitro* transcription (Srisawat and Engelke, 2002; Chapter 3.6). RNA was transcribed by run-off transcription using T7 RNA polymerase at 37°C for 2 to 4 hours. The product of this reaction was treated with DNAse (Promega) for 0.5 hours and electrophoresed on 10% denaturing acrylamide gel. RNA bands were resolved by UV shadowing, excised, and eluted from the gel in elution buffer containing 500 mM ammonium acetate, 0.1% SDS, and 20 mM EDTA. RNA was precipitated with 100% ethanol and 0.1% sodium acetate, resuspended in ddH_2O, and filtered through Sephadex-G50 resin (Amersham). Filtrates were ethanol precipitated and resuspended in ddH_2O to a final concentration of 50 pmol/μl as determined spectrophotometrically by measurement at 260 nm.

EMSA analysis.

EMSA experiments were performed as described in Chapter 4.6. To determine K_D, increasing amounts of PSF and a constant, trace amount of radiolabeled R199G RNA were used. The proportion of PSF binding was calculated densitometrically using phosphorimager scanning and ImageQuant software analysis (Molecular Dynamics). GraphPad Prism version 4.0 graphing software was used to generate the K_D binding curve.

Co-immunoprecipitation analysis.

Co-immunoprecipitation was performed using the Protein-G Immunoprecipitation Kit (Sigma-Aldrich) and R199G RNA that had been 5' end radiolabeled with [γ-^{32}P]ATP as previously described in Chapter 2.6 and according to manufacturer's protocol. Briefly, each reaction contained ~50 μg of HeLa NE (Promega), ~1 pmol of radiolabeled RNA, and 5 μg of α-RNAP II antibody 8WG16 (directed against the CTD of the Rpb1 subunit of RNAP II; Upstate). Competitors included a 50-fold molar excess of P11.60 RNA, unlabeled R199G, S1 RNA, S17 RNA, $S_{consensus}$ RNA, or 1 μg of purified RNAP II CTD (contains amino acids 1600-1700 of Rpb1 protein from *Saccharomyces cerevisiae* and is comprised of the consensus heptad repeat YSPTSPS; Abcam). Competitors were pre-incubated for 5 minutes at room temperature prior to the addition of radiolabeled R199G. Complexes were allowed to form for 1 hour at 4°C prior to addition of antibody. Co-immunoprecipitation reactions were then incubated overnight at 4°C with agitation, washed five times with RIPA buffer, and eluted from the columns with acrylamide loading dye (Sambrook et al., 1989). Samples were subjected to electrophoresis on 10% denaturing acrylamide gels and results visualized using autoradiography and phosphorimager scanning.

S1-affinity purification of ribonucleoprotein (RNP) complexes.

Species of R199G RNA were transcribed containing the S1 aptamer tag (Srisawat and Engelke, 2002; Figure 5.4A) embedded at the 3' end. S1 affinity chromatography was performed using 30 μl of pre-washed streptavidin agarose beads (Invitrogen) incubated with ~200 μg of HeLa NE and 1 nmol of RNA (either R199G-S1 or P11.60-S1) in a total reaction volume of 200 μl completed with S1 binding buffer [50 mM HEPES (pH 7.4), 10 mM $MgCl_2$, 100 mM NaCl]. Binding was allowed to proceed for 1 h at 4°C followed by a 3 hour

wash in 500 µl binding buffer. Samples were eluted with S1 binding buffer containing 5 mM *d*-biotin (Sigma-Aldrich) for 1 hour and then subjected to co-immunoprecipitation using the Protein-G Immunoprecipitation Kit (Sigma-Aldrich) and 5 µg of either α-PSF (mouse monoclonal IgG SFPQ antibody [B92]; Abcam) or α-RNAP II (8WG16; Upstate). Following co-immunoprecipitation, SDS loading dye was added to elute the samples and they were subjected to SDS-PAGE and Western blotting using the One Hour Complete Western Kit (Genscript) to detect RNP complexes using the previously described, reciprocal antibody (i.e. for the co-immunoprecipitation performed with α-RNAP II, the α-PSF antibody was used for Western blot analysis and vice versa).

5.4 References

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Gozani, O., Patton, J.G., Reed, R., 1994. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. EMBO J. 13, 3356-3367.

Greco-Stewart, V.S., Thibault, C.S., Pelchat, M., 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology 356, 35-44.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology 357, 68-78.

Ji, Z.L., Chen, X., Zhen, C.J., Yao, L.X., Han, L.Y., Yeo, W.K., Chung, P.C., Puy, H.S., Tay, Y.T., Muhammad, A., Chen, Y.Z., 2003. KDBI: kinetic data of bio-molecular interactions database. 31, 255-257.

Kaneko, S., Rozenblatt-Rosen, O., Meyerson, M., Manley, J.L., 2007. The multifunctional protein $p54^{nrb}$-PSF complex recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. Genes Dev. 21, 1779-1789.

Kumar, P., Han, B.C., Shi, Z., Jia, J., Wang, Y.P., Zhang, Y.T., Liang, L., Liu, Q.F>, Ji, Z.L., Chen, Y.Z., 2009. Update of KDBI: kinetic data of bio-molecular interaction database. Nucleic Acids Res. 37 (Database issue), D636-41.

Liang, S., Lutz, C.S., 2006, $p54^{nrb}$ is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. RNA 12, 111-121.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P., Nadal-Ginard, B., 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev. 7, 393-406.

Peng, R., Dye, B.T., Perez, I., Barnard, D.C., Thompson, A.B., Patton, J.G., 2002. PSF and $p54^{nrb}$ bind a conserved stem in U5 snRNA. RNA 8, 1334-1347.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual, 2^{nd} ed. Cold Spring Harbor Press, New York.

Srisawat, C., Engelke, D. R., 2002. RNA affinity tags for purification of RNAs and ribonucleoprotein complexes. Methods 26, 156-161.

CHAPTER 6:
DISCUSSION

6.1 Multiple DNA-dependent RNA polymerases are involved in the HDV life cycle

Much data exist supporting a role for RNAP II in HDV RNA synthesis, and evidence showing the involvement of RNAP I, at least in antigenome synthesis, has been reported as described in Chapter 1.3. The research presented in my doctoral thesis shows the direct interaction of RNAP II with both polarities of HDV RNA, strengthening the argument that RNAP II is involved in HDV replication. However, similar interaction with RNAP I was also demonstrated, providing further support to the theory that RNAP I might be involved in replication or transcription of HDV. A novel finding resulting from my research was the capacity of RNAP III to undergo similar interactions with HDV RNA; RNAP III has not previously been reported to be involved in any stage of the HDV life cycle. These findings contribute to a growing body of evidence regarding use of RNA templates by DNA-dependent RNA polymerases and provide the foundation for further research into the specific involvement of each polymerase in the HDV life cycle.

6.1.1 RNAP II interaction confirmed: evidence for multiple HDV RNA promoters

In Chapter 2, I show that RNAP II interacts with both polarities of HDV RNA both *in vitro* and in cells actively replicating HDV (Greco-Stewart et al., 2007). These data corroborate and expand on several other studies pertinent to the involvement of RNAP II in HDV replication. While I used the RIPA approach to determine interaction of RNAP II with HDV RNA in cells replicating HDV, data from the Taylor laboratory recently corroborated these findings using similar ChIP methodology (Chang et al., 2008). Two additional studies have

subsequently confirmed the interaction of the left terminal domain of the antigenome with RNAP II (Haussecker et al., 2008; Lehmann et al., 2007). The antibody used for the RIPA analysis that I performed recognizes the hypophosphorylated form of RNAP II, implying that HDV binding occurs during PIC formation or shortly thereafter during transcription initiation. In addition to using this antibody, researchers in the Taylor laboratory also performed the experiment using an antibody that recognizes the hyperphosphorylated form of RNAP II, providing evidence of interaction of RNAP II with HDV RNA during the elongation phase of transcription (Chang et al., 2008). They also observed a 4-fold enrichment in the amount of antigenome recovered by ChIP using the same antibody employed in my RIPA analysis, further corroborating my observations. This phenomenon might be due to differential stabilities of the two RNA species since the RIPA assay detects only the presence of RNA and cannot distinguish between species which are actively being transcribed from those accumulating post-transcriptionally. Since active transcription cannot be presumed based on these sets of experiments, evidence of RNAP II involvement in HDV replication is highly suggestive but not completely conclusive.

Using a co-immunoprecipitation approach, I delineated the RNAP II interaction to both extremities of both polarities of the HDV genome within a 38 nt (right terminal domain) and 48 nt (left terminal domain) region. Bioinformatic analyses of 81 sequences of the HDV quasi-species indicate extreme conservation of the structure of these terminal regions and the uniform presence of a CUG motif and high level of purine/pyrimidine polarization. My experiments further indicate that alteration of these termini resulting in abrogation of the base-paired structure or inversion of the upper and lower strands of the hairpins (the "flip mutants") results in loss of RNAP II binding. It is thus tempting to speculate that these terminal extremities might constitute four different promoter regions for HDV, two regions

per polarity. Previous studies have demonstrated the importance of conservation of the stem, loop, and bulge structures of the right terminal stem-loop domain and shown that RNA accumulation is reduced when these structures are altered (Gudima et al., 1999; Wu et al., 1997). Recent data also demonstrate that the stem-loop domain of the extreme left terminal of the antigenome of HDV can be used as a template by RNAP II, and that structure is the predominant determinant of the RNA-RNAP II interaction (Lehmann et al., 2007). Although structure is thought to be most important in RNAP II recognition, flip mutants were found not to be replication competent in previous reports, and my data have shown that these mutant RNA species fail to interact with RNAP II *in vitro* (Greco-Stewart et al., 2007; Wu et al., 1997). These findings could indicate that either the sequence of this area is important for RNAP II recognition or that movement of the three nucleotide external bulge to the opposite strand is a structural disruption significant enough to cause loss of RNAP II interaction and subsequent replication deficiency.

Since RNAP II has been shown to possess RdRp activity, my findings, combined with existing data regarding the requirements of HDV RNA necessary for HDV RNA synthesis, provide insight into which RNA features might serve as promoters for RNA-templated synthesis by RNAP II. Evidence for RdRp activities of T7 bacteriophage RNA polymerase, *E. coli* RNAP, plant RNAP II, plant NEP, and plant PEP have all been reported and characteristics required for these RNAPs to transcribe from RNA templates are summarized in Table 6.1 with additional, supplementary discussion provided in Chapter 7.1. Several features of these putative RNA promoters are similar for all RNAPs with RdRp activity. While naturally-occurring RNAs that serve as templates for RNAPs range in size from 50 to 1,700 nt, segments as small as 38 nt can bind RNAP II (Greco-Stewart et al., 2007). Small synthetic structures can also serve as templates for RNA-templated RNA

Table 6.1: Characteristics of RNA-templated synthesis by RNAPs.

Polymerase	RNA Template Features				Examples
	Size[a]	Structure	Sequence	Features	
BacteriophageT7	60 - 120 nt	• Single-stranded • Hairpin • 5' extensions	• A/U-rich • pGG- and pCC-primed	• Complimentary strands[b] • Palindromic character	X RNA Y RNA
E. coli RNAP	50-200 nt	• Single-stranded • Rod-like • Internal bulges • Central "bubble"	• A/U-rich • U start site • Intramolecular base-pairing	• 14-22 nt product *in vivo*	6S RNA
Plant NEP	~ 250-350 nt	• Single-stranded • Complex secondary structure	• A/U-rich • U start site (ASBVd) • A/G and U start sites (PLMVd)	• Complementary strands • Conserved GUC motif • Promoter located in terminal stem-loop	ASBVd PLMVd
Plant PEP	~ 250-350 nt	• Single-stranded • Complex secondary structure	• A/U rich • A/G and U start site • Conserved GUC motif • CAGACG triplet	• Complementary strands • Promoter located in terminal stem-loop	PLMVd
RNAP II (Plant)	~360 nt	• Single-stranded • Rod-shaped	• G/C-rich helical region • C/G start site	• Complementary strands • Promoters located at termini of rod-like structure • Short stem-loop domain can serve as a promoter	PSTVd CEV
RNAP II[c] (Human)	~1,700 nt	• Single-stranded • Rod-shaped • *delta* ribozymes	• G/C-rich helical regions • U/A start site, A is the first nucleotide incorporated in transcript • Conserved CUG triplet with purine/pyrimidine polarization	• Complementary strands • >74% base pairing • Conserved UUA bulge, terminal stem-loop domain at promoter • Multiple start sites for different transcripts	HDV

(a) Total size of RNA species reported to be used as a template.

(b) Transcription intermediates include a complimentary strand, one species is more abundant.

(c) RNAP I and possibly RNAP III are also involved in replicating/transcribing HDV.

synthesis by RNAP II *in vitro* (Lehmann et al., 2007). Minimal structures required for RNAP II-mediated transcription include features such as a 5 to 6 nt stem, a 4 to 6 nt terminal loop, and short 5' extension (Lehmann et al., 2007). All known templates for the RdRp activity of RNAPs are single-stranded but form hairpin or stem-loop structures due to high levels of intramolecular base pairing, and putative promoter regions for RNAP II, *E. coli* RNAP, and plant phage-like NEP seem to be located in such terminal, stem-loop domains (Beard et al., 1996; Filipovska and Konarska, 2000; Fels et al., 2001; Motard et al., 2008; Navarro and Flores, 2000; Pelchat et al., 2002). In the case of HDV, there is a strong selective pressure to maintain this secondary structure and, although there is sequential variability among quasi-species, there is preservation of base pairing by co-variation to retain the hairpin structure (Greco-Stewart et al., 2007). Additional structures, such as internal and external bulges, can be found in these RNA species, and in some cases these structures are essential for RNAP binding and RNA-templated synthesis. For example, bacterial 6S RNA contains a large central "bubble" which mimics a DNA promoter open complex to stimulate RNAP binding (Gildehaus et al., 2007; Wassarman and Saeker, 2006). Furthermore, most RNA species transcribed by RNAPs have a less abundant, complimentary strand which can serve as a replication intermediate. Together, these data provide a picture of characteristics requisite for RNAP recognition and RNA-templated synthesis and provide insight into what might constitute an RNA promoter for RNAPs.

All known RNA promoters for RNAPs contain double-stranded RNA regions at putative promoter sites. The terminal stem-loop domains of HDV RNA that bind to RNAP II contain double-stranded RNA regions which might contribute to recognition by RNAP II as well as other eukaryotic RNAPs. Although B-helical DNA is the conventional conformation thought to be present in human genomic DNA, it has been shown that alternate

conformations, such as A- and Z-helices, are present in the genome (Konopka et al., 1985; Ng et al., 2000). While RNAP II prefers the B-helical DNA conformation as its template, A-RNA, a conformation acquired by double-stranded RNA, has been shown to also be recognized by RNAP II (Kettenberger et al.; 2006; Lehmann et al., 2007). Furthermore, helix conformation seems to be dictated by sequence composition. It has been shown that regions containing polypurine stretches of greater than ten consecutive purine residues adopt an A-helical structure while purine/pyrimidine sequences that contain alternating purine and pyrimidine residues extending longer than six nucleotides acquire a Z-helical conformation (Konopka et al., 1985; Ng et al., 2000). It has subsequently been shown that the human genome, as well as other eukaryotic genomes, demonstrates a significant bias towards containing these stretches, and that this enrichment for these particular sequences is greater than one that would be expected to occur simply by chance (Ussery et al., 2002). While most of these alternate helical conformations occur in intergenic and non-coding regions, certain genes, such as rRNA and tRNA genes, are heavy in A-helical content (Antony et al., 1999). While the precise significance of these structures is unknown, it has been speculated that they might serve regulatory roles, such as mediating polymerase stalling and regulating gene expression, or might confer additional stability to certain genomic regions (Antony et al., 1999; Ussery et al., 2002). From these observations, it is implied that eukaryotic RNAPs are capable of recognizing these alternate helices as templates for transcription. Remarkably, the HDV genome contains a higher frequency of long polypurine and polypyrimidine tracts than would be expected to occur solely by chance despite its composition of an approximately equal percentage of purine and pyrimidine bases (Branch et al., 1993). Specifically, the extreme terminal tips of the HDV genome and antigenome demonstrate a significant amount purine/pyrimidine polarization (Figure 2.4). Polypurine and polypyrimidine enrichment

might not only contribute to the secondary structure of these regions of HDV RNA but also be preferentially recognized by eukaryotic RNAPs, to facilitate RNA-templated HDV synthesis.

6.1.2 RNAP I and RNAP III bind HDV RNA: evidence for additional polymerase involvement in the HDV life cycle

In addition to showing the binding of RNAP II to HDV RNA species both *in vitro* and in cells replicating HDV RNA, I showed that RNAP I and RNAP III are capable of this interaction (Greco-Stewart et al., 2009; Chapter 3). Binding could be localized to the terminal domains of the genome and the antigenome to regions of 199 nt (right terminal domains) and 213 nt (left terminal domains; Figure 3.2). Furthermore, it was shown that complexes containing each RNAP form independently of one another on R199G RNA, a species of RNA derived from the right terminal stem-loop domain of HDV RNA which is reported to have promoter activity for RNAP II and which I identified to bind to all three RNAPs (Beard et al., 1996; Greco-Stewart et al., 2007; 2009). These data not only support a role for RNAP I in HDV biology but also present the possibility that RNAP III might be involved in HDV RNA synthesis, an observation that has previously not been reported. Although binding of HDV RNAs to RNAPs I and III was demonstrated, these data are preliminary and it is premature to assign a definitive role to these polymerases in the HDV life cycle. While these data contribute to the growing body of evidence that multiple polymerases might be capable of HDV RNA synthesis in addition to RNAP II, they do not show whether or not transcription by these polymerases does occur and, in the event that RNAP I and RNAP III can use HDV RNA as a template for RNA-dependent RNA synthesis,

whether or not these polymerases play an active role in HDV transcription and replication in HDV-infected cells.

My data showing the binding of RNAP I to HDV RNA corroborate data suggestive of RNAP I-mediated HDV RNA synthesis. In Chapter 1.3.5, I describe observations supporting a role for RNAP I in synthesis of at least one HDV transcript, namely the antigenome, including differences in α-amanitin sensitivity, sub-nuclear localization, and requirement for HDAg modification in transcription of different HDV RNA species (summarized in Table 1.3). Furthermore, as previously discussed, both nucleophosmin and nucleolin interact with HDAg; both proteins have also been found in association with RNAP I (Huang et al., 2001; Lee et al., 1998; Murano et al., 2008; Rickards et al., 2007). Recent data has shown that both nucleophosmin and nucleolin enhance RNAP I-directed transcription (Murano et al., 2008; Rickards et al., 2007). Although this enhancement has been attributed to chromosomal remodeling, it is possible that interaction with numerous host proteins such as nucleolin and nucleophosmin in an RNP complex with HDV could facilitate recruitment of RNAP I to the nucleoli to promote transcription. HDV RNA accumulation is enhanced by nucleophosmin, and antigenome synthesis is particularly sensitive to the presence of nucleolin (Huang et al., 2001; Lee et al., 1998). It thus appears that numerous host factors and polymerases associate with HDV RNA for transcription to occur, and a more comprehensive picture of HDV RNA synthesis in the cellular environment is now emerging.

Although my data show that components of the RNAP I and RNAP III transcriptional machinery bind HDV RNA, the biological significance of these interactions and capacity of these RNAPs to act as RdRps cannot be demonstrated from these observations alone. The RIPA procedure was used to show interaction of all three polymerases with both polarities of

HDV RNA in cells replicating HDV. While the RIPA is an invaluable tool for the elucidation of RNA-protein interactions, it is not without certain limitations. While RIPA analysis shows the binding of a protein to an RNA species, it does not reveal the physiological consequence of the interaction. The binding of HDV RNA to RNAP I and RNAP III might thus be a passive interaction and not result in *de novo* RNA synthesis by these polymerases. It is possible that a specific polymerase associates with HDV RNA only following synthesis by a different polymerase. Additionally, RIPA co-immunoprecipitation is performed under native conditions, and a particular RNA species co-immunoprecipitated with an antibody directed against a specific protein might actually be bound to a different component of an immunoprecipitated complex. For example, when the antibody directed against the CTD of the Rpb1 subunit of RNAP II is used for RIPA analysis, it is likely that the entire core polymerase, peripheral GTFs, and even additional transcription factors are co-immunoprecipitated, and it is possible that some of these proteins are concurrently or independently interacting with HDV RNA. As previously discussed, TBP is a common subunit of GTFs essential to all three RNAPs, and interaction of RNAP I, RNAP II, and RNAP III with HDV RNA might be occurring through a component of the RNAP complex such as TBP (Cormack and Struhl, 1992). Indeed, additional RIPA experiments showed that HDV RNA can be co-immunoprecipitated with TBP and that TBP is a component found in independent complexes containing each polymerase (Figure 3.3). In corroboration with these observations, it has recently been shown that TBP interacts with R199G by EMSA, implying that binding of HDV RNA to all three RNAPs might occur through interaction of this common subunit during initial promoter recognition (Abrahem and Pelchat, 2008). Whether or not this interaction leads to active PIC formation and transcriptional initiation, particularly in the case of RNAPs I and III, remains to be shown.

It will be interesting to determine if all three polymerases have RdRp activity and if they can participate in HDV RNA synthesis. Since TBP is common to all three polymerases and has been shown to bind HDV RNA, investigation of the TBP-HDV RNA interaction might provide insight into the manner in which RNAP I, RNAP II, and RNAP III recognize HDV RNA for non-canonical RNA-templated transcription. TBP is a small, saddle-shaped molecule that binds the minor groove of DNA (Kim et al., 1993). In the RNAP II complex, TBP is a component of the GTF TFIID, although TBP can be found in GTFs of all three eukaryotic polymerases (Cormack and Struhl, 1992). In each polymerase complex, TBP associates with unique TBP-associated factors (TAFs) which dictate the specificity of promoter elements recognized by TBP (reviewed by Gill, 1994 and Goodrich and Tijan, 1994). Although typically associated with binding the 8 nt TATA-box motif upstream from eukaryotic genes, TBP is extremely flexible in its ability to bind DNA (reviewed in Hahn, 2004). While TBP is necessary for promoter recognition and PIC formation in all three polymerases, it has been shown that only approximately 30% of genes in *Drosophila* contain the TATA-box motif, attesting to the lack of sequence specificity required for TBP binding to occur and the capacity of TAFs to affect the site of DNA binding by TBP (Gill, 1994; Ohler et al., 2002). HDV RNA does not contain a TATA box sequence, thus TBP recognition must be mediated by other sequential or structural features or involve interaction with additional GTFs and/or TAFs. One type of core promoter element, the initiator element (INR), has been shown to be recognized by TBP-containing complexes; located at or around the site of transcriptional initiation, INR sequences show a high amount of variability (reviewed by Gill, 1994). It is thus possible that HDV possesses some non-conventional sequence features that mediate TBP binding. Interestingly, the transcription factor YY1 which has been shown to interact with HDAg also binds to eukaryotic INRs; it is thus

possible that proteins such as YY1 assist in forming PICs on TATA-less HDV templates (Huang et al., 2008; Seto et al., 1991). Finally, G/C-rich helical regions upstream of many eukaryotic genes are sufficient for RNAP II recognition in the absence of TATA-box and INR elements, thus it is possible that the G/C-rich character of HDV RNA might contribute to RNAP binding through TBP (Smale, 1997). Though the present studies have shown that TBP can interact with HDV RNA and form unique complexes with all three polymerases, transcription from these polymerases cannot be concluded from this interaction since TBP binding is somewhat indiscriminate. It is possible that TBP is binding to HDV RNA and recruiting the three different polymerases to this template without causing the assembly of a functional PIC or resulting in transcription initiation. While transcription by RNAP II has been demonstrated and evidence supporting a role for RNAP I in these processes also exists, further analysis will be needed to definitively demonstrate transcription of HDV by RNAP I and RNAP III, since binding to these polymerases or their TBP subunits might not be sufficient for RNA synthesis.

Although my studies have shown that all three eukaryotic RNAPs bind the terminal stem-loop domains of both polarities of HDV RNA, differences exist in the precise regions of HDV RNA bound by these polymerases. While interaction of RNAP II with HDV RNA could be localized to 38 to 48 nt regions of the extreme tips of the HDV rod-like structure, *in vitro* binding of RNAP I and RNAP III could not be refined to these regions (Figure 3.3). A different set of transcription factors and activators is necessary for transcription from each polymerase (discussed by Cramer et al., 2008), and it is thus possible that these polymerases, although extremely homologous, bind different portions of R199G RNA. As aforementioned, TBP associates with different GTFs and TAFs in each polymerase complex, and at TATA-less promoters, as in the case of HDV RNA, these additional proteins provide sequence-

specific recognition for TBP binding (Goodrich and Tijian, 1994). If TBP is the mediator of interaction between each RNAP and HDV RNA, it would initially seem puzzling that the interaction could not be localized to precisely the same region. However, as aforementioned, TBP can bind DNA in a flexible manner, and its binding might not serve to initiate transcription but to recruit other proteins to a particular site through protein-protein interactions (discussed by Hanh, 2004). While these proteins are recruited to the convex outer surface of TBP, the concave, DNA-binding TBP surface interacts differently with nucleic acids when employed by different polymerase complexes (Burley and Roeder, 1996; Fan et al., 2004). While TBP is present in RNAP I, RNAP II, and RNAP III complexes, unique surfaces of TBP are used for DNA binding in each polymerase (Fan et al., 2005). For example, several RNA species are capable of tight interaction with TBP using distinct modes of interaction (Fan et al., 2004). While transcription by RNAP II is extremely sensitive to the presence of certain TBP-binding RNAs, transcription by RNAP I is resistant to this mode of inhibition and RNAP III-mediated transcription is only impeded if these RNAs are added during PIC formation (Fan et al., 2005). These observations attest to the specificity of the TBP interface used during PIC formation and transcriptional initiation by each of the three polymerases and could explain why all three RNAPs can bind HDV RNA without necessarily facilitating transcription. Although the present studies definitively demonstrate the capacity of all three eukaryotic RNAPs to interact with HDV RNA, the physiological relevance of the RNAP III-HDV interaction and, to a lesser extent, the RNAP I-HDV interaction requires further clarification.

6.1.3 Challenges and future directions

Ultimately, to determine whether or not RNAP I and RNAP III actually possess RdRp activity and not merely the capacity to bind HDV RNA, transcription assays such as those previously developed for RNAP II should be designed for these polymerases (Beard et al., 1996; Filipovska and Konarska, 2000). However, there are several obstacles to be overcome to develop these assays. While the Lai laboratory performed transcription assays on HDV using NE that had been immunodepleted of the RNAP I complex to demonstrate its necessity in antigenome synthesis, these data require further investigation since immunodepletion of this protein complex might have resulted in loss of other HDV-associated factors necessary for transcription (Li et al., 2006). Lack of antigenome synthesis might thus result from concurrent depletion of other nuclear factors, not necessarily RNAP I, so the necessity of RNAP I for antigenome synthesis cannot be concluded from these experiments alone. Immunodepletion of any single polymerase might concurrently remove subunits, transcription factors, or accessory proteins needed for HDV replication and transcription, leading to misinterpretation of the role of a specific polymerase in HDV RNA synthesis. Inhibitors such as α-amanitin have traditionally been used in transcription assays to differentiate between RNAPs responsible for the synthesis of a specific transcript, however α-amanitin treatment might not have the same effect on RNAP II-directed synthesis when an RNA template is used instead of DNA. For example, it has been shown that the general transcription inhibitor actinomycin D, which inhibits DNA-dependent RNA synthesis by all three RNAPs, does not prevent HDV RNA synthesis (Macnaughton et al., 2002). Since both RNAP I and RNAP III are resistant to higher doses of α-amanitin, it might also be difficult to quantitatively differentiate between synthesis by these two RNAPs, especially when considering that RNA template-driven RNA synthesis might not be subject to the same

constraints as those observed for DNA-dependent synthesis. For these reasons, it has been difficult to address the physiological relevance of the HDV-RNAP I and HDV-RNAP III interactions, and future experimentation will be necessary to elucidate whether or not these RNAPs have the capacity to act as RdRps and, if so, the physiological relevance of this activity with respect to HDV biology.

One approach to developing a transcription assay for determination of the ability of RNAP I and RNAP III to use HDV RNA as a template would involve the development of RNA aptamers that specifically inhibit each polymerase. An aptamer that specifically inhibits RNAP II while allowing RNAP I- and RNAP III-directed synthesis to proceed was developed for RNAP II of *S. cerevisiae* (Thomas et al., 1997). The RNAP II inhibitory aptamer prevents PIC formation by binding the Rpb1 and Rpb2 subunits at the active center of the RNAP II cleft, preventing template DNA from entering and transcriptional initiation from occurring (Kettenberger et al., 2006; Thomas et al., 1997). This inhibitory RNA acquires an A-helical double stem-loop structure and interacts with a separate RNAP II interface than that which binds DNA (Kettenberger et al., 2006; Thomas et al., 1997). Using this information, it might be possible to generate similar aptamers that inhibit human RNAPs I, II, and III, and these aptamers could be synthesized *in vitro* and included in conventional transcription assays prior to addition of HDV template. By inhibiting two of the three RNAPs, it would be possible to determine whether or not the remaining RNAP is capable of transcription from an HDV template. However, caution must be used in the development of these aptamers to ensure that they occlude hairpin RNA, and not only canonical template DNA, from entering the active site of the polymerase and initiating transcription. It is possible that an RNA template interacts with a different surface of RNAP II than does a DNA template; it has been reported that template RNA occupies the site normally occupied

by the DNA-RNA hybrid duplex during RNA-templated transcription, thus aptamers should be designed specifically to inhibit RNA-directed synthesis (Lehmann et al., 2007). Alternatively, aptamers that bind TBP and inhibit transcription from specific RNAPs could be used in transcription assays if TBP is determined to be essential for HDV RNA synthesis (Fan et al., 2004). Data acquired from such transcription assays would not only demonstrate a potential role for each RNAP in the HDV life cycle but also provide information regarding template requirement for RNA-templated RNA synthesis by the RNAPs. Identification of RNA promoters for RNA-dependent synthesis by RNAPs and characterization of this process should reveal insights not only into the replication of subviral pathogens but also provide evidence suggesting a novel flow of genetic information from RNA to RNA, hinting at the ancestral origin of modern day RNAPs and possible uses of their RdRp activity *in vivo*.

6.1.4. Global significance: identification of RNAPs with RdRp activity and characteristics of their RNA templates

Aside from providing novel data to the field of HDV biology, a significant implication of the present study is the potential ability of RNAP I, RNAP II, and RNAP III to initiate transcription and replication from RNA instead of the canonical DNA template. RNA synthesis by eukaryotic RNAPs is conventionally perceived to be a DNA-dependant process; this fundamental dogma of molecular biology has, for decades, provided the foundation for all nucleic acid-based research. Recently, much evidence has surfaced supporting a role for RNAPs in transcription from RNA templates, a previously heretical concept with respect to the basic tenets of biochemistry. The association of HDV RNA with these three polymerases in the present study provides insight into what might constitute an RNA promoter that can be

utilized by cellular RNAPs, adding a layer of complexity not only to the HDV life cycle but also to RNA-templated RNA synthesis by RNAPs.

In addition to HDV RNA synthesis, RdRp activity of RNAP II might explain several current developments in cellular RNA biology. A study of approximately 30% of the human genome has revealed the presence of an exceptionally high number of unannotated transcripts that do not correspond to any known genes (Cheng et al., 2005). These unconventional RNAs that neither encode proteins nor serve recognized biological functions have been termed non-coding RNAs (ncRNAs), and it is now speculated that these RNAs are the predominant type of cellular RNA. Polyadenylated, non-polyadenylated, and bimorphic transcripts are found both in the nucleus and in the cytoplasm, and it is estimated that conventional mRNAs comprise only a small percentage of all genomic transcripts (Cheng et al., 2005; Guttman et al., 2009; He et al., 2008). Unannotated transcripts that have been characterized correspond to introns, exons, and intergenic regions of the genome, and many overlap with or are complimentary to characterized mRNAs (Cheng et al., 2005; He et al., 2008; Kapranov et al., 2007). Transcription can occur from both the positive and the negative strand of genomic DNA, and transcription from the minus strand ("antisense" transcripts) typically occurs at or near its cognate "sense" transcript in a non-random, cell-specific manner (He et al., 2008). Several multi-exonic long intervening non-coding RNAs are found in intergenic regions and have been implicated in modulating transcription through chromatin remodeling and interaction with transcription factors (Guttman et al., 2009). Small (>200 nt), capped RNAs have been shown to regulate gene expression either through transcriptional interference or disruption of transcriptional events such as splicing (Affymetrix/Cold Spring Harbor Laboratory, 2009; Kapranov et al., 2007). These RNAs have also been associated with the delineation of transcriptional boundaries of protein-coding

genes and their abundance is sometimes correlated with gene expression state; their presence can thus be reflective of gene regulation as opposed to being directly involved in the process (Kapranov et al., 2007). Mirror-spliced antisense transcripts (MSATs) that are completely complimentary to known transcripts have also been discovered, and such transcripts might facilitate post-transcriptional gene regulation by interacting with their protein-coding counterparts to induce RNA silencing (Cheng et al., 2005). Since these transcripts do not all have genomic promoters or canonical splice sites, it is speculated that these novel regulatory RNAs might be produced by RdRp activity of cellular polymerases such as RNAP II. Characterization of determinants necessary for recognition of RNA promoters by RNAP II and other cellular polymerases might thus provide insight into the synthesis of ncRNAs and their subsequent regulation of gene expression.

Additional ncRNAs have been discovered in several systems and have been shown to play roles in the regulation of gene expression. As previously discussed, bacterial 6S RNA is capable of utilizing bacterial RNAP as an RdRp to facilitate the regulation of gene expression at certain stages of the bacterial life cycle (Wassarman and Saecker, 2006; Wassarman and Storz, 2000). 6S RNA interacts directly with the σ^{70} subunit of bacterial RNAP to suppress gene expression until conditions for active growth are favourable, at which point a small RNA species is synthesized from the 6S template to relieve transcriptional repression and enable outgrowth from stationary phase (Wassarman and Saecker, 2006; Wassarman and Storz, 2000). An RNA species capable of binding RNAP II in mammalian systems to mediate stress response has also been discovered (Fornace et al., 1989). Murine B2 RNA is a 178 nt ncRNA transcribed by RNAP III that is produced in response to heat shock (Espinoza et al., 2004). A 51 nt section of predominantly single-stranded RNA from B2 binds directly to RNAP II during formation of the PIC to prevent

RNA synthesis, and it has been speculated that B2 might act in a similar manner as 6S RNA (Espinoza et al., 2007). Various other endogenous cellular and viral RNAs have been shown to interact with transcription factors to modulate transcription, providing additional evidence for regulatory roles of ncRNAs and viral RNAs in transcriptional regulation (reviewed by Barrandon et al., 2007). These findings illustrate not only an increasing level of complexity for uncharacterized and newly-discovered RNAs in cellular systems but also show that ncRNAs might either result from RdRp activity of RNAP II or play roles in post-transcriptional gene regulation by employing the RdRp activity of RNAP II. Evidence obtained from studies of HDV-templated synthesis by host RNAPs should provide insight into the mechanism by which these novel regulatory pathways might be achieved.

Finally, eukaryotic RNAP transcription of HDV using an RNA template could provide insight into the origin and evolution of modern day DNA-dependent RNA polymerases. In 1986, Walter Gilbert established the concept of the "RNA World" – the theory that life originated through the activities of multifunctional RNAs and that only later in evolution was DNA adapted as the material of genetic storage and proteins as building blocks and specialized enzymes for a cellular existence (Gilbert, 1986). Following the rationale of this theory, the last universal common ancestor (LUCA) of modern organisms would have possessed an RNA genome and an RdRp from which modern RNAPs have descended (reviewed by Poole and Logan, 2005). Although sequentially divergent, these RNAPs, as well as viral RdRps, contain conserved motifs at their core that are essential for nucleic acid polymerization (Butcher et al., 2001; Delarue et al., 1990; O'Reilly and Kao, 1998). The core subunits of bacterial, archaeal, and eukaryotic RNAPs are structurally conserved, and all possess an active center containing two metal ions which are essential for polymerization (Poole and Logan, 2005). Furthermore, modern RNAPs have the capacity for

transcript proofreading and repair, a phenomenon which would be critical to the successful maintenance of a large RNA genome in the LUCA (Poole and Logan, 2005). These observations highlight the possibility that modern day RNAPs evolved from an ancestral multisubunit RdRp and that some residual RdRp capacity might be retained by modern RNAPs. By examining the manner in which modern RNAPs transcribe from RNA templates, it might be possible to obtain a better understanding of the evolutionary progenitor polymerase. My research has shown that RNAP I, RNAP II, and RNAP III can all interact with HDV RNA, potentially using them as templates for RNA-driven RNA synthesis. These preliminary data provide a foundation for investigation into whether or not RNAP I and RNAP III can use RNA as a template, whether or not this function is biologically relevant in the context of HDV biology, and whether or not the RdRp property of these RNAPs is used in endogenous cellular transcription. Studies of modern RNAP utilization of RNA templates might not only highlight their evolutionary relationship to RdRps but also reveal modern uses for RNA-templated transcription by RNAPs.

6.2 Involvement of PSF in HDV biology

In the present study, PSF was identified as a novel HDV RNA-interacting protein when it was discovered as part of a high molecular weight complex formed on HDV RNA in HeLa NE (Greco-Stewart et al., 2006; Chapter 4 and Chapter 5). In HeLa NE, PSF co-immunoprecipitates with several HDV RNA species; since PSF has numerous interacting partners in the nucleus including RNAP II, this binding could be interpreted to occur due to HDV interaction with another PSF-interacting protein in the co-immunoprecipitated complex (Emili et al., 2002; Greco-Stewart et al., 2006). To confirm the specificity of the HDV RNA-PSF interaction, purified, recombinant PSF was shown to bind PSF in the absence of other

NE proteins. Furthermore, interaction occurred rapidly and with a dissociation constant of 0.372 ± 0.031 μM, a value consistent with other biological RNA-protein interactions and those observed between viral RNAs and host proteins. RIPA analysis demonstrated the binding of PSF in a physiologically-relevant context by showing the association of PSF with both polarities of HDV RNA in HeLa cells replicating HDV. The data presented in Chapter 4 definitively demonstrate the capacity of PSF to interact with HDV RNA but do not establish a role for PSF in the HDV life cycle. Due to the abundance of PSF in the nucleus and its capacity to interact with numerous RNA and protein partners, its interaction with HDV RNA might be coincidental and not result in biological consequence with regards to HDV replication and/or pathogenicity. In order to assess the physiological relevance of the HDV-PSF interaction, further experiments, described in Chapter 5, were performed to develop a hypothesis for the role of PSF in HDV biology.

6.2.1 PSF binds viral RNAs in a sequence- and structure-dependent manner

PSF is an abundant, multifunctional protein that interacts with many types of nucleic acids. It can thus be speculated that PSF would be a good candidate as an interacting partner for various viruses and perhaps play roles in their biological activities. PSF has recently been shown to interact with HBV RNA to enhance splicing during maturation of HBV mRNA (Heise et al., 2006). Both RRM2 and the N-terminal domain of PSF are required for this process, and it is thus speculated that PSF binds HBV RNA through RRM2 and interacts simultaneously with other nuclear proteins through its N-terminal protein-protein interaction domain to form a functional splicing complex (Heise et al., 2006). These observations demonstrate how the multiple sites for protein and nucleic acid interaction contained within

PSF might facilitate its role in mediating processes necessary to viral survival (reviewed by Shav-tal and Zipori, 2002). Interestingly, it was observed that PSF recognizes and binds a post-transcriptional regulatory element and adjacent polypyrimidine tract located at the 3' splice site on HBV RNA (Heise et al., 2006). HBV is comprised of several stem-loop domains in this region and contains a polypyrimidine tract; these features are similar to those observed to bind to HDV RNA in my research (Kidd-Ljunggren et al., 2000). PSF has also been shown to interact with HIV, binding the instability elements (INSs) of its RNA genome in a structure-specific manner to repress expression of HIV mRNA (Zolotukhin et al., 2003). In contrast to its role in the HBV life cycle, PSF binds HIV mRNA to prevent gene expression and possibly facilitate nuclear retention and protein shuttling (Zolotukhin et al., 2003). The HIV genome contains several hairpin and stem-loop domains with which many host and viral proteins interact, and the specific INS element to which PSF binds is A/U-rich, corroborating the binding requirements for PSF on a viral template that I observed in the present study (Le et al., 1988; Schwartz et al., 1992).

In addition to binding viral RNAs, PSF has been shown to bind the mouse retrotransposon VL30. Retrotransposons or retroelements are ubiquitous sequences present in multiple copies per genome in higher organism and are related to tumorigenic mammalian retroviruses (reviewed by French and Norton, 1997). Binding of PSF to VL30 has been shown to occur through the RRMs of PSF, similar to what has been observed for HBV RNA binding (Heise et al., 2006; Song et al., 2004). Mouse VL30 contains a polypurine tract and is similar to retrovirus proviruses in sequence (Itin and Keshet, 1983). Song and associates (2004; 2005) propose that VL30, previously purported to be "junk" DNA, and other similar retroelements might be transcriptionally active and regulate PSF activity; by alleviating PSF-mediated inhibition of oncogene expression, VL30 could exert its effect as a tumorigenic

agent. These three examples demonstrate how PSF might serve as an important host factor in the viral life cycle and show that secondary structural features, as well as A/U-rich sequences and polypyrimidine tracts, are attractive RNA targets for PSF binding. Determination and characterization of the RNA features that facilitate the PSF-HDV interaction might help to elucidate other viral pathogens that use PSF in their life cycles in processes such as transcription, RNA processing, and localization.

My data, combined with the above observations, could represent viral recognition sequences and/or features necessary for PSF binding. Specifically, the HDV segments that bound PSF in my experiments were determined to be localized to the extreme 38 nt and 48 nt tips of the right and left genome and antigenome, respectively. Disruption of base pairing and significant alteration of the primary sequence while maintaining secondary structure both resulted in the inability of PSF to bind these RNA species. Although shown to photocross-link to polypyrimidine tracts, the optimal binding sequence for PSF was determined to be rich in purine residues (Patton et al., 1993; Peng et al., 2002). The HDV consensus sequence from 81 quasi-species shows extensive purine/pyrimidine polarization in the regions bound by PSF (Figure 2.4) and might represent sequential requirements necessary for PSF recognition.

My studies show that PSF binding is dependent on both structure and sequence since the "flip" mutants failed to interact significantly with PSF. However, the sequence requirement for PSF binding might be less stringent than the structural requirement - it is possible that the reason for the lack of interaction is due to changes in local structure. For example, the right terminal genomic domain contains a trinucleotide bulge and the left terminal antigenomic domain contains a dinucleotide bulge (Figure 1.8A and Figure 1.8B). The local structural environment comprising the stem-loop domain might be thus responsible

for the RNA-protein interaction as opposed to the sequence itself. Alteration of the side of the stem containing the bulge region or of the high purine/pyrimidine content might be responsible for abrogating the PSF-RNA interaction, making these features responsible for binding and not the exact nucleotide sequence *per se*. Additionally, it is possible that alternate structural conformations are adopted by these regions of HDV RNA *in vivo*; it is likewise possible that the predicted structures of the HDV RNA regions bound by PSF vary slightly from those predicted. Because of these variables, it difficult to conclusively define the specific RNA structure that is optimal for PSF binding. To further delineate the precise requirements for binding, co-immunoprecipitation and EMSA experiments could be repeated with *in vitro*-transcribed HDV RNAs containing more subtle mutations such as point mutations, single or dinucleotide insertions or deletions, or elimination of small proximal bulges. Nevertheless, the data acquired in the present investigation demonstrate the basic RNA requirements for PSF binding to HDV RNA, providing insight into RNA features required for PSF recognition.

6.2.2 HDV might use molecular mimicry to coerce PSF binding

In a cellular environment, PSF has been shown to bind endogenous U5 snRNA (Peng et al., 2002). The stem-loop structure of U5 snRNA is necessary for this interaction, and this structure is highly similar to those adopted by the terminal tips of the HDV genome and antigenome (Peng et al., 2002; Figure 5.3A). The similarity between the HDV terminal tips and cellular U5 snRNA bound by PSF suggests that HDV RNA might use molecular mimicry to attract cellular factors to bind its RNA. Many viral pathogens engage in molecular mimicry as a mechanism for the subversion of host cellular factors to facilitate their own replication, encapsidation, and dissemination. Viruses have demonstrated the

ability to mimic both host proteins and nucleic acids to achieve these goals and perpetuate their life cycles.

With respect to RNA mimicry, numerous viruses use this strategy to hijack host proteins involved in processes such as transcription and translation. It has been shown that viral RNA genomes have evolved to imitate those of their hosts. For example, CpG composition, which is underrepresented in the vertebrate genome, is mimicked by DNA viruses such that the CpG content of their DNA and corresponding mRNA are similar to that of their host organism (Greenbaum et al., 2008). This mimicry would imply an evolutionary pressure to assimilate the viral nucleic acids to be congruent in composition to those of their host organism. Many viruses also contain structural RNA elements in their genomes which resemble those found in host nucleic acids and use these structures to actively subvert host proteins for viral propagation. One unique strategy used by some RNA pathogens of higher vertebrates involves the use of complex RNA secondary and tertiary structures to enable translation of viral proteins. While positive sense ssRNA viruses can be used directly as mRNA, they lack the 5' methylguanine cap characteristic of the eukaryotic mRNAs ; the cap structure facilitates recognition by translation factors and delivery of the transcript to the ribosome. To circumvent this problem, viral RNAs have internal ribosome entry site (IRES) elements to enable cap-independent translational initiation. Several types of viral IRESs exist, although they all contain elaborate secondary structures and domains comprised of pseudoknot and stem-loop motifs (reviewed by Kieft, 2008). In the *Dicistroviridae* family, one such highly-structured domain of the RNA genome mimics the anticodon loop of tRNA which interacts with mRNA to facilitate binding to the ribosome decoding site to initiate translation (Costantino et al., 2008). This extremely specific structure mimics its eukaryotic counterparts so precisely that it can bind the ribosome and initiate translation in a manner

that is independent of both the 5' methylguanine cap and AUG start codon. Direct interactions between other viral IRESs and the ribosome occurring through stem-loops and other complex structures have been observed, and it might be possible that these structures mimic other endogenous RNAs to fulfill their role in accomplishing viral translation. Finally, many positive-sense ssRNA plant viruses contain tRNA-like structures in their 3' ends that can interact with aminoacyl tRNA synthetase to cause aminoacylation of their 3' ends (reviewed by Hammond et al., 2009; Pinck et al., 1970). In this manner, these viruses are capable of synthesizing proteins in the absence of ribosome recognition of a 5' methylguanine cap. This mimicry also enables interaction with host tRNA-interacting proteins such as eukaryotic translation elongation factor 1 alpha 1 (eEF1A1; Hammond et al., 2009). Remarkably, eEF1A1 was recently demonstrated to interact with HDV RNA, though the consequence of the HDV-eEF1A1 interaction is currently unknown (Sikora et al., 2009; Chapter 7.2). These examples show some of the mechanisms by which viral pathogens mimic host nucleic acids to enable their proliferation. It is thus reasonable to speculate that HDV employs a similar strategy to carry out its life cycle, especially given its limited protein-coding capacity and lack of RdRp. From these examples it is also apparent that RNA structure is critical to such mimicry, while sequence is important insofar as maintaining the structure that is requisite for the mimicry to occur.

6.2.3 PSF mediates RNAP II recruitment to HDV RNA

An interesting observation arising from my research is that at least two of the regions of HDV RNA bound by PSF are the same as those involved in RNAP II binding (regions R38G and L48AG; Figure 2.3C and Figure 5.2). Since PSF has been shown to interact directly with the CTD of RNAP II, it is possible that binding of HDV to PSF enables recruitment of

RNAP II to this RNA template (Emili et al., 2002). PSF could thus act like a transcription factor by providing a direct physical link between polymerase and template to promote RNAP II binding to a non-DNA template and initiate HDV RNA synthesis.

My preliminary experiments showed that RNAP II could be co-immunoprecipitated with the R199G HDV RNA from HeLa NE (Greco-Stewart et al., 2007; Chapter 2). However, since PSF is abundant in the NE used for these experiments as confirmed by Western blot analysis, it is possible that PSF and RNAP II are both present in co-immunoprecipitated complexes and that PSF, or another protein, is actually responsible for the apparent R199G-RNAP II interaction. However, the TBP subunit of the TFIID GTF of RNAP II has subsequently been shown to interact with HDV RNA in cells replicating HDV and to specifically bind HDV RNA *in vitro* by EMSA, assuring that components of the RNAP II transcriptional apparatus are capable of HDV RNA interaction independently of other known HDV RNA-interacting proteins (Abrahem and Pelchat, 2008; Greco-Stewart et al., 2009). It might then be pondered why PSF is needed to mediate the RNAP II-HDV interaction if components of the RNAP II complex can interact independently with HDV RNA *in vitro*. Since the nucleus is a complex microenvironment in which a myriad nucleic acid-interacting proteins and canonical DNA promoters exist with the capacity to attract GTFs and transcriptional machinery, much competition for polymerase recruitment must occur. In a physiological context, there is an abundance of DNA promoters in the nucleus which are preferred by RNAP II, thus the direct interaction between RNAP II subunits and HDV RNA might be negated by an overwhelming amount of intermolecular competition with more favourable DNA templates. It can be speculated that TBP and other components of the RNAP II complex would preferentially interact with conventional DNA templates and transcription factors, and that significant coercion must occur to promote RNAP II binding to

an RNA template. It is possible that interaction of HDV RNA with PSF is necessary for recruitment of RNAP II to an HDV template *in vivo* since PSF can interact simultaneously with RNA and protein, with the CTD of RNAP II, and with many regions of HDV RNA (Emili et al., 2002; Greco-Stewart et al., 2006; Shav-Tal and Zipori, 2002). PSF might coordinately facilitate the recruitment of HDV RNA to the RNAP II complex containing TBP, a phenomenon that might be necessary for RNAP II-HDV RNA interaction in a cellular context in which abundant preferred, non-HDV templates are available. The present studies intended not only to identify novel HDV-protein interactions but to determine their biological relevance to the HDV life cycle, thus further experiments were designed to address whether or not PSF facilitated the HDV-RNAP II interaction.

A co-immunoprecipitation approach was employed to determine the necessity of PSF for the RNAP II interaction with HDV RNA. Immunodepletion of PSF prior to co-immunoprecipitation of RNAP II was not a viable option since proteins that interact with PSF, such as RNAP II and $p54^{nrb}$, might be concurrently removed upon immunodepletion of PSF. Results observed might thus be due to removal of these putative HDV RNA-interacting proteins and not directly result from the absence of PSF. An alternative approach I considered was the selective inhibition of the PSF-RNA interaction. This would require an inhibitor that would specifically bind and occupy the RNA binding sites (i.e. RRMs) of PSF, precluding its ability to interact with HDV RNA, without affecting its ability to engage in protein-protein interactions such as the interaction with the CTD of RNAP II. A co-immunoprecipitation experiment performed with α-RNAP II in the presence of such a selective inhibitor would demonstrate whether or not the PSF-HDV RNA interaction is necessary for RNAP II to bind HDV RNA. To this end, the SELEX-derived PSF-binding small RNA species S1, S17, and $S_{consensus}$ were synthesized *in vitro* to use as selective PSF-

binding agents that could theoretically saturate the PSF RRMs and competitively occlude HDV RNA binding while leaving the N-terminal domain of PSF free to engage in protein-protein interactions (Peng et al., 2002). These three RNA species have been determined by SELEX to have a high affinity for PSF and their short sequence (20 to 30 nt) make them appealing candidates for binding to PSF without significantly altering the protein or affecting its ability to engage in protein-protein interactions.

The results of these experiments are shown in Chapter 5, Figure 5.3. When a 50-fold molar excess of each of these small RNAs were individually pre-incubated with the HeLa Ne prior to addition of radiolabeled R199G for co-immunoprecipitation, they prevented R199G from immunoprecipitating with RNAP II (Figure 5.3B). This would seemingly indicate that the PSF-HDV RNA interaction is necessary for RNAP II-HDV RNA interaction, at least in this *in vitro* system. It is possible that the interaction of HDV RNA with RNAP II observed by co-immunoprecipitation in Chapter 2 occurs as a result of a protein such as PSF interacting with RNAP II and HDV RNA simultaneously, giving the appearance of RNAP II interaction upon co-immunoprecipitation of the RNAP II-containing complex. Although high levels of purified RNAP II holoenzyme and TBP are capable of direct interaction *in vitro*, it would appear that in a more complex environment containing several nuclear proteins PSF enables, or at least enhances, HDV RNA binding to RNAP II. NE closely approximates the biologically-relevant composition of the nucleus, suggesting a role for the PSF-HDV RNA interaction in a cellular context. This data implies that PSF is required for the RNAP II-HDV RNA interaction, possibly acting as a transcription factor that physically links the polymerase to the template in NE. The proximity of RNAP II to the HDV RNA after binding to the PSF-HDV RNA complex might thus stimulate transcription from a non-DNA template.

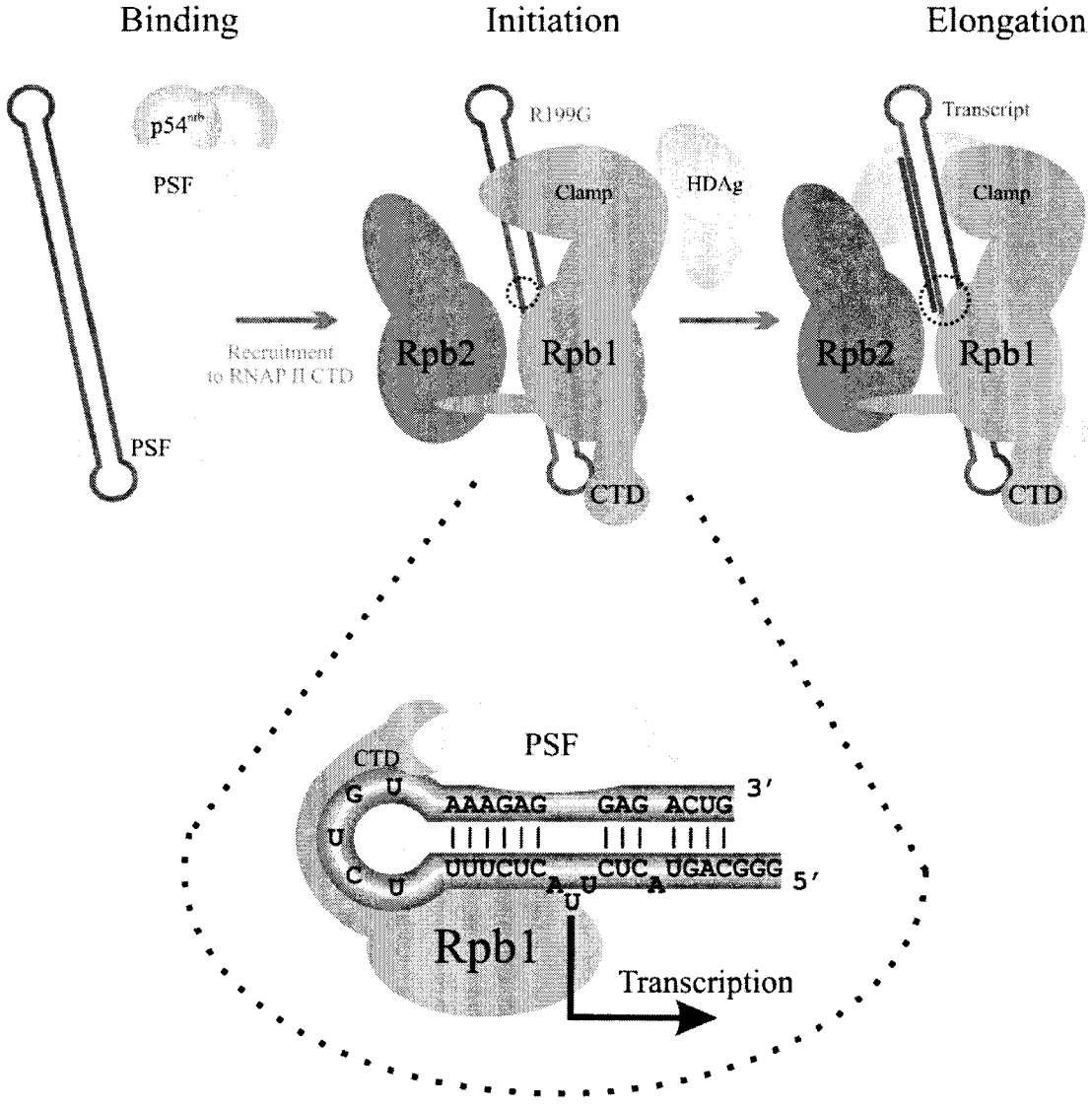
Another manner in which the necessity of PSF in the RNAP II-HDV RNA interaction was assessed was through selective competition for the RNAP II CTD-binding moiety of PSF. Theoretically, if the portion of PSF that interacts with the CTD is competitively prevented from interacting with RNAP II by introducing an abundance of CTD prior to co-immunoprecipitation, the PSF-RNAP II interaction would be abrogated and the effect on RNAP II binding to HDV RNA could be assessed. To test this theory, a peptide of approximately 100 aa containing the YSPTSPS consensus heptad repeat of RNAP II from *S. cerevisiae* was pre-incubated with HeLa NE prior to co-immunoprecipitation of RNAP II. It was shown that, upon addition of an excess of CTD, RNAP II was unable to interact with R199G. These observations would further corroborate a role for PSF in RNAP II recruitment to HDV RNA showing that both PSF binding to HDV RNA and PSF binding to RNAP II are necessary for the RNAP II-HDV RNA interaction. However, this data must be interpreted with some caution. The monoclonal antibody used in these experiments (8WG16) is directed against the CTD of RNAP II; it is possible that excess CTD polypeptide is interacting with this antibody, saturating the antibody and preventing it from binding to RNAP II in the co-immunoprecipitation experiment, giving a false negative result (i.e. that R199G was not co-immunoprecipitated due to the antibody not binding to RNAP II as opposed to PSF being saturated with CTD and being unable to bind RNAP II to mediate the RNAP II-HDV RNA interaction). To attempt to prevent this undesirable effect, the CTD peptide, as with previous inhibitors, was added prior to R199G and incubated for 5 minutes at room temperature or 1 hour at 4°C prior to co-immunoprecipitation with and α-RNAP II antibody; the same results were obtained in both circumstances. This data could be confirmed in the future by performing the experiment using an antibody directed against a different epitope of the RNAP II core protein Rpb1 that contains the CTD (such as POLR2A antibody available

from Abnova) or by using an antibody directed against a different subunit of the RNAP II holoenzyme such as Rpb4, Rpb8, or Rpb9 (POLR2D, POLR2H, and POLR2I antibodies; also available from Abnova). However, based on the data acquired from the SELEX-derived RNA and CTD competition co-immunoprecipitation experiments, it can be speculated that PSF must bind both the RNAP II CTD and HDV RNA for RNAP II to associate with HDV RNA in nuclear extracts.

To determine if PSF indeed acts as a physical bridge between RNAP II and HDV RNA, experiments were conducted to determine if both PSF and RNAP II were present simultaneously on HDV RNA. An RNA affinity chromatography strategy was used as described in Chapter 3 to determine if the complex formed on the HDV RNA species R199G contained both PSF and RNAP II, and the results are presented in Chapter 5, Figure 5.4. It was shown that PSF and RNAP II are present simultaneously on R199G, implying that PSF might mediate RNAP II recruitment to HDV RNA for transcription. The co-immunoprecipitation experiments with the SELEX-derived RNAs and purified CTD peptide demonstrate that PSF must be able to bind both HDV RNA and the RNAP II CTD for the RNAP II complex to interact with HDV RNA, and the RNA affinity chromatography further demonstrates that both PSF and RNAP II associate with HDV RNA in the heterogeneous environment of HeLa NE. While HDAg was not present in this system, a role for HDAg in recruitment of HDV RNA to RNAP II for transcriptional initiation cannot be excluded since studies have suggested its necessity *in vivo* for initiation of HDV genome synthesis and RNA accumulation (Chao et al., 1990). Based on observations from this combination of *in vitro* experimental approaches, a hypothetical model of the HDV RNA-PSF-RNAP II complex has been developed.

Figure 6.1 illustrates my working hypothetical model of PSF-mediated RNAP II recruitment to the right terminal stem-loop domain of the HDV genome for transcriptional initiation. This domain has been demonstrated to possess promoter activity for RNAP II, and binding of RNAP II has been shown to be refined to the extreme 38 nt tip of this RNA species (Beard et al., 1996; Greco-Stewart et al., 2007). In my model, PSF binds HDV RNA at locations such as the extreme right terminal stem-loop domain of the genome; in a cellular context, the possibility of association of the entire PSF/p54^{nrb} heterotetramer might occur, especially since it has recently been shown that p54^{nrb} also has the capacity to bind HDV RNA (Sikora et al., 2009; Chapter 7.2). PSF association with the CTD of RNAP II brings HDV RNA to the RNAP II transcriptional machinery for HDV RNA synthesis. While HDAg is not required for the proposed HDV RNA-PSF-RNAP II interaction *in vitro*, it is presumed that HDAg-S binds the clamp module of RNAP II *in vivo* to facilitate transcriptional elongation as it has been shown to be required for HDV RNA accumulation in HDV infected cells (Chao et al., 1990; Yamaguchi et al., 2001; 2007). While PSF and RNAP II can exist concurrently on HDV RNA, it cannot be discerned whether or not PSF remains associated with the RNAP II-HDV RNA transcriptional complex following PSF-mediated HDV RNA recruitment. PSF thus serves as the physical link between HDV RNA and RNAP II and could explain the ability of RNAP II to use an unconventional RNA template in the nuclear environment which is rich in preferred DNA templates. PSF would thus serve as a transcription factor in the HDV life cycle, providing insight into the mechanism by which HDV is capable of replication and transcription using a DNA-dependent RNA polymerase.

Figure 6.1: Working hypothetical model of HDV transcription initiation facilitated by PSF *in vivo*. PSF, alone or in a heterotetrameric complex with $p54^{nrb}$, binds HDV RNA and mediates its recruitment to the RNAP II complex by interaction with the RNAP II CTD. Since PSF interacts with both the hypo- and hyperphosphorylated form of the CTD, PSF could remain associated with RNAP II and HDV RNA during elongation. An enlarged representation of the simultaneous presence of RNAP II and PSF on HDV RNA at the right terminal hairpin domain of the HDV genome is shown.


Binding
Initiation
Elongation
p54nrb
PSF
PSF
Recruitment
to RNAP II CTD
R199G
Clamp
HDAg
Rpb2
Rpb1
CTD
Transcript
Clamp
Rpb2
Rpb1
CTD
PSF
CTD
G
U
U
C
U
AAAGAG
GAG ACUG
3′
UUUCUC
CUC UGACGGG
5′
A
U
A
U
Rpb1
Transcription

6.2.4 Challenges and future directions

Ultimately, *in vivo* studies of cells replicating HDV will be necessary to determine the precise role of PSF in the HDV life cycle. Post-transcriptional gene silencing of PSF is one approach; cells in which PSF gene expression has been reduced or halted (i.e. "knocked-down") using small interfering or small hairpin RNAs (siRNAs or shRNAs) could be used in conjunction with the HDV transfection model system developed in Chapter 2 and the effect on HDV replication, such as the ability to produce antigenome, could be determined by RT-PCR. Additionally, the amount of active transcription and replication of HDV in PSF knock-down cells could be measured using a quantitative, real-time PCR (qRT-PCR) approach such as that developed by Tseng et al. (2008) to detect changes in levels of HDV RNA synthesis. If HDV RNA synthesis is altered by PSF knock-down, RIPA analysis could show whether or not RNAP II is still capable of association with HDV RNA in PSF-deficient cells. Unfortunately, obtaining cells in which PSF expression has been reduced or halted could be problematic since PSF is multifunctional and integral to numerous cellular processes. My attempts to inhibit PSF production in HeLa cells by transfection with four different PSF-targeted shRNAs resulted in HeLa cell death; similar results have be reported following PSF knock-down in HeLa cells using traditional siRNAs (Kaneko et al., 2007). It is possible that either PSF is essential for cell viability or that certain types of interfering RNAs have appreciable off-target effects which are deleterious to cell survival. Fortunately, successful silencing of PSF expression has recently been reported in HeLa cells using an alternative form of siRNAs (Sasaki et al., 2009). Stealth siRNA (Invitrogen) employs chemical modification of the antisense strand of the siRNA duplex to prevent off-target effects caused by sense strand complimentarity to non-desired transcripts; silencing using this method was highly-efficient though incomplete, yet cell viability was not effected (Sasaki et al., 2009).

This approach could thus be used to reduce PSF expression and determine the effects on HDV RNA synthesis.

Although this strategy is appealing, it should be noted that decreasing PSF expression could cause alteration in the expression of numerous other genes due to the multifunctional nature of PSF and its role in transcriptional regulation, complicating the interpretation of data with respect to direct effect of PSF silencing on HDV RNA synthesis. It has been shown that cells expressing a low level of PSF demonstrate decreased expression of 298 genes and increased expression of 184 genes (Song et al., 2005). Any observations with regards to HDV replication obtained from such model might be a result of downstream effects and not a direct consequence of PSF reduction. Furthermore, cell viability could be adversely affected by these changes, and complete inhibition of PSF expression might prove to be a significant impediment to cell survival. It would thus be interesting to determine the effect of decreased PSF on HDV transcription and replication, but caution should be exercised when using a PSF knock-down cell system.

Another more conservative approach to the investigation of the role of PSF a cellular context would be to transfect cells with plasmids expressing mutant PSF alleles and determine the effect on HDV replication. The Blencowe laboratory has generously provided constructs containing PSF with deletions of the N-terminal domain, RGG motif, P/Q-rich region, P-rich region, RRM1, RRM2, both RRMs, and charged C-terminal residues (for PSF structure see Figure 1.5; Rosonina et al., 2005). By using our HDV infection model in cells expressing mutant PSF, it would be possible to gauge the effect of each region of PSF in HDV replication by performing qRT-PCR analysis. It would also be useful to perform these experiments in a hepatic cell line such as HepG2 or Huh7 since liver cells are traditionally the site of HDV replication and because HeLa cells contain a deletion in the DNA-binding

region of PSF (Song et al., 2005). My RIPA assay was able to show interaction of PSF with HDV RNA in HeLa cells because binding likely occurs through one or both RRM motifs, however a cell line with PSF intact would be more valuable for biologically-relevant functional interaction studies since there are multiple nucleic acid and protein partners with which PSF interacts to exert its cellular effect. These potential future lines of research should help clarify whether or not PSF is essential to or enhances HDV replication in cells infected with HDV.

6.2.5 Global significance: a role for PSF in HDV pathogenesis

Finally, elucidation of PSF as an HDV-interacting protein and proposal of its involvement in RNAP II-mediated HDV RNA synthesis might provide insight into the mechanism of HDV-mediated cellular pathogenesis. In patients co-infected with HDV, HBV disease progression is significantly exacerbated with an increased propensity for the development of hepatocellular carcinoma (Romeo et al., 2009). It has been shown that mutations of the genomic region encoding PSF are associated with increased incidence of cancer and mutations in the PSF gene have been identified in both cervical cancer and papillary renal carcinoma cell lineages (Benn et al., 2000; Clark et al., 1997; Song et al., 2005). PSF can suppress the expression of numerous cellular oncogenes and diminished cellular expression of PSF is associated with altered expression profiles of several hundred additional genes (Song et al., 2004; Song et al., 2005). PSF binds DNA upstream from cellular oncogenes to suppress their expression, however introduction of PSF-binding RNA species such as VL30 causes PSF to dissociate from the DNA upon RNA binding, allowing oncogene transcription (Song et al., 2004). The observation that PSF interacts with HDV RNA might explain the increased incidence of hepatocellular carcinoma patients with HDV; HDV RNA binds PSF

with high affinity and thus might effectively preclude PSF binding to repressor elements of oncogenes, resulting in their overexpression. This disregulation of cellular gene expression could contribute to the pathogenesis of HDV by causing the expression of several PSF-regulated oncogenes. A recently published long-term study of approximately 300 patients has shown that active HDV replication is consistently positively associated with liver failure and that severity of HDV infection is a significant predictor of the development of hepatocellular carcinoma (Romeo et al., 2009). However, pre-existing data regarding the oncogenic capacity of HDV has been conflicting. While some studies have shown that HDV is acutely cytotoxic to hepatocytes, others have suggested a more subtle growth disadvantage (Cole et al., 1991; Wang et al., 2001). Studies in insect cells have shown that cell cycle arrest occurs in the presence of HDAg, whereas in human kidney cells HDV RNA accumulation, as opposed to presence of HDAg, has been shown to cause cell cycle arrest (Chang et al., 2005; Hwang and Park, 1999). Since the complete, human patient represents the natural system in which HDV exists and causes disease, it can be speculated that a three-fold increase in hepatocellular carcinoma risk in patients with severe HDV-mediated liver injury demonstrates the potential of HDV to promote oncogenesis (Romeo et al., 2009). PSF interaction with HDV RNA might provide a potential mechanism by which HDV causes certain types of pathologies including predisposition to the development of hepatocellular carcinoma.

Due to the versatile and multifunctional nature of PSF, sequestration of this protein by HDV RNA might also result in disruption of other nuclear processes, resulting in additional pathogenic effects. PSF has been referred to as a "sticky protein" due to its ability to interact simultaneously with DNA, RNA, and protein, and it has been proposed to serve as a scaffold to link numerous aspects of nucleic acid expression, transport, and metabolism

(Shav-Tal and Zipori, 2002; Wu et al., 2006). Since PSF plays roles not only in gene expression but also in localization and cellular shuttling of nucleic acids and retention of defectively edited RNAs, it is possible that loss of available PSF due to HDV RNA binding could result in aberrant localization of cellular transcripts and release of promiscuously hyperedited RNAs (Zhang and Charmichael, 2001; Zolotukhin et al., 2003). Both of these situations could be associated with the abnormal, diseased state of HDV-infected cells. The hepatocyte tropism shown by HBV is partially explained due to enrichment of certain transcription factors requisite for the HBV lifecycle, and it would seem that HDV also usurps such proteins to survive in the cellular environment (reviewed by Lazinski, 1999). Since PSF has many natural RNA targets, the intracellular presence of HDV RNA could cause significant disregulation of essential host activities, contributing to deterioration of liver function observed in patients with chronic HDV infection. Due to its ubiquitous localization and numerous roles, PSF is definitely a protein essential for proper nuclear function and cellular viability, and it is understandable that lack of PSF activity as a result of HDV RNA-mediated sequestration could cause significant cytopathic effects.

6.3 Concluding remarks

In the present series of investigations, PSF has definitively been proven to interact with HDV RNA *in vitro* and in cells replicating HDV, and a potential role for PSF in mediating RNAP II binding has been proposed. Together, the observations documented herein identify a novel eukaryotic protein involved in HDV biology and propose a role for this protein which provides insight into the manner in which HDV RNA is capable of subverting host proteins and stimulating RNA synthesis by a DNA-dependent RNA polymerase from an RNA template. These studies also propose a mechanism by which host transcription factors and

scaffold proteins could act in the life cycle of RNA viruses, and preliminary investigations have indeed revealed that HDV can interact with other multifunctional RNA-associated proteins such as ASF and hnRNP L. These findings also suggest a manner in which HDV infection results in hepatic disease progression in persons suffering from HBV co-infection. In addition to the potential deleterious effects of PSF sequestration mentioned above, HDV utilization of PSF, RNAP II, and other cellular factors, both those described in the present study and those yet to be discovered, could result in significant disregulation of numerous nuclear and cellular processes, resulting in the disease pathology associated with HDV infection. These experiments have provided a foundation for future research not only into HDV biology but that of other RNA viruses and have proposed a manner by which RNAPs might be recruited to non-preferred templates to facilitate RNA-directed RNA synthesis.

6.4 References

Abrahem, A., Pelchat, M., 2008. Formation of an RNA polymerase II preinitiation complex on an RNA promoter derived from the hepatitis *delta* virus RNA genome. Nucleic Acids Res. 36, 5201-5211.

Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project, 2009. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature 457, 1028-1032.

Antony, T., Thomas, T., Shirahata, A., Thomas, T., 1999. Selectivity of polyamines on the stability of RNA-DNA hybrids containing phosphodiester and phosphorothioate oligodioxyribonucleotides. Biochemistry 38, 10775-10784.

Barrandon, C., Spiluttini, B., Bensaude, O., 2008. Non-coding RNAs regulate the transcriptional machinery. Biol. Cell 100, 83-95.

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Benn, D.E., Dwight, T., Richardson, A.L., Delbridge, L., Bambach, C.P., Stowasser, M., Gordon, R.D., Marsh, D.J., Robinson, B.G., 2000. Sporadic and familial pheochromocytomas are associated with loss of at least two discrete intervals on chromosome 1p. Cancer Res. 60, 7048-7051.

Branch, A.D., Lee, S.E., Neel, O.D., Robertson, H.D., 1993. Prominent polypurine and polypyrimidine tracts in plant viroids and in RNA of the human hepatitis *delta* agent. Nucleic Acids Res. 21, 3529-3535.

Burley, S.K., Roeder, R.G., 1996. Biochemistry and structural biology of transcription factor IID (TFIID). Annu. Rev. Biochem. 65, 769-799.

Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H., Stuart, D.I., 2001. A mechanism for initiating RNA-dependent RNA polymerization. Nature 410, 235-240.

Chang, J., Gudima, S.O., Tarn, C., Nie, X., Taylor, J.M., 2005. Development of a novel system to study hepatitis *delta* virus genome replication. J. Virol. 79, 8182-8188.

Chang, J., Nie, X., Chang, H.E., Han, Z., Taylor, J., 2008. Transcription of hepatitis *delta* virus RNA by RNA polymerase II. J. Virol. 82, 1118-1127.

Chao, M., Hsieh, S.Y., Taylor, J., 1990. Role of two forms of the hepatitis *delta* virus antigen: evidence for a mechanism of self-limiting genome replication. J. Virol. 64, 5066-5069.

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., Sementchenko, V., Piccolboni, A., Bekiranov, S., Bailey, D.K., Ganesh, M., Ghosh, S., Bell, I., Gerhard, D.S., Gingeras, T.R., 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science 308, 1149-1154.

Clark, J., Lu, Y.J., Sidhar, S.K., Parker, C., Gill, S., Smedley, D., Hamoudi, R., Linehan, W.M., Shipley, J., Cooper, C.S., 1997. Fusion of splicing factor genes PSF and NonO ($p54^{nrb}$) to the TFE3 gene in papillary renal cell carcinoma. 51, 2233-2239.

Cole, S.M., Gowans, E.J., Macnaughton, T.B., Hall, P.D., Burell, C.J., 1991. Direct evidence for cytotoxicity associated with expression of hepatitis *delta* virus antigen. Hepatology 13, 845-851.

Cormack, B.P., Struhl, K., 1992. The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. Cell, 69, 685-696.

Costantino, D.A., Pfingsten, J.S., Rambo, R.P., Kieft, J.S., 2008. tRNA-mRNA mimicry drives translation initiation from a viral IRES. Nat. Struct. Mol. Biol. 15, 57-64.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengle, S., Geiger, S.R., Jasiak, A.J., Jawhari, A., Jennebach, S., Kamenski, T., Kettenberger, H., Kuhn, C.D., Lehmann, E., Leike, K., Sydow, J.F., Vannini, A., 2008. Structure of eukaryotic polymerases. Annu. Rev. Biophys. 37, 337-352.

Delarue, M., Poch, O., Tordo, N., Moras, D., Argos, P., 1990. An attempt to unify the structure of polymerases. Protein Eng. 3, 461-467.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Espinoza, C.A., Allen, T.A., Hieb, A.R., Kugel, J.F., Goodrich, J.A., 2004. 11, 822-829.

Espinoza, C.A., Goodrich, J.A., Kugel, J.F., 2007. Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. RNA 13, 583-596.

Fan, X., Shi, H., Adelman, K., Lis, J.T., 2004. Probing TBP interactions in transcription initiation and reinitiation with RNA aptamers that act in distinct modes. Proc. Natl. Acad. Sci. 191, 6934-6939.

Fan, X., Shi, H., Lis, J.T., 2005. Distinct transcriptional responses of RNA polymerases I, II, and III to aptamers that bind TBP. Nucleic Acids Res. 33, 838-845.

Fels, A., Hu, K., Reisner, D., 2001. Transcription of potato spindle tuber viroid by RNA polymerase II starts predominantly at two specific sites. Nucleic Acids Res. 29, 4589-4597.

Fornace, A.J., Alamo, I., Hollander, M.C., and Lamoreaux, E., 1989. Induction of heat shock protein transcripts and B2 transcripts by various stresses in Chinease hamster cells. Exp. Cell Res. 182, 61-74.

Filipovska, J., Konarska, M., 2000. Specific HDV RNA-templated transcription by pol II *in vitro*. RNA 6, 41-54.

French, N.S., Norton, J.D., 1997. Structure and functional properties of mouse VL30 retrotransposons. Biochim. Biophys. Acta. 1352, 33-47.

Gilbert, W., 1986. Origin of life: the RNA world. Nature 319, 618.

Gildehaus, N., Neusser, T., Wurm, R., Wagner, R., 2007. Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding inhibition of *in vitro* transcription and synthesis of RNA-directed *de novo* transcripts. Nucleic Acids Res. 35, 1885-1896.

Gill, G., 1994. Transcriptional initiation. Taking the initiative. Curr. Biol. 4, 374-376.

Goodrich, J.A., Tijian, R., 1994. TBP-TAF complexes: selectivity factors for eukaryotic transcription. Curr. Opin. Cell Biol. 6, 403-409.

Greco-Stewart, V.S., Thibault, C.S., Pelchat, M., 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology 356, 35-44.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology 357, 68-78.

Greco-Stewart, V.S., Schissel, E., Pelchat, M., 2009. The hepatitis *delta* virus RNA genome interacts with the human RNA polymerases I and III. Virology 386, 12-15.

Greenbaum, B.D., Levine, A.J., Bhanot, G., Rabadan, R., 2008. Patterns of evolution and host gene mimicry in influenza and other RNA viruses. PLoS Pathog. 4, e1000079.

Gudima, S., Dingle, K., Wu, T.T., Moraleda, G., Taylor, J., 1999. Characterization of the 5' ends for polyadenylated RNAs synthesized during the replication of hepatitis *delta* virus. J. Virol. 73, 6533-6539.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., Cabili, M.N., Jaenisch, R., Mikkelsen, T.S., Jacks, T., Hacohen, N., Bernstein, B.E., Kellis, M., Regev, A., Rinn, J.L., Lander, E.S., 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 223-227.

Hahn, S., 2004. Structure and mechanism of the RNA polymerase II transcription machinery. Nat. Struct. Mol. Biol. 11, 394-403.

Hammond, J.A., Rambo, R.P., Filbin, M.E., Kieft, J.S., 2009. Comparison and functional implications of the 3D architectures of viral tRNA-like structures. RNA 15, 294-307.

Haussecker, D., Cao, D., Huang, Y., Parameswaran, P., Fire, A.Z., Kay, M.A., 2008. Capped small RNAs and MOV10 in human hepatitis *delta* virus replication. Nat. Struct. Mol. Biol. 15, 714-721.

He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., Kinzler, K.W., 2008. The antisense transcriptomes of human cells. Science 322, 1855-1857.

Heise, T., Sommer, G., Reumann, K., Meyer, I., Will, H., Schaal, H., 2006. The hepatitis B virus PRE contains a splicing regulatory element. Nucleic Acids Res. 34, 353-363.

Huang, C., Chang, S.C., Yu, I.C., Tsay, Y.G., Chang, M.F., 2007. Large hepatitis *delta* antigen is a novel clathrin adaptor-like protein. J. Virol. 81, 5985-5994.

Huang, W.H., Mai, R.T., Lee, Y.H., 2008. Transcription factor YY1 and its associated acetyltransferase CBP and p300 interact with hepatitis *delta* antigens and modulate hepatitis *delta* virus RNA replication. J. Virol. 82, 7313-7324.

Hwang, S.B., Park, K.J., 1999. Cell cycle arrest mediated by hepatitis *delta* antigen. FEBS Lett. 449, 41-44.

Itin, A., Keshet, E., 1983. Nucleotide sequence analysis of the long terminal repeat of murine virus-like DNA (VL30) and its adjacent sequences: resemblance to retrovirus proviruses. J. Virol. 47, 656-659.

Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccoloboni, A., Sementchenko, V., Tammana, H., Gingeras, T.R., 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science, 316, 1484-1488.

Kato, M., McAllister, C.J., Hokabe, S., Shimizu, N., Lyubchenko, Y.L., 2002. Structural heterogeneity of pyrimidine/purine-biased DNA sequence analyzed by atomic force microscopy. Eur. J. Biochem. 269, 3632-3636.

Kieft, J.S., 2008. Viral IRES RNA structures and ribosome interactions. Trends Biochem. Sci. 33, 274-283.

Kettenberger, H., Eisenführ, A., Brueckner, F., Theis, M., Famulok, M., Cramer, P., 2006. Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by noncoding RNAs. Nat. Struct. Mol. Biol. 13, 44-48.

Kidd-Ljunggren, K., Zuker, M., Hofacker, I.L., Kidd, A.H., 2000. The hepatitis B virus pregenome: Prediction of RNA structure and implications for the emergence of deletions. Intervirology 43, 154-164.

Kim, Y., Geiger, J.H., Hahn, S., Sigler, P.B., 1993. Crystal structure of a yeast TBP/TATA-box complex. Nature 365, 512-520.

Kaneko, S., Rozenblatt-Rosen, O., Meyerson, M., Manley, J.L., 2007. The multifunctional protein $p54^{nrb}$-PSF complex recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. Genes Dev. 21, 1779-1789.

Konopka, A., Reiter, J., Jung, M., Zarling, D., Jovin, T., 1985. Concordance of experimentally mapped or predicted Z-DNA sites with positions of selected alternating purine-pyrimidine tracts. Nucleic Acids Res. 13, 1683-1701.

Lazinski, D.W., 1999. Viral Hepatitis. In: Schaechter, M., Engleberg, N.C., Eisenstein, B.I., and Medoff, G. (Eds), Mechanisms of microbial disease, 3rd edition. Lippincott, Williams &Wilkins, Baltimore, pp. 391-402.

Le, S.Y., Chen, J.H., Braun, M.J., Gonda, M.A., Maizel, J.V., 1988. Stability of RNA stem-loop structure and distribution of non-random structure in the human immunodeficiency virus (HIV-1). Nucleic Acids Res. 16, 5153-5168.

Lee, C.H., Chang, S.C., Chen, C.J., Chang, M.F., 1998. The nucleolin binding activity of hepatitis *delta* antigen is associated with nucleolus targeting. J. Biol. Chem. 273, 7650-7656.

Lehmann, E., Brueckner, F., Cramer, P., 2007. Molecular basis of RNA-dependent RNA polymerase II activity. Nature 450, 445-449.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M.M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

Motard, J., Bolduc, F., Thompson, D., Perreault, J.P., 2008. The peach latent mosaic viroid replication initiation site is located at a universal position that appears to be definite by a conserved sequence. Virology 373, 362-375.

Murano, K., Okuwaki, M., Hisaoka, M., Nagata, K., 2008. Transcription regulation of the rRNA gene by multifunctional nucleolar protein, B23/nucleophosmin, through its histone chaperone activity. Mol. Cell. Biol. 28, 3114-3126.

Navarro, J.A., Flores, R., 2000. Characterization of the initiation sites of both polarity strands of a viroid RNA reveals a motif conserved in sequence and structure. EMBO J. 268, 218-225.

Ng, H., Kopka, M., Dickerson, R., 2000. The structure of a stable intermediate in the A to B DNA helix transition. Proc. Natl. Acad. Sci. 97, 2035-2039.

Ohler, U., Liao, G.C., Niemann, H., Rubin, G.M., 2002. Computational analysis of core promoters in the *Drosophila* genome. Genome Biol. 3, RESEARCH0087.

O'Reilly, E.K., Kao, C.C., 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. Virology 252, 287-303.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P., Nadal-Ginard, B., 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev. 7, 393-406.

Pelchat, M., Grenier, C., Perreault, J.P., 2002. Characterization of a viroid-derived RNA promoter for the DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 41, 6561-6571.

Peng, R., Dye, B.T., Perez, I., Barnard, D.C., Thompson, A.B., Patton, J.G., 2002. PSF and $p54^{nrb}$ bind a conserved stem in U5 snRNA. RNA 8, 1334-1347.

Pinck, M., Yot, P., Chapeville, F., Duranton, H.M., 1970. Enzymatic binding of valine to the 3' end of TYMV-RNA. Nature 226, 954-956.

Poole, A.M., Logan, D.T., 2005. Modern mRNA proofreading and repair: clues that the last universal common ancestor possessed an RNA genome? Mol. Biol. Evol. 22, 1444-1455.

Romeo, R., Ninno, E.D., Rumi, M., Russo, A., Sangiovanni, A., Franchis, R.D., Ronchi, G., Columbo, M., 2009. A 28-year study of the course of hepatitis *delta* infection: a risk factor for cirrhosis and hepatocellular carcinoma. Gastroenterology 136, 1629-1638.

Rosonina, E., Ip, J.Y., Calarco, J.A., Bakowski, M.A., Emili, A., McCracken, S., Tucker, P., Ingles, C.J., Blencowe, B.J., 2005. Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing *in vivo*. Mol. Cell. Biol. 25, 6734-6746.

Rickards, B., Flint, S.J., Cole, M.D., LeRoy, G., 2007. Nucleolin is required for RNA polymerase I transcription *in vivo*. Mol. Cell. Biol. 27, 937-948.

Sasaki, Y.T., Ideue, T., Sano, M., Mituyama, T., Hirose, T., 2009. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proc. Natl. Acad. Sci. 106, 2525-2530.

Schwartz, S., Campbell, M., Nasioulas, G., Harrison, J., Felber, B.K., Pavlakis, G.N., 1992. Mutational inactivation of an inhibitory sequence in human immunodeficiency virus type 1 results in Rev-independent gag expression. J. Virol. 66, 7176-7182.

Seto, E., Shi, Y., Shenk, T., 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. Nature 352, 24-245.

Shav-Tal, Y., Zipori, D., 2002. PSF and p54(nrb)/NonO--multi-functional nuclear proteins. FEBS Lett. 531, 109-114.

Sikora, D., Greco-Stewart, V., Miron, P., Pelchat, M., 2009. The hepatitis *delta* virus RNA genome interacts with eEF1A1, $p54^{nrb}$, hnRNP-L, GAPDH, and ASF/SF2. Virology 390, 71-78.

Smale, S.T., 1997. Transcription from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta. 1351, 73-88.

Song, X., Sui, A., Garen, A., 2004. Binding of mouse VL30 retrotransposon RNA to PSF protein induces genes repressed by PSF: effects on steroidogenesis and oncogenesis. Proc. Natl. Acad. Sci. 101, 621-626.

Song, X., Sun, Y., Garen, A., 2005. Roles of PSF protein and VL30 RNA in reversible gene regulation. Proc. Natl. Acad. Sci. 102, 12189-12193.

Thomas, M., Chédin, S., Carles, C., Riva, M., Famulok, M., Sentenac, A., 1997. Selective targeting and inhibition of yeast RNA polymerase II by RNA aptamers. J. Biol. Chem. 272, 27980-27986.

Ussery, D., Soumpasis, D.M., Brunak, S., Staerfeldt, H.H., Worning, P., Krogh, A., 2002. Bias of purine stretches in sequenced chromosomes. Comput. Chem. 26, 531-541.

Wang, D., Pearlberg, J., Liu, Y.T., Ganem, D., 2001. Deleterious effects of hepatitis *delta* virus replication on host cell proliferation. J. Virol. 75, 3600-3604.

Wassarman, K.M., Saecker, R.M., 2006. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. Science 314, 1601-1603.

Wassarman, K.M., Stortz, G., 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell 101, 613-623.

Wu, T.T., Netter, H.J., Lazinski, D.W., Taylor, J.M., 1997. Effects of nucleotide changes on the ability of hepatitis *delta* virus to transcribe, process, and accumulate unit-length, circular RNA. J. Virol. 71, 5408-5414.

Wu, X., Yoo, Y., Okuhama, N.N., Tucker, P.W., Liu, G., Guan, J.L., 2006. Regulation of RNA-polymerase II-dependent transcription by N-WASP and its nuclear binding partners. Nat. Cell Biol. 8, 756-763.

Yamaguchi, Y., Filipovska, J., Yano, K., Furuya, A., Inukai, N., Narita, T., Wada, T., Sugimoto, S., Konarska, M.M., Handa, H., 2001. Stimulation of RNA polymerase II elongation by hepatitis *delta* antigen. Science 293, 124-217.

Yamaguchi, Y., Mura, T., Chanarat, S., Okamoto, S., Handa, H., 2007. Hepatitis *delta* antigen binds to the clamp of RNA polymerase II and affects transcriptional fidelity. Genes Cells 12, 863-875.

Zhang, Z., Carmichael, G.G., 2001. The fate of dsRNA in the nucleus: a $p54^{nrb}$-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. Cell 106, 465-475.

Zolotukhin, A.S., Michalowski, D., Bear, J., Smulevitch, S.V., Traish, A.M., Peng, R., Patton, J., Shatsky, I.N., Felber, B.K., 2003. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Mol. Cell. Biol. 23, 6618-6630.

CHAPTER 7:
APPENDIX

7.1 RNA-templated transcription by DNA-dependent RNA polymerases

The following section elaborates on the information presented in Table 6.1 regarding the ability of different RNAPs to use RNA as a template. This text, in conjunction with the discussion of eukaryotic RNAP transcription from RNA templates, presented in Chapter 1.3, will be used to create a review publication regarding the ability of RNAPs to act on non-canonical RNA templates.

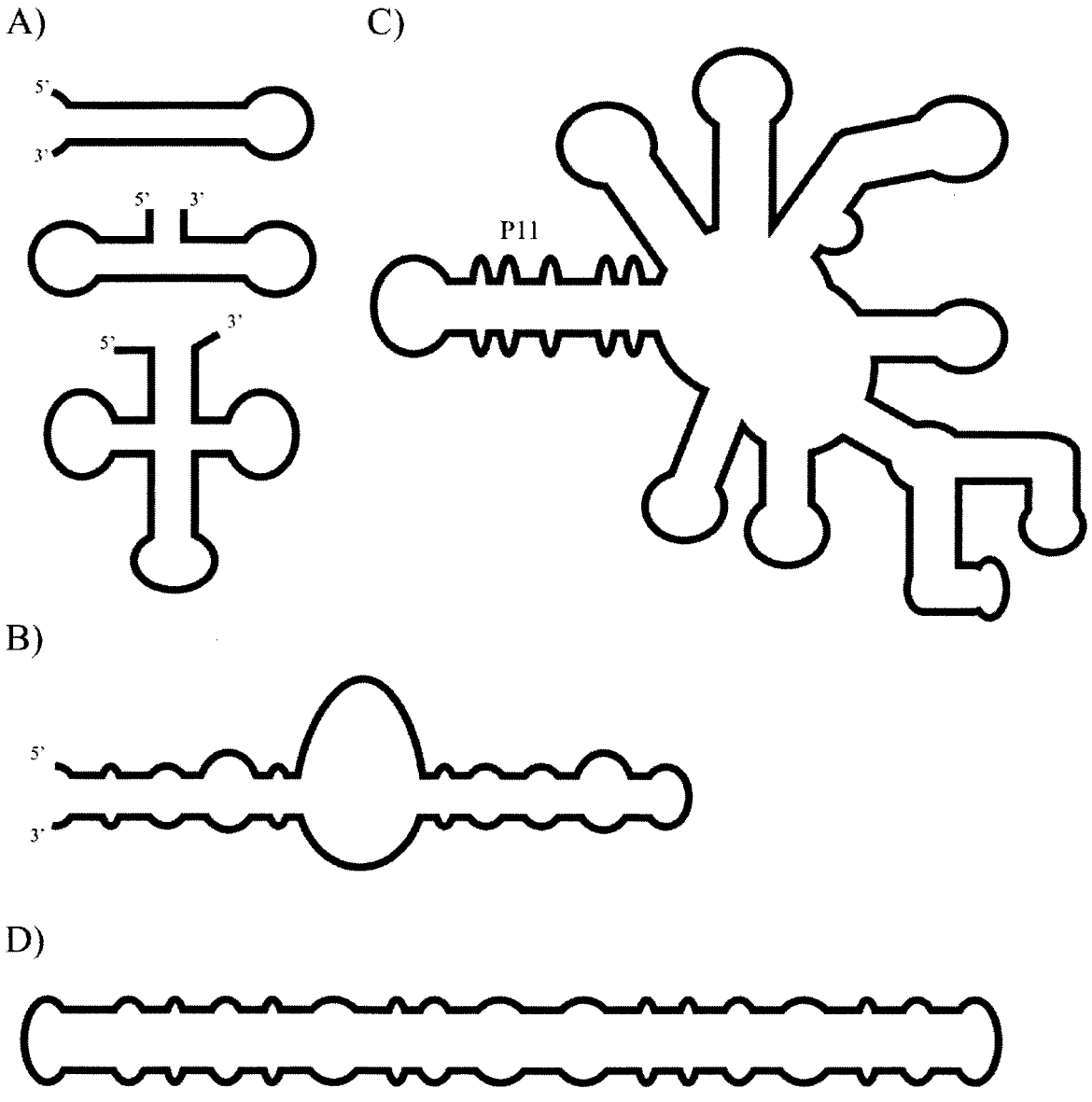
7.1.1 T7 RNA polymerase

The bacteriophage T7 possesses a DNA-dependent RNA polymerase (T7 RNA polymerase) comprised of a single polypeptide chain that efficiently transcribes from a well-characterized DNA promoter of approximately 20 nt (reviewed by Steitz, 2004). The conventional T7 promoter must be double-stranded, although the template strand may be single stranded and reused in subsequent reactions (Milligan and Uhlenbeck, 1989). Furthermore, it has been shown that the polymerase binds one face of the DNA duplex upstream from the promoter and that the promoter region must be in the double-helical conformation for recognition by T7 RNA polymerase (Schick and Martin, 1995). It has long been established that DNA-dependent RNA polymerases can transcribe from RNA templates under specific conditions *in vitro*, but with extremely limited efficiency (Chamberlin, 1974). Only recently has the capacity of T7 RNA polymerase to transcribe from a non-native RNA template been more thoroughly investigated and the precise requirements of the RNA promoter determined.

In 1989, an RNA contaminant was identified in a commercial preparation of T7 RNA polymerase that was preferentially synthesized over DNA templates during *in vitro* transcription reactions (Konarska and Sharp, 1989). Designated X RNA, this species was characterized to determine the nature of an RNA species that could be used as a template by a DNA-dependent RNA polymerase. It was found that X RNA is comprised of two complimentary strands, each of which has palindromic character and a dual axis of symmetry with a high A/U content (Konarska and Sharp, 1989). These characteristics enable X RNA to adopt numerous secondary structures including hairpins, dumbbells, and cloverleaves (Figure 7.1A). While all of these structures are present within a population of X RNA, the hairpin is the most energetically-favourable, abundant conformation. Despite great sequential heterogeneity within X RNA populations, the hairpin structure is present in all replication-competent species, indicating the necessity of this structural motif for the RNA-dependent activity of T7 RNA polymerase (Konarska and Sharp, 1990). Transcription of X RNA is resistant to inhibition by actinomycin D, a general inhibitor of DNA-dependent RNA transcription which binds DNA during initiation and prevents elongation (Konarska and Sharp, 1990; Sobell, 1985). Synthesis of X RNA is primed by pGG and pCC dinucleotides, giving rise to the nomenclature of the "G" and "C" strands (Konarska and Sharp, 1990). The G strand accumulates with a greater abundance, accounting for approximately 90% of the X RNA population. Although two strands of X RNA are synthesized, T7 RNA polymerase utilizes only single-stranded RNA as a template for RNA-dependent RNA transcription. While X RNA can be synthesized as a multimeric concatamer, a single, unit-length RNA of approximately 64 nucleotides is the favoured template for the transcription reaction.

Another species, designated Y RNA, was later discovered in another preparation of T7 RNA polymerase (Konarska and Sharp, 1990). The strand that was characterized was

Figure 7.1: Secondary structure of RNA species replicated by DNA-dependent RNA polymerases. A) Predicted conformations adopted by the C strand of X RNA including the hairpin, dumbbell, and cloverleaf (adopted from Konarska and Sharp, 1990). B) Predicted secondary structure of 6S RNA from *E. coli* (adopted from Trotochaud and Wassarman, 2005). C) Proposed secondary structure of the avsunviroid PLMVd (adopted from Pelchat et al., 2002). D) Secondary structure of the (+) strand of pospiviroid PSTVd (adapted from Kolonko et al., 2006).


A)
5'
3'
5'
3'
5'
3'
C)
P11
B)
5'
3'
D)

comparable in length to X RNA and found to adopt similar secondary structures. Interestingly, the cloverleaf conformation adopted by both X and Y RNA are reminiscent of those adopted by some plant RNA viruses which use free 3' nucleotides as initiation sites for negative strand synthesis. X and Y RNA also resemble tRNAs; the C strand of X RNA is capable of being utilized as weak substrate by bacterial tRNA nucleotidyl transferase, supporting the observation that X RNA might be recognized by bacterial proteins through molecular mimicry. Although neither X nor Y RNA is synthesized *in vivo*, *in vitro* experiments demonstrated that T3 RNA polymerase, SP6 RNA polymerase, and *E. coli* RNAP are incapable of replicating X and Y RNA, attesting to the specific requirements of T7 RNA polymerase for transcription from an RNA template and indicating that template differences for RNA-templated RNA synthesis occur among RNAPs.

Additional requirements for T7 RNA polymerase to transcribe from an RNA template were later characterized (Biebricher and Luce, 1996). It was previously discovered that the *E. coli* RNAP can select RNA species from random oligonucleotide pools for RNA-templated transcription (Biebricher and Orgel, 1973). Additionally, Qβ replicase, a viral RNA-dependent RNA polymerase, is capable of condensing NTPs into random oligonucleotides in the presence of high concentrations of replicase and NTPs; it is subsequently able to replicate these species in an RNA-dependent manner (Biebricher 1986; Sumper and Luce, 1975). Based on these findings, Beibricher and Luce (1996) investigated the ability of T7 RNA polymerase to spontaneously condense NTPs into oligonucleotides and observed the abilities of these oligonucleotides to serve as templates for RNA-dependent RNA synthesis by T7 RNA polymerase. It was demonstrated that spontaneous catalysis of NTPs into oligonucleotide polymers occurs in the presence of high concentrations of NTPs

and polymerase and, in serial reactions, each separate reaction gives rise to the emergence of one dominant RNA species of 60 to 120 nt (Beibricher and Luce, 1996). These RNAs were purified and shown to act as suitable templates for RNA-dependent synthesis by T7 RNA polymerase. While the initial lag time for RNA-templated reaction is increased as compared to that observed with a DNA template, the reaction proceeds with high fidelity, although some sequence heterogeneity occurs among each RNA population. Both double- and single-stranded RNA species were detected among the RNA products, and the abundant species contained a GG dinucleotide at the 5' terminal while the complimentary strand possessed a CC dinucleotide at this position, consistent with the termini of X and Y RNA (Konarska and Sharp, 1989; 1990). Sequencing of these RNAs revealed that they were all nearly palindromic, formed stable hairpin structures, and contained A/U-rich sequences adjacent to the hairpin (Biebricher and Luce, 1996). In competitive reactions with DNA templates, the DNA template is initially favoured over the RNA template, although at later times in reaction progression the RNA species dominates. This phenomenon is likely indicative of the preferential binding of T7 RNA polymerase to a DNA promoter, however at later times in the reaction the amount of RNA template is increased, perpetuating the RNA-dependent transcription reaction. Other polymerases, such as T3 RNA polymerase, SP6 RNA polymerase, and Qβ replicase, show limited ability to transcribe from T7 RNA polymerase-derived RNA templates, indicating that, as previously reported, each RNAP has unique requirements when acting as an RdRp.

A final series of investigations into the RdRp activity of T7 RNA polymerase were performed by Arnaud-Barbe et al. (1998) to further delineate its template requirements for RNA-dependent transcription. Using chimeric templates comprised of the T7 DNA promoter and RNA species starting 1 or 18 nt downstream from the promoter to investigate initiation

and processivity of the transcription reaction, respectively, it was found that both species are suitable for productive transcription by T7 RNA polymerase. Single-stranded, homo-duplex, and hetero-duplex RNA templates are all accepted by T7 RNA polymerase under standard reaction conditions, and the sequence of the RNA used as a template does not affect the transcription reactions. These findings indicate the ability of T7 RNA polymerase to transcribe RNA from initiation through elongation and to melt an RNA-RNA duplex during the transcription process. It also demonstrates that base composition and not specific sequence requirements could play a role in RNA-dependent RNA-transcription by T7 RNA polymerase.

Although RNA-dependent RNA polymerization by T7 RNA polymerase is unlikely to play a crucial role *in vivo* in the presence of more attractive DNA promoters, the ability of T7 RNA polymerase to reliably transcribe from an RNA template provides insight into the potential evolution of modern DNA-dependent RNA polymerases from the archaic, hypothetical "RNA world" and poses interesting questions about the ability of DNA-dependent RNA polymerases to selectively utilize certain RNA templates.

7.1.2 6S RNA: a regulator of bacterial transcription

The bacterial RNAP is a DNA-dependent RNA polymerase that has been largely characterized in species such as *E. coli*. The holoenzyme contains a conserved catalytic core comprised of two α, one β, and one β' subunit which associate with a σ specificity factor; the structures of these subunits have been elucidated and the core and active centre are highly-conserved among polymerases (Malhotra et al., 1996; Polyakov et al., 1995). In *E. coli*, seven σ factors modulate gene expression in response to environmental stimuli, including a housekeeping σ factor and alternate σ factors which are predominantly responsible for

mediating transcription of genes in response to environmental stressors such as nutrient availability and heat shock (reviewed by Gruber and Gross, 2003).

As aforementioned, studies have shown that *E. coli* RNAP is capable of spontaneously condensing NTPs and transcribing from these resulting oligomeric RNA templates (Biebricher and Orgel, 1973). Using the NTPs A, C, I, and U in the presence of Mn^{2+}, it was found that, in the absence of template, polyA/U and polyI/C are synthesized spontaneously by *E. coli* RNAP following prolonged incubation. Using random co-polymeric RNA species containing these four NTPs, it was further shown that this RNAP can replicate certain RNA species in the presence of Mn^{2+}. RNAs replicated under these conditions are highly-structured species of approximately 100 to 150 nt and are double-stranded with one strand present in excess. The population of replicated RNA species is comprised of a single species or several closely-related species. It was later shown that RNA-templated synthesis by *E. coli* RNAP can occur under near-physiological conditions using native NTP-containing co-polymers (containing G instead of I) and Mg^{2+} in random template reactions (Wettich and Biebricher; 2001). This reaction results in a population of RNA species of approximately 50 to 100 nt comprised of replication intermediates of different lengths. Reaction kinetics for this RNA-templated polymerization are similar to those observed for viral RNA replicases and RNA-templated synthesis by T7 RNA polymerase, indicating that template binding is likely the rate-limiting step since enzyme saturation increases reaction kinetics. Although heterogeneous in sequence, replicating RNAs were found to be single-stranded, A/U-rich, hairpins with 5' extensions. These species have low melting temperatures despite their high levels of intramolecular base pairing. Interestingly, these RNAs bear no similarity to known bacterial promoters, and synthesis can occur in the

absence of a σ specificity factor. While these data were intriguing, the physiological significance in *E. coli* biology remained unknown.

In 1967, an abundant RNA species was observed in *E. coli* total RNA preparations and was later determined to be a component of RNP complexes (Hindley, 1967; Lee et. al, 1978). This first non-coding bacterial RNA, named 6S RNA, was sequenced in 1971, although its function remained enigmatic until recently when it was discovered to be a regulator of *E. coli* RNAP activity (Brownlee, 1971; Wassarman and Stortz, 2000). 6S RNA binds directly to *E. coli* RNAP, specifically making contact with the β/β' subunits of the core polymerase and σ^{70}, the main housekeeping σ factor in *E. coli* (Gildehaus et al., 2007; Wassarman and Stortz, 2000). 6S RNA increases almost 10-fold during the stationary phase of bacterial growth and is speculated to mediate the transition from stationary to exponential phase growth (Wassarman and Stortz, 2000). Furthermore, 6S RNA has been shown to repress transcription from many, but not all, σ^{70} promoters, indicating an increased level of complexity to 6S-mediated regulation (reviewed by Wassarman, 2007). Although 6S RNA does not bind to free σ subunits, some binding to alternate σ factors such as σ^{S} has been observed; these data are controversial since 6S seems to activate σ^{S} promoters *in vivo* but inhibit their transcription *in vitro* (Gildehaus et al., 2007; Trotochaud and Wassarman, 2004). Together, these data demonstrate a biological activity for an ncRNA species in bacterial growth regulation through direct interaction with *E. coli* RNAP, providing insight into possible roles for small RNAs in RNAP-mediated transcriptional regulation.

6S RNA homologues have been identified in over 100 bacterial species including γ-proteobacteria, β-proteobacteria, and Gram-positive species (reviewed by Wassarman, 2007). Although they lack significant sequential homology, bioinformatic and biochemical approaches have shown that all of these RNAs of about 180 to 200 nt are single-stranded and

adopt a complex secondary structure in which the RNA is folded into a rod-like structure containing several small internal bulges, a large central "bubble" region, and a variable number of smaller bulges flanking the central bubble (Trotochaud and Wassarman, 2005; Figure 7.1B). This bubble is highly reminiscent of an *E. coli* DNA promoter open complex, which alludes to the fact that 6S might be able to compete for RNAP binding by DNA mimicry (Gildehaus et al., 2007; Wassarman and Saeker, 2006). Furthermore, alteration of the central bubble by reducing the amount of single-stranded RNA on either the top or bottom of the bubble results in inability of the mutant 6S RNA to bind σ^{70} and repress transcription (Trotochaud and Wassarman, 2005). Other mutations including strand inversion of the single-stranded bubble or sequence alteration have only minor effects on transcriptional repression, indicating some flexibility in the content of the bulge but demonstrating that its presence is crucial for 6S function (Trotochaud and Wassarman, 2005).

Perhaps the most intriguing discovery in 6S biology is the manner by which it has recently been proposed that 6S RNA regulates bacterial gene transcription. When bound to *E. coli* RNAP, the central bubble of 6S RNA has been shown to be situated in the active site of the polymerase where it competitively prevents DNA promoter binding (Gildehaus et al., 2007; Wassarman and Saeker, 2006). Promoters containing a weak -35 element or an extended -10 element are preferentially bound by 6S while promoters containing the consensus -35 element are 6S-resistant (Cavanagh et al., 2008). While engaged in the active site, 6S RNA directs the *de novo* transcription of a 14 to 22 nt RNA product, sometimes referred to as pRNA, in the absence of DNA (Gildehaus et al., 2007; Wassarman and Saeker, 2006). Transcription starts at the U_{44} position and proceeds optimally at high NTP concentrations. pRNA synthesis occurs similarly in RNAP complexes containing different σ

subunits and in the presence of the core polymerase alone. The pRNA transcript is complimentary to the 20 to 40 nt region of 6S RNA, a region that is highly conserved among 6S RNAs from different bacterial species (Wassarman and Saeker, 2006). Furthermore, mutant 6S RNA species in which the top portion of the central bubble has been removed are unable to direct synthesis of pRNA. High concentrations of 6S RNA result in an increase in short pRNA production, while high polymerase concentrations favour increased transcription of pRNA products of many lengths (Gildehaus et al., 2007). Evidence suggests that the σ^{70} subunit is released during transcription, followed by the dissociation of the 6S RNA-pRNA complex from the rest of the polymerase, possibly due to the inability of the bubble region to "reanneal" during the transition from transcription initiation to elongation (Wassarman and Saeker, 2006). The current model of 6S regulation thus speculates that transcription of pRNA occurs during bacterial outgrowth from stationary phase upon nutrient acquisition and subsequent increase in available nucleotide pools, resulting in disengagement of the 6S RNA-pRNA complex, enabling the polymerase to transcribe genes required for bacterial growth. 6S levels drop following outgrowth, possibly due to accessibility of this RNA to RNAses upon release from the RNAP complex. These data reveal a previously-undiscovered method of gene regulation in bacteria involving RNA-dependent RNA transcription and provide a physiological role for the RdRp activity of RNAPs.

Although future research will be necessary to fully characterize the precise mechanism of 6S-mediated transcriptional regulation, a crucial biological function can definitively be assigned to this bacterial ncRNA, providing a basis for future study into this complex, RNA synthesis-dependent regulatory mechanism in bacteria. While 6S RNA is the only known RNA species transcribed by bacterial RNAP *in vivo*, evidence for additional

capabilities of RNAPs to undergo RNA-templated transcription have also been discovered with RNA from a unique type of plant pathogen - the viroid.

7.1.3 Viroids: a diverse pantheon of RNAP usurpers

Viroids are an assorted group of plant pathogens that are responsible for significant agricultural and economic loss. They consist of a small (approximately 250 to 400 nt), circular RNA genome and do not contain any open reading frames, making their mechanism of pathogenesis somewhat enigmatic (Davies et al., 1974; Hall et al., 1974). Viroids can be categorized into two major taxa: *Avsunviroidae* and *Pospiviroidae* (Flores et al., 1998). Avsunviroids adopt intricate secondary structures containing many stem-loop domains and bulges and contain hammerhead ribozyme motifs which are essential for their replication (recently reviewed by Daròs et al., 2006 and Tsagris et al, 2008). They replicate in plant chloroplasts by a symmetrical rolling circle mechanism (Branch and Robertson, 1984; Diener et al., 1996). Pospiviroids assume long, rod-shaped structures containing a conserved central region and replicate in the nucleus by an asymmetrical rolling circle mechanism which relies heavily on host enzymes (Daròs et al., 2006 and Tsagris et al., 2008). The most abundant viroid RNA species in infected cells is designated the (+) polarity by convention, while the less-abundant replication intermediate, which occurs as linear multimers in pospiviroids, is referred to as the (-) polarity (Branch and Robertson, 1984). Due to their distinct compartmentalization upon infection, avsunviroids and pospiviroids rely on different host polymerases for their replication and propagation.

Chloroplasts contain at least two DNA-dependent RNA polymerases: a nuclear-encoded phage-like RNAP (nuclear-encoded polymerase; NEP) and a plastid-encoded multisubunit RNAP similar to the bacterial RNAP (plastid-encoded polymerase; PEP; Sakai

et al., 2004). Members of the avsunviroid group can thus be speculated to require one or both of these polymerases for their replication. Evidence for involvement of NEP in replication of the avocado sunblotch viroid (ASBVd) comes principally from the observation that transcription of ASBVd is resistant to inhibition by tagetitoxin, a compound that inhibits transcription by bacterial RNAPs but not phage RNA polymerases (Navarro et al., 2000). The predicted structure of the 247 nt genome of ASBVd contains a right terminal stem-loop domain, an internal region with much base pairing and a small bulge, and a left terminal domain containing two branched, stem-loop domains; it was determined that transcription initiation sites for both polarities of ASBVd are located in the right terminal stem-loop domains (Navarro et al., 2000; Symons, 1981). Both initiation sites are located at U residues flanked by A/U-rich regions, reminiscent of initiation sites found in DNA-dependent RNA promoters from bacteriophages and those used by T7 bacteriophage polymerase when transcribing from a non-canonical RNA template (Navarro et al., 2000). Although naturally-occurring variants of ASBVd demonstrate heterogeneity in this region, the A/U-rich character of this domain is preserved in all infectious species, indicating the necessity of these residues in promoter recognition and transcriptional initiation (Pallás et al., 1988; Rakowski and Symons, 1989). Additionally, positive-stranded plant RNA viruses and satellites contain terminal hairpins which are essential for promoter activity, supporting a role for the hairpin structure of these right terminal domains in ASBVd replication (Carpenter and Simon, 1998; Lauber et al., 1997). It is thus implied that both sequence composition and secondary structure could play a role in RNA promoter recognition by RNAPs. Although these data are suggestive of RNA features essential for chloroplastic viroid replication, conflicting evidence regarding promoter elements and polymerase

utilization in avsunviroid RNA synthesis has emerged through studies of the peach latent mosaic viroid (PLMVd).

PLMVd consists of an RNA genome of 336 to 351 nt and adopts a complex secondary structure including a long, left terminal stem-loop domain containing the hammerhead ribozyme, several smaller branching stem-loop domains, and a complex pseudoknot (Bussière et al., 2000; Fekih Hassen et al., 2007; Figure 7.1C). Although in the same family as ASBVd, PLMVd differs significantly in G+C content, hammerhead ribozyme motif, and secondary structural morphology (Hernández and Flores, 1992). It was shown that PLMVd is capable of initiating transcription *in vitro* using a commercial preparation of *E. coli* RNAP holoenzyme, supporting a role for PEP in PLMVd RNA synthesis (Pelchat et al., 2001). Elements contained in left-terminal hairpin domain encompassing the hammerhead ribozyme motif alone are sufficient to initiate transcription from a PLMVd template. Furthermore, *in vitro* transcription is resistant to tagetitoxin, indicating that this inhibitor of bacterial RNAP might not inhibit transcription from an RNA template as it does from a DNA template. Using *in vitro*-transcribed segments of the PLMVd genome, it was further shown that the region responsible for initiating transcription is located within a 60 to 75 nt double-stranded segment of the left-terminal stem-loop domain denoted the P11 stem (Pelchat et al., 2002; Figure 7.1C). Specifically, a minimal loop of 6 nt adjacent to a base-paired uridine and unstable stem in the P11 stem are necessary for transcription. The σ^{70} subunit of RNAP is essential for transcription of PLMVd, suggesting that the semi-stable nature of the stem could represent the portion of this "transcription bubble" that must be melted for initiation to occur. Transcription from this putative promoter is less susceptible to intercalating agents such as ethidium bromide than are DNA promoters, indicating

potential differences between bacterial RNAP utilization of DNA and RNA templates (Pelchat et al., 2004). Finally, the β/β' subunits of *E. coli* RNAP have been shown to bind specifically to the P11 stem of PLMVd RNA, providing convincing evidence that PEP is the enzyme responsible for PLMVd transcription and proposing this region of the genome as a putative RNA promoter for an RNAP.

Recently, another set of data has emerged, further characterizing the initiation site of PLMVd and proposing that NEP is the polymerase involved in PLMVd replication. The site of initiation of transcription for both polarities of PLMVd was refined to a region within 6-7 nt at the base of the P11 stem domain characterized by a conserved GUC triplet upstream from the hammerhead ribozyme cleavage site (Delgado et al., 2005). The initiating nucleotides were determined to be C_{51} and U_{286} for the (+) and (-) strands, respectively, which are located only a few nucleotides upstream from the hammerhead ribozyme cleavage site. It was proposed that NEP was the enzyme responsible for transcription due to the previous observations of ASBVd RNA synthesis, the ability of PLMVd to replicate efficiently in leaves with impaired PEP synthesis, and the observation that NEPs exhibit some plasticity in promoter requirements as demonstrated by their ability to transcribe many chloroplastic genes in PEP-deficient tobacco plants (Delgado et al., 2005; Krause et al., 2000). It is thus unclear which polymerase, NEP or PEP, is principally involved in PLMVd RNA synthesis or whether both enzymes play a role in PLMVd replication in infected chloroplasts.

Evidence confirming the existence of a universal PLMVd initiation site on each polarity of the PLMVd genome has recently emerged (Motard et al., 2008). These sites were localized to the A_{50}/C_{51} and U_{284} positions of the left terminal P11 stem region of the (+) and

(-) strands, respectively. These nucleotides were determined to be the first nucleotides present in RNA-templated transcripts due to their possession of a 5' triphosphate cap since it has been previously reported that chloroplastic transcripts are 5'triphosphorylated (Meng et al., 1991; Motard et al., 2008). An *in vitro* SELEX experiment using *E. coli* RNAP was performed to elucidate the precise requirements for the initiation site and, after 11 rounds of selection, the final pool of RNAs transcribed by this RNAP were shown to contain a conserved CAGACG box reminiscent of another reported single-stranded DNA promoter for *E. coli* RNAP (Motard et al, 2008; Ohmichi et al., 2002). This sequence also corresponds to the PLMVd consensus sequence for the P11 stem region conserved among PLMVd isolates (Motard et al., 2008). A rolling circle replication assay using *E. coli* RNAP further showed that this site is essential for RNA-templated replication by this polymerase. These data clearly delineate the PLMVd initiation site and its surrounding sequence and suggest the potential for replication of PLMVd by the PEP enzyme, although the capacity of NEP to act on an RNA template cannot be excluded. Replication of avsunviroids thus remains controversial with respect to which chloroplastic polymerase is responsible for replication of these small RNA pathogens.

Unlike members of the *Avsunviroidae* taxa, members of the *Pospiviroidae* family of viroids, including PSTVd, replicate in the nucleus of infected cells. Pospiviroids are thus thought to use the multisubunit eukaryotic RNAP II for their replication (reviewed by Darós et al., 2006 and Diener et al., 1996). PSTVd is comprised of a circular RNA molecule of approximately 359 nt and acquires a rod-like structure due to its large proportion of base-paired residues (Gross et al., 1978; Figure 7.1D). Many lines of evidence have emerged supporting a role for RNAP II in PSTVd replication including the observations that replication of PSTVd is sensitive to α-amanitin *in vivo* and that PSTVd can be transcribed *in*

vitro by both tomato and wheat germ nuclear extracts containing RNAP II (Mühlbach and Sänger, 1979; Rackwitz et al., 1981). Interestingly, transcription of PSTVd is not inhibited by actinomycin D, a general inhibitor of DNA-dependent RNA synthesis, implying that RNA-templated transcription might not proceed in the same manner as that templated by DNA (Schindler and Mühlbach, 1992). Transcription in the presence of low doses of α-amanitin has also been reported, suggesting basal transcription of PSTVd might occur from other host RNAPs or that RNA-templated synthesis by RNAP II might differ from DNA-templated synthesis (Kolonko et al., 2006). Direct binding of wheat germ RNAP II to PSTVd was demonstrated both by analytical ultracentrifugation and electron microscopy, further substantiating the involvement of RNAP II in PSTVd transcription (Goodman et al., 1984). RNAP II binds PSTVd RNA significantly more tightly than non-related ssRNA species and binding is localized to the termini of the rod-like PSTVd genome. Together, these data demonstrate the ability of RNAP II to recognize, bind, and initiate transcription from an RNA template and strongly suggest that this polymerase is involved in pospiviroid replication.

Characteristics of the putative promoters for PSTVd RNA synthesis emerged from experiments involving *in vitro* transcription from (+) strand, circular PSTVd RNA isolated from infected tomato plants (Fels et al., 2001). Start sites of for (-) strand synthesis were localized to both terminal hairpin domains of the (+) polarity of the PSTVd genome, consistent with previous observations that RNAP II binding is localized to these regions (Fels et al., 2001; Goodman et al., 1984). It was initially determined that at least two start sites located at positions A_{111} of the right terminal domain and A_{325} of the left terminal domain exist, although more recent data suggests that a sole start site is located at position C_1/G_{359} of the left terminal hairpin domain while A_{111} and A_{325} serve to enhance transcription

(Fels et al., 2001; Kolonko et al., 2006). Sequence analysis of PSTVd showed that this viroid contains neither TATA-boxes nor INR elements, sequences commonly required for transcription from eukaryotic promoters (Fels et al., 2001). However, the terminal domains of PSTVd contain G/C-rich helical regions which are conserved among PSTVd isolates and have been proposed to be necessary for synthesis of the (+) strand (Fels et al., 2001; Loss et al., 1991; Qu et al., 1993). Mutants in which this G/C-rich structure has been disrupted are unstable and quickly revert to wild-type sequence (Kolonko et al., 2006). G/C-rich helical regions are common features of TATA-less promoters of numerous eukaryotic housekeeping genes and could explain the ability of RNAP II to initiate transcription from the unconventional PSTVd template (Roeder, 1996; Smale, 1997). RNAP II has also been shown to transcribe from genes lacking both TATA-boxes and INR elements, supporting the hypothesis that the G/C-rich helical region might provide a structural basis for transcription initiation for PSTVd (Novina and Roy, 1996). The start site of PSTVd RNA synthesis might thus be determined by the structure of this region and not the specific nucleotide sequence (Kolonko et al., 2006). These data suggest the necessity of the hairpin structure for RNAP II transcription of PSTVd, corroborating similar reports of the hairpin structural requirement for RNA template utilization by RNAPs.

Another line of evidence supporting a role for RNAP II in the replication of *Pospiviroidae* family members emerges from examination of the citrus exocortis viroid (CEVd). CEVd is a 371 nt single-stranded, circular viroid that is polypurine-rich and adopts a hairpin secondary structure reminiscent of PSTVd (Gross et al., 1982). Both CEVd and the related coconut cadang cadang viroid (CCCVd) have been shown to localize to the nucleoplasm in infected plant cells, implicating a role for host polymerases such as RNAP II in their life cycles (Bonfiglioli et al., 1994). RNAP II was specifically suspected to be the

polymerase involved in CEVd replication since CEVd RNA synthesis is sensitive to α-amanitin-mediated inhibition (Semancik and Harper, 1984). Furthermore, both (+) and (-) sense CEVd species are present in RNAP II-containing complexes co-immunoprecipitated from nuclear extract isolated from CEVd-infected tomatoes (Warrilow and Symons, 1999). The RNA bound by RNAP II comprises only a small proportion of total viroid RNA present in the extract, and the (-) strand is 4.5-fold more abundant, implying that the bound species are comprised mainly of template and nascent transcript RNA. Furthermore, the hyperphosphorylated form of the RNAP II CTD is present, indicating that not only RNA binding but also transcriptional elongation is occurring in these complexes. These data provide additional evidence that RNAP II is involved in pospiviroid RNA synthesis, although the specific RNA features required for recognition of CEVd and CCCVd have not been explored.

The present body of evidence strongly suggests a role for RNAP II in the nuclear transcription of *Pospiviroidae* family members. Furthermore, it can be ascertained that the rod-shaped secondary structure of these viroids is responsible for template recognition by RNAP II, and that structural rather than sequential determinants are required for recruitment of RNAP II to a non-DNA template. Due to recent discoveries regarding the ability of RNAP II to act as an RdRp on short A-helical RNA templates, it is reasonable to assume that this polymerase might play a role in pospiviroid replication. However, the involvement of other plant polymerases including RNAP I, RNAP III, RNAP IVa, RNAP IVb, and RdRp have yet to be examined in the context of viroid replication, and a much more complex story of the subviral life cycle involving multiple host polymerases is emerging from studies of the hepatitis *delta* virus and its mammalian host.

7.1.4 References

Arnaud-Barbe, N., Cheynet-Sauvion, V., Oriol, G., Mandrand, B., Mallet, F., 1998. Transcription of RNA templates by T7 RNA polymerase. Nucleic Acids Res. 26, 3550-3554.

Bonfiglioli, R.G., McFadden, G.I., Symons, R.H., 1994. *In situ* hybridization localizes avocado sunblotch viroid on chloroplast thylakoid membranes and coconut cadang cadang viroid in the nucleus. Plant J. 6, 99-103.

Biebricher, C.K., Luce, R., 1996. Template-free generation of RNA species that replicate with bacteriophage T7 RNA polymerase. EMBO J., 15, 3458-3465.

Biebricher, C.K., Orgel, L.E., 1973. An RNA that multiplies indefinitely with DNA-dependent RNA polymerase: selection from a random copolymer. Proc. Natl. Acad. Sci. 70, 934-938.

Biebricher, C.K., Eigen, M., Luce, R., 1986. Template-free RNA synthesis by Q beta replicase. Nature 321, 89-91.

Branch, A.D., Robertson, H.D., 1984. A replication cycle for viroids and other small infectious RNAs. Science 223, 450-455.

Brownlee, G.G., 1971. Sequence of 6S RNA of *E. coli*. Nat. New Biol. 229, 147-149.

Bussière, F., Ouellet, J., Côté, F., Lévesque, D., Perreault, J.P., 2000. Mapping in solution shows the peach latent mosaic viroid to possess a new pseudoknot in a complex, branched secondary structure. J. Virol. 74, 2647-2654.

Cavanagh, A.T., Klocko, A.D., Liu, X., Wassarman, K.M., 2008. Promoter specificity for 6S regulation of transcription is determined by core promoter sequences and competition for region 4.2 of sigma70. Mol. Microbiol. 67, 1242-1256.

Carpenter, C.D., Simon, A.E., 1998. Analysis of sequences and predicted secondary structures required for viral satellite RNA accumulation *in vivo* genetic selection. Nucleic Acids Res. 26, 2426-2432.

Chamberlin, M.J., 1974. The selectivity of transcription. Annu. Rev. Biochem. 43, 721-775.

Daròs, J.A., Elena, S.F., Flores, R., 2006. Viroids: an Ariadne's thread into the RNA labyrinth. EMBO Rep. 7, 593-598.

Davies, J.W., Kaesberg, P., Diener, T.O., 1974. Potato spindle tuber viroid XII. An investigation of viroid RNA as a messenger for protein synthesis. Virology 61, 281-286.

Delgado, S., Martínez de Alba, A.E., Hernández, C., Flores, R., 2005. A short double-stranded RNA motif of peach latent mosaic viroid contains the initiation and the self-cleavage sites of both polarity strands. J. Virol. 79, 12934-12943.

Diener, T.O., 1996. Understanding replication mechanisms in viroids and viroidlike- RNAs. Trends Microbiol. 4, 85-87.

Fekih Hassen, I., Massart, S., Motard, J., Roussel, S., Parisi, O., Kummert, J., Fakhfakh, H., Marrakchi, M., Perreault, J.P., Jijakli, M.H., 2007. Molecular features of new peach latent mosaic viroid variants suggest that recombination may have contributed to the evolution of this infectious RNA. Virology. 360, 50-57.

Fels, A., Hu, K., Reisner, D., 2001. Transcription of potato spindle tuber viroid by RNA polymerase II starts predominantly at two specific sites. Nucleic Acids Res. 29, 4589-4597.

Flores, R., Randles, J. Bar-Joseph, M., Diener, T., 1998. A proposed scheme for viroid classification and nomenclature. Arch. Virol. 143, 623-629.

Gildehaus, N., Neusser, T., Wurm, R., Wagner, R., 2007. Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding inhibition of *in vitro*

transcription and synthesis of RNA-directed *de novo* transcripts. Nucleic Acids Res. 35, 1885-1896.

Goodman, T.C., Nagel, L., Rappold, W., Klotz, G., Reisner, D., 1984. Viroid replication: equilibrium association constant and comparative activity measurements for the viroid-polymerase interaction. Nucleic Acids Res. 12, 6231-6246.

Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H., Sänger, H.L., 1978. Nucleotide sequence and secondary structure of potato spindle tuber viroid. Nature 273, 203-208.

Gross, H.J., Krupp, G., Domdey, H., Raba, M., Jank, P., Lossow, C., Alberty, H., Ramm, K., Sänger, H.L., 1982. Nucleotide sequence and secondary structure of citrus exocortis and chrysanthemum stunt viroid. Eur. J. Biochem. 121, 249-257.

Gruber, T.M., Gross, C.A., 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. 57, 441-466.

Hall, T.C., Wepprich, R.K., Davies, J.W., Weathers, L.G., Semancik, J.S., 1974. Functional distinctions between the ribonucleic acids from citrus exocortis viroid and plant viruses: cell-free translation and aminoacylation reactions. Virology 61, 486-492.

Hernández, C., Flores, R., 1992. Plus and minus RNAs of peach latent mosaic viroid self-cleave *in vitro* via hammerhead structures. Proc. Natl. Acad. Sci. 89, 3711-3715.

Hindley, J., 1967. Fractionation of 32-P-labeled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. J. Mol. Biol. 30, 125-136.

Kolonko, N., Bannach, O., Aschermann, K., Hu, K.H., Moors, M., Schmitz, M., Steger, G., Riesner, D., 2006. Transcription of potato spindle tuber viroid by RNA polymerase II starts in the left terminal loop. Virology 347, 392-404.

Konarska, M.M., Sharp, P.A., 1989. Replication of RNA by the DNA-dependent RNA polymerase of phage T7. Cell 57, 423-431.

Konarska, M.M., Sharp, P.A., 1990. Structure of RNAs replicated by the DNA-dependent T7 RNA polymerase. Cell 63, 609-618.

Krause, K., Maier, R.M., Kofer, W., Krupinska, K., Herrmann, R.G., 2000. Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. Mol. Gen. Genet. 263, 1022-1030.

Lauber, E., Guilley, H., Richards, K., Jonard, G., Gilmer, D., 1997. Conformation of the 3'-end of beet necrotic yellow vein benevirus RNA 3 analyzed by chemical and enzymatic probing and mutagenesis. Nucleic Acids Res. 25, 4723-4729.

Lee, S.Y., Bailey, S.C., Apirion, D., 1978. Small stable RNAs from *Escherichia coli*: evidence for the existence of new molecules and for a new ribonucleoprotein particle containing 6S RNA. J. Bacteriol. 133, 1015-1023.

Lehmann, E., Brueckner, F., Cramer, P., 2007. Molecular basis of RNA-dependent RNA polymerase II activity. Nature 450, 445-449.

Loss, P., Schmitz, M., Steger, G., Riesner, D., 1991. Formation of a thermodynamically metastable structure containing hairpin II is critical for infectivity of potato spindle tuber viroid RNA. EMBO J., 10, 719-727.

Malhotra, A., Severinova, E., Darst, S.A., 1996. Crystal structure of sigma70 subunit fragment from *E. coli* RNA polymerase. Cell 87, 127-136.

Meng, B.Y., Wakasugi, T., Sugiura, M., 1991. Two promoters within the psbK-psbI-trnG gene cluster in tobacco chloroplast DNA. Curr. Genet. 20, 259-264.

Milligan, J.F., Uhlenbeck, O.C., 1989. Synthesis of small RNAs using T7 RNA polymerase. Methods Enzymol. 180, 51-62.

Motard, J., Bolduc, F., Thompson, D., Perreault, J.P., 2008. The peach latent mosaic viroid replication initiation site is located at a universal position that appears to be defined by a conserved sequence. Virology 373, 362-375.

Mühlbach, H.P., Sänger, H.L., 1979. Viroid replication is inhibited by α-amanitin. Nature 278, 185-188.

Navarro, J.A., Flores, R., 2000. Characterization of the initiation sites of both polarity strands of a viroid RNA reveals a motif conserved in sequence and structure. EMBO J. 268, 218-225.

Novina, C.D., Roy, A.L., 1996. Core promoters and transcriptional control. Trends Genet. 12, 351-355.

Ohmichi, T., Maki, A., Kool, E.T., 2002. Efficient bacterial transcription of DNA nanocircle vectors with optimized single-stranded promoters. Proc. Natl. Acad. Sci. 99, 54-59.

Pallás, V., García-Luque, I., Dominigo, E., Flores, R., 1988. Sequence variability in avocado sunblotch viroid (ASBVd). Nucleic Acids Res. 26, 9864.

Pelchat, M., Côté, F., Perreault, J.P., 2001. Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol. 146, 1753-1763.

Pelchat, M., Grenier, C., Perreault, J.P., 2002. Characterization of a viroid-derived RNA promoter for the DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 41, 6561-6571.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636-642.

Polyakov, A., Severinova, E., Darst, S.A., 1995. Three-dimensional structure of *Escherichia coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. 83, 365-373.

Qu, F., Heinrich, C., Loss, P., Steger, G., Tien, P., Riesner, D., 1993. Multiple pathways of reversion in viroids for conservation of structural elements. EMBO J., 12, 2129-2139.

Rackwitz, H.R., Rohde, W., Sänger, H.L., 1981. DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies. Nature 291, 297-301.

Rakowski, A.G., Symons, R.H., 1989. Comparative sequence studies of variants of avocado sunblotch viroid. Virology. 173, 352-356.

Roeder, R.G., 1996. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. 21, 327-335.

Sakai, A., Takano, H., Kuroiwa, T., 2004. Organelle nuclei in higher plants: structure, composition, function, and evolution. Int. Rev. Cytol. 238, 59-118.

Schick, C., Martin, C.T., 1995. Tests of a model of specific contacts in T7 RNA polymerase-promoter interactions. Biochemistry 34, 666-672.

Schindler, I.M., Mühlbach, H.P., 1992. Involvement of nuclear DNA-dependent RNA polymerases in potato spindle tuber viroid replication: a re-evaluation. Plant Sci. 84, 221-229.

Semancik, J.S., Harper, K.L., 1984. Optimal conditions for cell-free synthesis of citrus exocortis viroid and the question of specificity of RNA polymerase activity. Proc. Natl. Acad. Sci. 81, 4429-4433.

Smale, S.T., 1997. Transcription from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta. 1351, 73-88.

Sobell, H., 1985. Actinomycin and DNA transcription. Proc. Natl. Acad. Sci. 82, 5328-5331.

Steitz, T.A., 2004. The structural basis of the transition from initiation to elongation phases of transcription, as well as translocation and strand separation, by T7 RNA polymerase. 14, 4-9.

Sumper, M., Luce, R., 1975. Evidence for *de novo* production of self-replicating and environmentally adapted RNA structures by bacteriophage Qbeta replicase. Proc. Natl. Acad. Sci. 72, 162-166.

Symons, R.H., 1981. Avocado sunblotch viroid: primary sequence and proposed secondary structure. Nucleic Acids Res. 9, 6527-6537.

Trotochaud, A.E., Wassarman, K.M., 2004. 6S RNA function enhances long-term cell survival. J. Bacteriol. 186, 4978-4985.

Trotochaud, A.E., Wassarman, K.M., 2005. A highly conserved 6S RNA structure is required for regulation of transcription. Nat. Struct. Mol. Biol. 12, 313-319.

Tsagris, E.M. Martinez de Alba, A.E. Gozmanova, M., Kalantidis, K., 2008. Viroids. Cell. Microbiol. 10, 2168-2179.

Warrilow, D., Symons, R.H., 1999. Citrus exocortis viroid RNA is associated with the largest subunit of RNA polymerase II in tomato *in vivo*. Arch. Virol. 144, 2367-2375.

Wassarman, K.M., 2007. 6S RNA: a regulator of transcription. Mol. Microbiol. 65, 1425-1431.

Wassarman, K.M., Saecker, R.M., 2006. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. Science 314, 1601-1603.

Wassarman, K.M., Stortz, G., 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell 101, 613-623.

Wettich, A., Biebricher, C.K., 2001. RNA species that replicate with DNA-dependent RNA polymerase from *Escherichia coli*. Biochemistry 40, 3308-3315.

7.2 The hepatitis *delta* virus RNA genome interacts with eEF1A1, p54^{nrb}, hnRNP-l, GAPDH, and ASF/SF2.

Dorota Sikora, Valerie S. Greco-Stewart, Paul Miron and Martin Pelchat

7.2.1 Statement of contribution

The manuscript entitled "The hepatitis *delta* virus RNA genome interacts with eEF1A1, p54^{nrb}, hnRNP-l, GAPDH, and ASF/SF2" was published in Virology in 2009 (390, 71-78). I contributed the data presented in Figure 7.2.6 and wrote the sections of the text pertinent to these experiments. Since these data were obtained using a technology acquired during the course of my doctorate studies but were not related to my core research objectives, this paper has been included as an appendix; the text is presented as it was upon submission with additional formatting and minor revisions.

7.2.2 Abstract

Because of its extremely limited coding capacity, the hepatitis *delta* virus (HDV) takes over cellular machineries for its replication and propagation. Despite the functional importance of host factors in both HDV biology and pathogenesis, little is known about proteins that associate with its RNA genome. Here, we report the identification of several host proteins interacting with an RNA corresponding to the right terminal stem-loop domain of HDV genomic RNA using mass spectrometry on a UV cross-linked ribonucleoprotein complex, RNA affinity chromatography, and screening of a library of purified RNA-binding proteins. Co-immunoprecipitation was used to confirm the interactions of eEF1A1, p54^{nrb}, hnRNP-L, GAPDH, and ASF/SF2 with the right terminal stem-loop domain of HDV genomic RNA *in vitro* and with both polarities of HDV RNA within HeLa cells. Our discovery that HDV

RNA associates with proteins involved in RNA processing pathways and translation machinery during its replication provides new insights into HDV biology and its pathogenicity.

Keywords: Hepatitis *delta* virus; RNA virus; alternative splicing factor ASF/SF2; heterogeneous nuclear ribonucleoprotein L; hnRNP-L; glyceraldehydes 3-phosphate dehydrogenase; GAPDH; $p54^{nrb}$; eukaryotic elongation factor 1A1; eEF1A1.

7.2.3 Introduction

The hepatitis *delta* virus (HDV) is the smallest known human RNA pathogen and requires the human hepatitis B virus (HBV) envelope proteins for virion production and transmission. HDV consists of a small (~1,680 nucleotides, nt), single-stranded, circular RNA molecule, which adopts an unbranched, rod-like structure due to a high degree of intramolecular complementarity (for a review see Taylor, 2006). Its genome is proposed to replicate via a double rolling circle mechanism in which multiple-length genomic and antigenomic linear transcripts are processed into unit-length RNAs by intrinsic *delta* ribozymes. These linear RNAs are then ligated to produce unit-length circular molecules. The HDV genome contains a single open reading frame encoding two viral proteins (HDAgs). These two proteins are mostly identical in sequence except that the large HDAg (HDAg-L) contains 19 additional amino acids at its C-terminus resulting from post-transcriptional RNA editing of the termination codon of the small HDAg (HDAg-S) gene. Each protein has a distinct function: HDAg-S (195 amino acids) is essential for HDV replication, while HDAg-L (214 amino acids) is necessary for virion assembly and is reported to be a dominant negative inhibitor of replication (Taylor, 2006).

The pathogenicity of HDV is poorly understood, and liver transplantation remains the only viable option for treatment of advanced HDV liver disease (Farci, 2003). Notwithstanding its apparent simplicity, HDV causes a severe, rapidly-progressive form of liver disease which culminates in chronic liver cirrhosis and is associated with the development of hepatocellular carcinoma (Fattovich et al., 2004; Romeo et al., 2009; Su et al., 2006; Taylor, 2006). The expression of the hepatitis *delta* antigen has been suggested to result in significant cytotoxic changes in HeLa and HepG2 cells (Cole et al., 1991). However, HDV RNA replication, rather than expression of the *delta* antigen alone, might be responsible for cytopathic effects (Wang et al., 2001). Because interaction of host factors with the HDV RNA genome during viral replication might adversely affect their normal cellular activities, identification of these factors is required to understand HDV pathogenicity.

Many studies using cultured cells and cell extracts have implicated RNAP II in HDV replication (Abrahem and Pelchat, 2008; Chang et al., 2008; Filipovska and Konarska, 2000; Greco-Stewart et al., 2007; Moraleda and Taylor, 2001). It has also been hypothesized that RNAP I, or an RNAP I-like polymerase, might be involved in HDV replication because some antigenomic HDV RNA has been observed in the nucleolus, HDV RNA synthesis was reported to be associated with the RNAP I-specific transcription factor SL1, and HDV RNA accumulation was resistant to a higher dose of the RNAP II inhibitor α-amanitin (Li et al., 2006; Macnaughton et al., 2002; Modahl et al., 2000). Recently, we further implicated RNAP I in HDV replication and also suggested a role for RNAP III in this process by reporting the association of both RNAP I and RNAP III with HDV RNA (Greco-Stewart et al., 2009). Another enzyme involved in the life cycle of HDV is the small isoform of adenosine deaminase acting on RNA (ADAR-1) which is believed to catalyze the editing

reaction that leads to the production of the large *delta* antigen (Wong and Lazinski, 2002). Other cellular proteins reported to interact with HDV RNA include the cellular double-stranded RNA-dependent protein kinase (Circle et al., 1997), the negative elongation factor (Yamaguchi et al., 2002), glyceraldehyde 3-phosphate dehydrogenase (Lin et al., 2000), and the polypyrimidine tract-binding (PTB) protein-associated splicing factor (PSF; Greco-Stewart et al., 2006). Finally, HDV RNA was shown to interact with the splicing factor SC35 and to transiently co-localize with it in nuclear speckles (Bichko and Taylor, 1996; Abrahem and Pelchat, 2008).

In this study, we used several experimental approaches to identify nuclear proteins that associate with the HDV RNA genome, specifically with an HDV-derived RNA that we and others have reported to act as an RNA promoter *in vitro* (Abrahem and Pelchat, 2008; Beard et al., 1996; Greco-Stewart et al., 2007). First, we revisited the mass spectrometry analysis of a ribonucleoprotein complex produced by UV cross-linking of the HDV-derived RNA fragment with HeLa nuclear extract (NE; Greco-Stewart et al., 2006). Next, we used an RNA affinity purification strategy followed by mass spectrometric analysis. Finally, we performed electrophoretic mobility shift assays (EMSAs) using a number of purified RNA-binding proteins. Using these approaches, we identified four new HDV-interacting cellular proteins (eEF1A1, $p54^{nrb}$, hnRNP-L, and ASF/SF2) and corroborated a previously-known HDV-interacting protein (GAPDH). Together, this information increases our knowledge of HDV biology and serves as a foundation to study HDV pathogenicity.

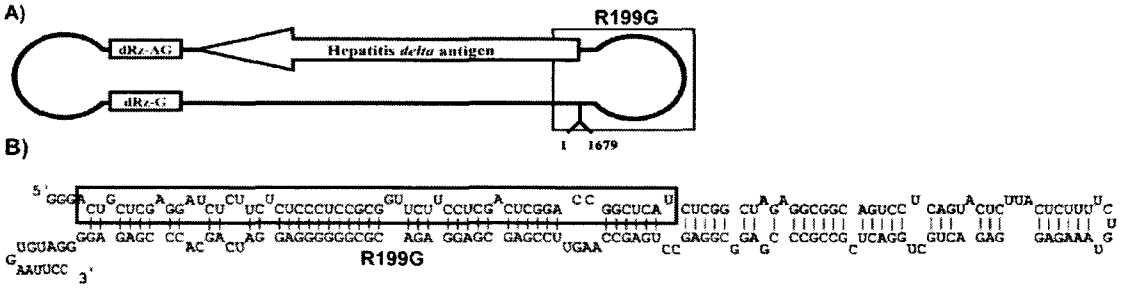
7.2.4 Results

Identification of HeLa NE proteins involved in the formation of a UV cross-linked complex with the right terminal stem-loop domain of HDV genomic RNA.

To identify human proteins that interact with the HDV RNA genome, and thus might be important for HDV biology, we used as bait a 199-nt HDV-derived RNA fragment corresponding to what we refer to as the right terminal stem-loop domain of the genomic polarity (R199G; Figure 7.2.1). We selected this RNA segment because it contains the proposed initiation site for HDAg mRNA transcription (Gudima et al., 2000), it was shown to direct synthesis of antigenomic RNA in an *in vitro* transcription reaction (Abrahem and Pelchat, 2008; Beard et al., 1996), and we have reported the specific binding of RNAP II to R199G (Abrahem and Pelchat, 2008; Greco-Stewart et al., 2007). All these features suggest that this part of the HDV genome includes an RNA promoter for RNAP II which might be involved in viral replication and/or transcription.

Recently, we reported the formation of a single, high molecular weight ribonucleoprotein complex following UV cross-linking of HeLa NE proteins to R199G (Greco-Stewart et al., 2006). Although preliminary mass spectrometry analysis of this complex led to the identification of PSF as a novel HDV RNA-interacting protein (Greco-Stewart et al., 2006), the presence of other proteins in the complex was likely. To identify additional proteins binding to R199G, we revisited the cross-linking experiment. R199G was synthesized *in vitro* in the presence of [α-^{32}P] GTP, incubated with HeLa NE proteins, and the ribonucleoprotein complexes were UV cross-linked at 365 nm. Following UV irradiation, the mixtures were fractionated by SDS-PAGE and the complex was detected by autoradiography. In agreement with our previous report (Greco-Stewart et al., 2006), a specific band migrating as a species of approximately 250 kDa was detected when the

Figure 7.2.1: Schematic representation of the HDV RNA genome and secondary structure of the right terminal region of the genomic polarity (R199G). R199G RNA fragment was *in vitro*-synthesized and used as bait in the identification of HDV RNA interacting proteins. The proposed initiation site for HDAg mRNA transcription and the region binding both RNAP II and PSF are indicated (Gudima et al., 2000; Greco-Stewart et al., 2006; Abrahem and Pelchat, 2008).


A)
R199G
dRz-AG
Hepatitis delta antigen
dRz-G
1
1679
B)
R199G

radiolabeled R199G was incubated with HeLa NE proteins and UV cross-linked (Figure 7.2.2). To test the specificity of the complex, we challenged its formation with P11.60, which is a small RNA species that folds into a hairpin (Pelchat and Perreault, 2004), thus providing an excellent competitive RNA to test for non-specific affinity of the proteins to double-stranded RNA or stem-loop secondary structures. Challenges of the complex with a molecular excess of this non-specific competitor did not significantly alter its formation. However, complex formation was greatly reduced by the addition of excess of non-radiolabeled R199G RNA, and no complex was observed in the absence of NE or UV irradiation. Taken together, these observations demonstrate that NE protein(s) interact specifically with R199G.

The complexes from three independent experiments were excised and analyzed by mass spectrometry (LC-MS/MS, Wemb Biochem) as previously described (Greco-Stewart et al., 2006). As a negative control, UV irradiation was performed on samples without R199G and subjected to the same treatment. The protonated masses of those peptides were then used to search human protein sequence databases. To be cautious, only the proteins for which at least two peptides were identified in the three samples, and which were not found the absence of R199G, were considered. A list of the putative HDV-interacting proteins is presented in Table 7.1. In accordance with our previous reports (Abrahem and Pelchat, 2008; Greco-Stewart et al., 2006; 2007) peptides corresponding to both the largest subunit of RNAP II and PSF were detected as constituents of the complex. In addition, numerous peptides corresponding to proteins involved in a wide range of cellular processes were detected in the sample containing R199G, suggesting their interaction with the right terminal stem-loop region of genomic HDV RNA. Among those were peptides specific to the mRNA splicing factor $p54^{nrb}$, the eukaryotic translation elongation factor 1 alpha 1 (eEF1A1), the F-

Figure 7.2.2: Detection of a specific UV cross-linked complex between R199G and HeLa nuclear extract proteins. Radiolabeled R199G was incubated with HeLa NE and subjected to UV irradiation at a wavelength of 365 nm. Complexes were resolved on SDS-PAGE. Specificity of the interaction was verified by addition of 50X molar excess of either a specific (R199G) or non-specific (P11.60) RNA competitor. The asterisk indicates the excised complex.

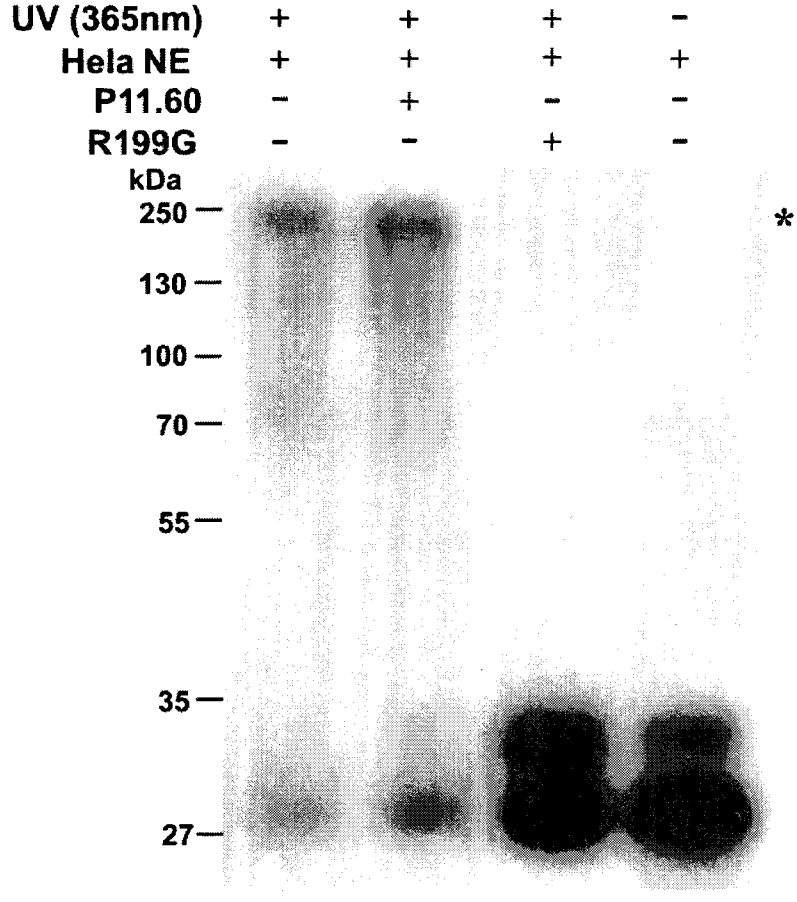

UV (365nm) + + + -
Hela NE + + + +
P11.60 - + - -
R199G - - + -
kDa
250
130
100
70
55
35
27
*

Table 7.1: Proteins identified by LC-MS/MS following analysis of a ribonucleoprotein complex obtained from a UV crosslinking experiment between HDV-derived RNA (R199G) and HeLa nuclear extract proteins.

Protein Identified	Biological Function	Accession No.	Peptides Identified	MW (kDa)
Eukaryotic translation elongation factor 1A1	Translation	4503471	4	50
Polypyrimidine tract binding protein-associated splicing factor (PSF)	Pre-mRNA processing	33879558	3	75
Nuclear mitotic apparatus protein 1	Mitotic spindle stabilization	71361682	3	238
$p54^{nrb}$	Pre-mRNA processing	543010	2	54
F-box and leucine-rich repeat protein 17	Protein-ubiquitin ligase complex	45238580	2	34
ANKS6 protein	Unknown function	39963545	2	50
DNA-directed RNA polymerase II polypeptide A	Transcription	4505939	2	217
Kalirin/RhoGEF kinase isoform 3	Guanine nucleotide exchange factor	68362740	2	144

box and leucine-rich repeat protein 17, the isoform 3 of kalirin/RhoGEF kinase, and the nuclear mitotic apparatus protein 1.

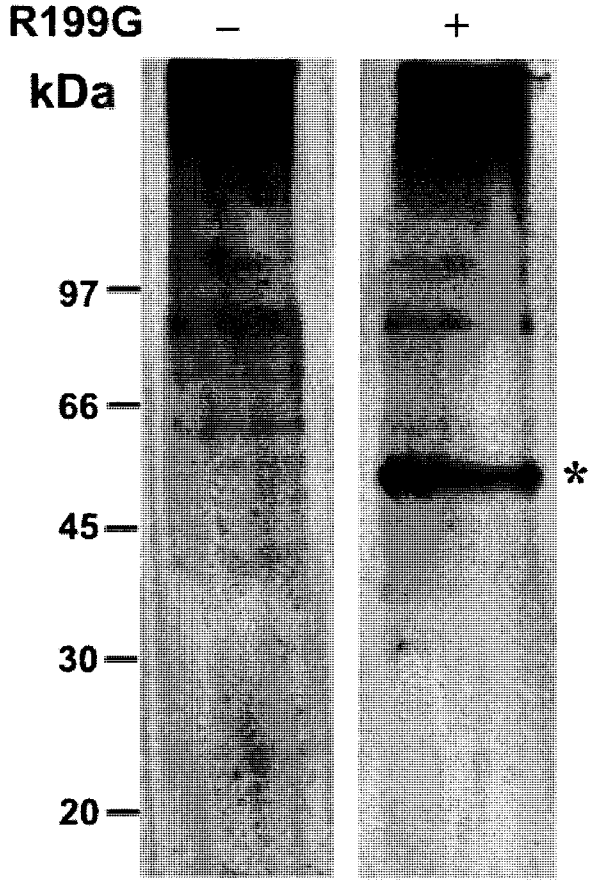
Purification of the splicing factor hnRNP-L using R199G affinity chromatography.

As an alternative to the UV cross-linking approach, we used an RNA affinity chromatography purification procedure that involves chemical cross-linking of R199G to adipic acid dehydrazide agarose beads (Vioque and Altman, 1986). R199G was synthesized *in vitro* and oxidized prior to being covalently linked to the beads. Nucleic acid-coupled beads were then incubated with pre-cleared HeLa NE. To assess the specificity of the complexes, an excess amount of yeast total tRNA was added to the extract. In addition, uncoupled adipic acid dehydrazide agarose beads were tested for protein interaction to rule out non-specific binding of NE proteins to the beads. Following incubation and washing to remove unbound proteins, the proteins were eluted with SDS loading dye and resolved on SDS-PAGE. Silver staining of the gel revealed a band corresponding to a protein of approximately 55 kDa (Figure 7.2.3). This band was clearly absent from the negative control, which consisted of uncoupled beads. The band was excised from the gel, in-gel trypsin digested, and analyzed by mass spectrometry (LC-MS/MS; Wemb Biochem). By searching human protein sequence databases, peptides matching the heterogeneous nuclear ribonucleoprotein L (60 kDa; hnRNP-L) were detected in the sample.

Identification of the alternative splicing factor ASF/SF2 as a R199G binding protein by screening a library of proteins.

HDV RNA was previously found to associate with the splicing factor SC35 and to co-localize with it in nuclear speckles (Abrahem and Pelchat, 2008; Bichko and Taylor, 1996).

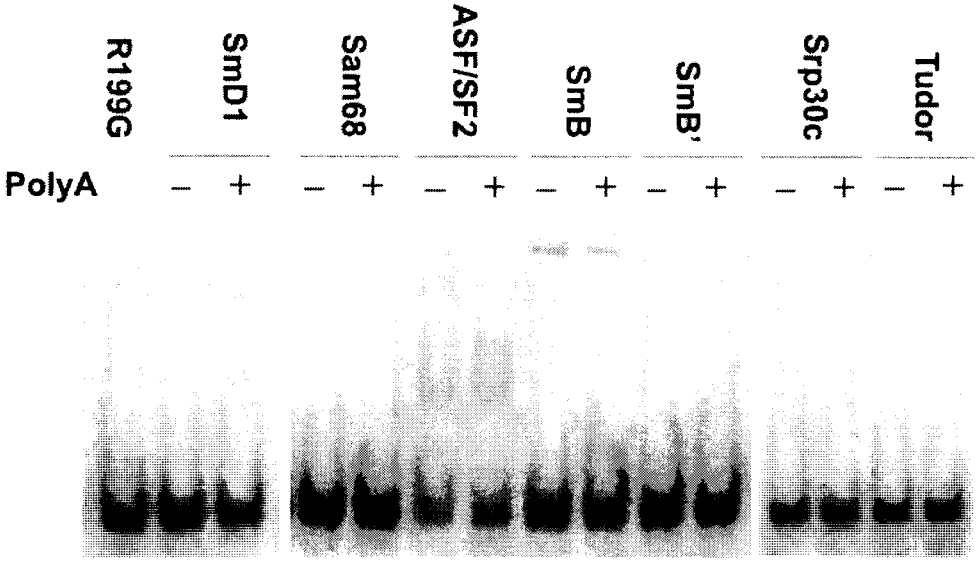
Figure 7.2.3: Interaction of hnRNP-L with R199G RNA. RNA-coupled beads were incubated with HeLa nuclear extract. The bound proteins were eluted with SDS loading dye and resolved by SDS-PAGE. The gel was silver-stained and the band corresponding to a protein of approximately 55 kDa (asterisk), and absent from the negative control, was excised and identified by mass spectrometry as hnRNP-L.


R199G
–
+
kDa
97
66
45
30
20
*

Our experiments led to the identification of three other pre-mRNA splicing factors as potential R199G-interacting partners, namely $p54^{nrb}$, hnRNP-L, and PSF (recently reported to interact specifically and directly with HDV RNA; Greco-Stewart et al., 2006). Accordingly, several purified recombinant proteins known to be involved in RNA metabolism, specifically during splicing events, were screened for their interaction with R199G by EMSA (Figure 7.2.4). We tested the Sm proteins B/B'and D1, which are ribonucleoproteins that collectively make up the Sm core of snRNPs (small nuclear ribonucleoproteins), the serine/arginine (SR)-rich proteins ASF/SF2 and SRp30, which play essential roles in constitutive as well as regulated splicing (Hastings and Krainer, 2001), Sam68, a member of the signal transduction and activation of RNA (STAR) family of proteins implicated in alternative splicing (Hartmann et al., 1999), and the Tudor domain, which is found in several splicing factors and RNA-binding proteins (Ponting, 1997).

Internally-radiolabeled R199G was synthesized *in vitro* and was allowed to bind to the purified recombinant proteins. To ensure specificity of the interactions, the samples were incubated either in the presence or absence of poly(A), a non-specific competitor. Following a 30-minute incubation at 4°C, the samples were subjected to electrophoresis through a non-denaturing 5% polyacrylamide gel. Under our conditions, retardation of the migration of R199G was observed with ASF/SF2, both in the presence and absence of molecular excess of poly(A), indicating the specific binding of this purified protein to the HDV-derived RNA molecule. Similar experiments showed that ASF/SF2 does not bind to P11.60 (data not shown), confirming that the binding of ASF/SF2 to R199G was not due to non-specific affinity to double-stranded RNA or stem-loop secondary structures. Finally, we observed an apparent shift of the radiolabeled R199G in the presence of SmB. However, because the

Figure 7.2.4: Direct interaction of R199G with GST-ASF/SF2 *in vitro*. Electrophoretic gel shift assays (EMSAs) were performed to analyze interaction of the genomic HDV RNA promoter with several recombinant human proteins. Radiolabeled R199G was incubated with each protein in the presence or absence of poly(A) (100ng).


R199G
SmD1
Sam68
ASF/SF2
SmB
SmB'
Srp30c
Tudor
PolyA
– + – + – + – + – + – + – +

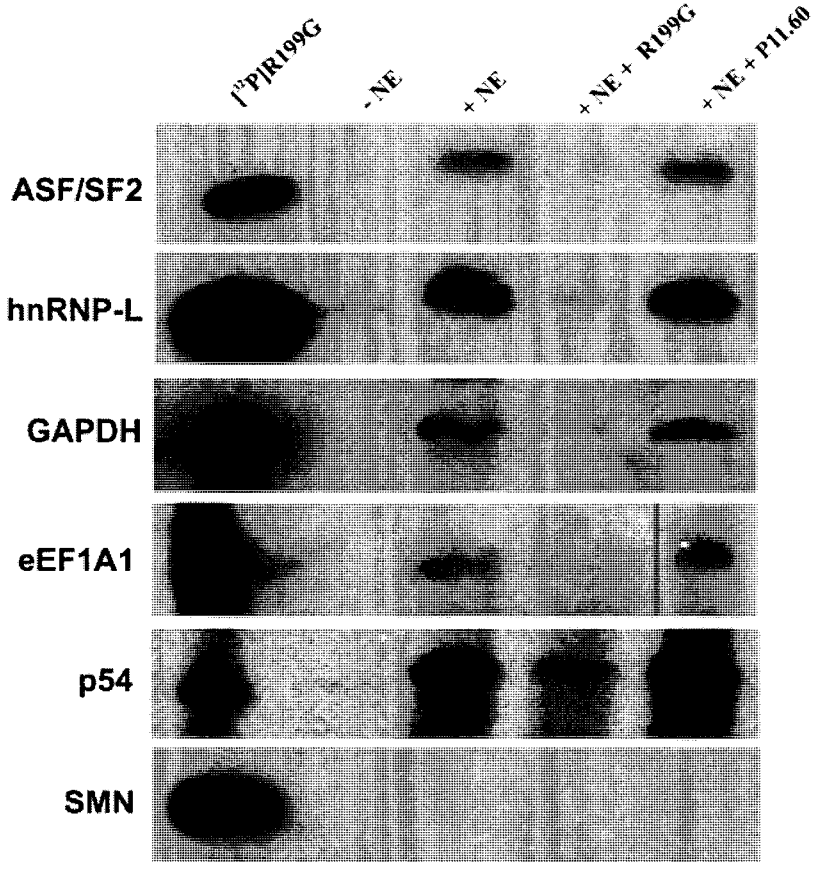
RNA was trapped in the wells of the gel, this finding was considered inconclusive. No apparent binding of R199G to the other proteins was observed, although activity of these proteins was not tested.

Interaction of p54^{nrb}, eEF1A1, GAPDH, hnRNP-L, and ASF/SF2 with R199G in HeLa NE.

To confirm the interactions between R199G and the newly identified R199G-interacting proteins, co-immunoprecipitation experiments were performed with R199G and nuclear extract proteins. Because our analysis was limited to commercial-availability of antibodies, only p54^{nrb}, eEF1A1, hnRNP-L, and ASF/SF2 were analyzed. We also included GAPDH in our analysis, based on previous findings indicating GAPDH association with the opposite extremity of the HDV genome (Lin et al., 2000) which also interacts with RNAP II (Filipovska and Konarska, 2000; Greco-Stewart et al., 2007). Radiolabeled R199G was incubated with HeLa NE proteins in the presence or absence of molecular excess of P11.60 or unlabeled R199G to verify the specificity of the interaction. Following co-immunoprecipitation using specific antibodies against those five proteins, the resulting mixture was resolved on denaturing PAGE and visualized by autoradiography (Figure 7.2.5).

R199G co-immunoprecipitated with p54^{nrb}, eEF1A1, GAPDH, hnRNP-L, and ASF/SF2. The presence of a molecular excess of the non-specific competitor P11.60 did not significantly alter the amount of R199G that was co-immunoprecipitated. However, the interactions were greatly reduced by the addition of molecular excess of non-radiolabeled R199G RNA, and no R199G RNA was detected in the absence of nuclear extract. Finally, no R199G was co-immunoprecipitated using an α-IgG antibody (data not shown), or an antibody specific for SMN (Figure 7.2.5), a protein containing a Tudor domain which has shown no interaction with HDV RNA by EMSA (Figure 7.2.4). Taken together, our

Figure 7.2.5: Interaction of R199G with various nuclear factors in HeLa nuclear extract. Radiolabeled R199G RNA was co-immunoprecipitated with each eEF1A1, $p54^{nrb}$, hnRNP-L, GAPDH, ASF/SF2, and SMN from HeLa nuclear extract using the corresponding antibody. Specificity of the interaction was verified by addition of 50X molar excess of either a specific (R199G) or non-specific (P11.60) RNA competitor.


[32P]R199G
- NE
+ NE
+ NE + R199G
+ NE + P11.60
ASF/SF2
hnRNP-L
GAPDH
eEF1A1
p54
SMN

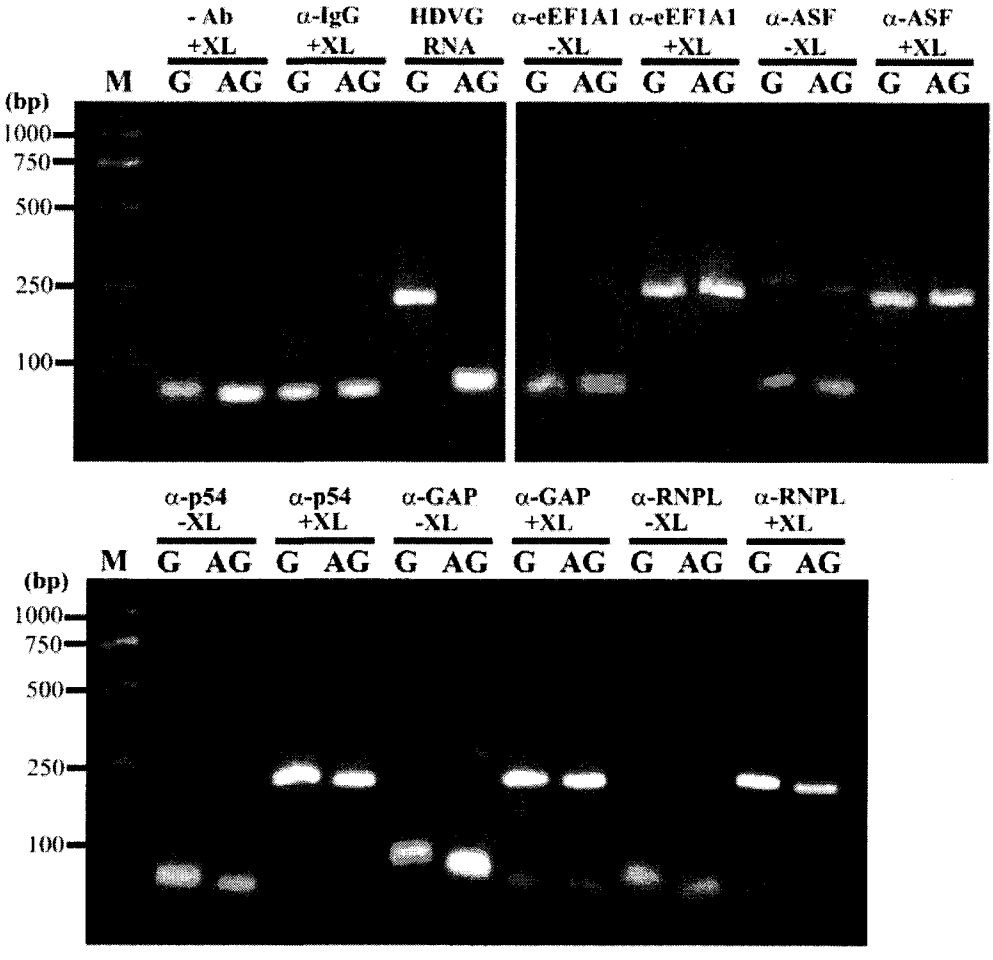
observations demonstrate that these proteins interact specifically with R199G within HeLa NE.

Interaction of $p54^{nrb}$, eEF1A1, GAPDH, hnRNP-L, and ASF/SF2 with both polarities of HDV RNA in HeLa cells replicating HDV RNA.

To establish the significance of these interactions in a physiological context, HeLa cells replicating HDV RNA were subjected to a ribonucleoprotein immunoprecipitation assay (RIPA; Niranjanakumari et al., 2002). This technique is both sensitive and specific, and has recently been employed by our laboratory to substantiate cellular interactions of both PSF and RNAP II with HDV RNA (Greco-Stewart et al., 2006; 2007).

HeLa cells were transfected with both a dimeric HDV transcript of genomic polarity and a plasmid transiently expressing HDAg-S from a CMV promoter; HDAg-S has been shown to be necessary for HDV synthesis (Kuo et al., 1989). Four days post-transfection, the cells were treated with formaldehyde, a reversible cross-linking agent. Following cell lysis and co-immunoprecipitations of the RNA-protein complexes using specific antibodies, the cross-links were heat-reversed. The RNA was isolated and analyzed by RT-PCR for the presence of HDV RNA (Figure 7.2.6). Bands corresponding to the R199 cDNA (i.e. 232 bp) were detected when the cross-linked extract was subjected to co-immunoprecipitation with antibodies against $p54^{nrb}$, eEF1A1, GAPDH, hnRNP-L, and ASF/SF2. R199 cDNA was not detected in the absence of antibody, or in the presence of α-IgG antibody, indicating that the RT-PCR products obtained were not the results of non-specific binding of HDV RNA to either the antibody or the protein-G agarose beads. Furthermore, no HDV RNA-protein complexes were recovered from non-cross-linked HeLa lysates. However, in the case of eEF1A1 and ASF/SF2, a small amount of PCR product was recovered after the amplification

Figure 7.2.6: Interaction of R199G with various cellular factors in cultured HeLa cells. Ribonucleoprotein immunoprecipitation assays were performed to confirm the interaction of HDV RNA of both genomic and antigenomic polarity (G or AG) with the candidate host proteins in HDV infected HeLa cells. HDV RNA was co-immunoprecipitated from lysates of formaldehyde cross-linked HDV-replicating cells with the indicated proteins using the corresponding antibodies. HDV RNA was detected by reverse transcription using sense-specific primers. Co-immunoprecipitations from non-cross-linked cell lysates serve as specificity controls. HDVG RNA lanes are controls using *in vitro* transcribed HDV RNA of genomic polarity.


- Ab
+XL
α-IgG
+XL
HDVG
RNA
α-eEF1A1
-XL
α-eEF1A1
+XL
α-ASF
-XL
α-ASF
+XL
M
G AG G AG G AG G AG G AG G AG G AG
(bp)
1000
750
500
250
100
α-p54
-XL
α-p54
+XL
α-GAP
-XL
α-GAP
+XL
α-RNPL
-XL
α-RNPL
+XL
M
G AG G AG G AG G AG G AG G AG
(bp)
1000
750
500
250
100

reactions, which might indicate strong associations of these proteins with HDV RNA. Taken together, these data indicate the specific interaction of these proteins with HDV RNA within HeLa cells.

Additionally, the primers used for the reverse transcription reactions were designed to distinguish between the polarities of HDV RNA. For instance, the antisense primer used to generate the R199G cDNA was only able to reverse transcribe from the genomic polarity of HDV RNA (Figure 7.2.6, HDVG RNA lanes). Because RNA fragments corresponding to both polarities of HDV were detected when the cross-linked extract was co-immunoprecipitated with specific antibodies against $p54^{nrb}$, eEF1A1, GAPDH, hnRNP-L, and ASF/SF2, we conclude that these proteins associate with both polarities of HDV RNA within HeLa cells. These data indicate that each polarity of the HDV genome has at least one region that interacts with these proteins, and that these interactions occur in a cellular context during HDV replication.

7.2.5 Discussion

HDV is unique among human viral pathogens in that it has a very limited coding capacity, and thus must rely on host factors for its replication. In addition, because the HDV RNA genome interacts with several host proteins, it is likely that these interactions affect normal host functions. Despite their functional importance in both viral replication and pathogenicity, only a few cellular proteins have been reported to interact with HDV RNA. Accordingly, this study was undertaken to identify additional host factors that interact specifically with HDV RNA and that might be involved in HDV transcription and/or replication. We used an HDV RNA fragment corresponding to the right stem-loop domain of the genomic polarity that has been suggested to include an RNA promoter for RNAP II with

a role in viral replication and/or transcription of the HDAg mRNA (Abrahem and Pelchat, 2008; Beard et al., 1996; Greco-Stewart et al., 2007; Gudima et al., 2000). Using a UV cross-linking approach and RNA affinity chromatography, we have identified several HDV RNA-interacting partners. As reported previously (Greco-Stewart et al., 2006), UV cross-linking of HeLa NE to R199G did not yield complexes corresponding to the molecular weight of the individual proteins. It is likely that several of these proteins might have cross-linked to R199G simultaneously, which would account for the detection of a single high molecular weight complex.

Because a number of these proteins are known to be involved in splicing events, we screened additional purified splicing factors in EMSAs. The interactions between HDV RNA and these proteins were further validated by co-immunoprecipitation both *in vitro* and in HeLa cells replicating HDV RNA. In addition to PSF and RNAP II (Abrahem and Pelchat, 2008; Greco-Stewart et al., 2006; 2007), our results demonstrate the interactions of eEF1A1, $p54^{nrb}$, hnRNP-L, GAPDH, and ASF/SF2 with HDV RNA. Since we identified many different proteins using these various experimental approaches it is possible that these proteins might represent only a small subset of the host proteins interacting with HDV RNA. On the other hand, we might have been too cautious in our analysis of the mass spectrometry results, as single peptides for both hnRNP-L and ASF/SF2 were also detected in some of the cross-linked complexes.

The identified proteins are frequently linked to viral replication and/or transcription. ASF/SF2 belongs to the family of SR (serine/arginine-rich) proteins and promotes exon inclusion by recognizing and binding to exonic splicing enhancers (ESEs), and this binding is suggested to be modulated by the C-terminal domain (CTD) of the largest subunit of RNAP II (Cramer et al., 1999; McCracken et al., 1997). $p54^{nrb}$ is related to PSF (Dong et al.,

1993), which we previously reported to interact specifically and directly with HDV RNA domains also bound by RNAP II (Greco-Stewart et al., 2006; 2007). PSF forms a heterotetramer with $p54^{nrb}$, and together they were shown to interact with the CTD of RNAP II (Emili et al., 2002). hnRNP-L belongs to a family of abundant nuclear proteins frequently linked to viral replication (for a review see Lai, 1998). GAPDH has been shown to greatly stimulate transcription by RNAP II in *Xenopus laevis* oocytes (Morgenegg et al., 1986) and has been reported to associate with RNA polymerase II in *Schizosaccharomyces pombe* through a direct interaction with the Rpb7 subunit of RNAP II (Mitsuzawa et al., 2005). GAPDH was also reported to be involved in RNA-dependent viral RNA synthesis by binding to either the 5'-UTR (+) of the hepatitis A virus or the 3'-UTR (+/-) of the human parainfluenza virus-3 (De et al., 1996; Schultz et al., 1996). The translation factor eEF1A1 was reported to be involved in viral RNA synthesis by associating with the RNA-dependent RNA polymerase activities of poliovirus (Harris et al., 1994), vesicular stomatitis virus (Das et al., 1998), turnip yellow mosaic virus (Joshi et al., 1986) and West Nile virus (Blackwell and Brinton, 1997). Thus, it is tempting to speculate that these proteins might function to enhance the RNAP II-mediated replication of HDV RNA.

It is also possible that some of these proteins are involved in HDV RNA intracellular shuttling. For instance, PSF, $p54^{nrb}$, and eEF1A have been identified as constituents of an RNA-transporting granule in the mouse brain (Kanai et al., 2004). Furthermore, hnRNP-L binds to the pre-mRNA processing enhancer (PPE) of the herpes simplex virus thymidine kinase intronless mRNA and enhances its cytoplasmic export (Liu and Mertz, 1995). Some SR proteins have also been shown to enhance cytoplasmic accumulation of an intronless mRNA by interacting with a *cis*-acting RNA element (Huang and Steitz, 2001). The

functions of these proteins with respect to HDV might be related to their ability to bind and export intronless mRNAs since HDV RNA itself is intronless.

The high level of accumulation of HDV RNA in infected cells (Chen et al., 1986) suggests that the interaction of HDV RNA with these factors might interfere with several of their normal cellular functions, thereby eliciting the virus' pathogenic effect. For instance, it is possible that HDV alters ASF/SF2, PSF, $p54^{nrb}$, and/or hnRNP-L–mediated alternative splicing, since disregulated splicing has been associated with many types of cancer (Castiglioni et al., 2006; Lee et al., 2006; Pospisil et al., 2006), and HDV is known to be associated with liver cancer (Fattovich et al., 2004; Romeo et al., 2009; Su et al., 2006). ASF/SF2 has also been shown to play a significant role in protecting chromosomal DNA from the damaging effects of R-loops that form between nascent RNA transcripts and template DNA; inactivation of this protein leads to genomic instability (Li and Manley, 2005; Li et al., 2005). PSF represses the expression of numerous oncogenes (Song et al., 2004). The interaction of PSF with HDV RNA might reverse this repression and promote metastasis (Song et al., 2002; Song et al., 2004; Song et al., 2005). eEF1A1 has been shown to bind the non-coding RNA HSR1 and to activate the heat shock transcription factor 1 (HSF1), which induces the expression of heat shock proteins. Interestingly, Hsp105 was reported to be downregulated in Huh7 cells transiently transfected with a plasmid encoding genomic HDV RNA (Mota et al., 2008). It is possible that HDV RNA might compete with HSR1 for eEF1A binding, thereby impeding a heat shock response. GAPDH has been proposed to play an essential role in apoptosis (Barbini et al., 2007; Fukuhara et al., 2001; Ishitani and Chuang, 1996; Ishitani et al., 1998). The potential disregulation of this important process by HDV might also contribute to liver disease. Although these proteins are very

abundant, their essential nature suggests that interfering with their ability to perform their cellular activities would have an adverse effect on many aspects of cellular functioning.

In summary, we have identified five proteins that interact specifically with HDV RNA: eEF1A1, $p54^{nrb}$, hnRNP-L, ASF/SF2, and GAPDH. Similar to what is observed for other RNA viruses, normal cellular components associated with RNA-processing pathways and translation machinery appear to be exploited by HDV (Lai, 1998). Although their precise functions in HDV biology are not fully understood, their interaction with the HDV RNA genome is consistent with their known biological properties. Further investigation of the interactions of these proteins with HDV RNA and on the consequences of these interactions on the various processes mediated by these proteins might lead to a better comprehension of both HDV replication and pathogenesis.

7.2.6 Material and methods

Synthesis of R199G RNA.

The DNA template used for the synthesis of R199G RNA was generated by PCR off the pHDVd2 plasmid, a derivative of pBluescriptKS$^+$ (Stratagene) harboring a dimer of HDV cDNA (Kuo et al., 1988). The forward primer was designed to contain the T7 promoter sequence. R199G was synthesized by *in vitro* transcription using the T7 RNA polymerase (New England Biolabs; NEB). Transcription products were subjected to DNAse I (Promega) digestion at 37°C for 30 min, resolved on denaturing 5% polyacrylamide gel, visualized under shortwave UV light, excised, eluted in 500 mM ammonium acetate, 0.1% SDS, 10 mM EDTA, and ethanol-precipitated. The precipitated RNA was resuspended in 100 µl H_2O, passed through a Sephadex G-50 column (GE Healthcare) and ethanol-precipitated.

RNA was quantified by spectrophotometry at a wavelength of 260 nm. P11.60 RNA was synthesized as previously described (Pelchat and Perreault, 2004).

Radiolabeling of R199G RNA was carried out as follows: ten pmol of R199G was dephosphorylated using calf intestinal phosphatase (NEB) and 5'-labeled using T4 polynucleotide kinase (NEB) in the presence of 10 μCi [γ-^{32}P] ATP (GE Healthcare). Radiolabeled R199G was purified by phenol/chloroform extraction, passed through Sephadex G-50 (GE Healthcare), ethanol-precipitated, and resuspended in H_2O.

UV cross-linking.

Radiolabeled R199G (100 fmol) was incubated with 10 μg of HeLa NE (Accurate Chemical and Scientific) in 20 mM HEPES pH 7.9, 100 mM KCl, 0.2 mM EDTA, 6.0 mM $MgCl_2$, 0.5 mM DTT and 20% glycerol in a final volume of 10 μl. P11.60 (50 pmol) or non-radiolabeled R199G (50 pmol) were added to certain reactions, where indicated. Reaction mixtures were irradiated for 20 min on ice in a UV Stratalinker 2400 (Stratagene) at a wavelength of 365 nm. Following addition of 10 μl of 2X SDS loading dye (100 mM Tris–HCl, pH 6.8, 4% SDS, 0.1% bromophenol blue, and 20% glycerol), RNA-protein complexes were resolved on 10% SDS-PAGE and visualized by autoradiography. To identify the proteins in the complex, the experiment was repeated, the gel was stained with methylene blue, the resulting bands were excised and three samples were sent individually to WEMB Biochem Inc. (Toronto, Canada) for protein analysis by LC-MS/MS, as previously described (Greco-Stewart et al., 2006).

RNA affinity chromatography.

Two hundred pmol of *in vitro*-synthesized R199G was oxidized in 20 mM Tris–HCl (pH 7.5) and 10 mM Sodium-m-periodate, precipitated with ethanol, and resuspended in 0.1 M sodium acetate (pH 5.0). Oxidized RNA was incubated overnight at 4°C with adipic acid dihydrazide agarose beads (Sigma) that had been pre-washed with 0.1 M sodium acetate (pH 5.0). RNA-coupled beads were then washed 3 times with wash buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA). Fifty μg HeLa nuclear extract (Accurate Chemical and Scientific) were pre-incubated with non-coupled beads partially equilibrated in binding buffer (10 mM Tris-HCl pH 7.5, 10 mM KCl, 150 mM NaCl, 1 mM $MgCl_2$, 50% glycerol) for 1 h at room temperature and centrifuged to remove unbound proteins. The eluted proteins were added to the RNA-coupled beads and incubated for 1 h at room temperature. The beads were washed 5 times with binding buffer, and the bound proteins were eluted following addition of 1X SDS loading buffer and incubation at 95°C for 5 min. Proteins were resolved on 10% SDS-PAGE and visualized by silver staining (BioRad). Resulting bands were excised, subjected to in-gel trypsin digestion and analyzed with LC-MS/MS.

Electrophoretic mobility shift assay.

^{32}P-labeled RNA (100 fmol) was mixed with 0.5 μg of purified, GST-tagged protein. Reactions were carried out in a final volume of 10 μl containing 20 mM Tris-HCl, pH 7.5, 100 mM KCl, 1 mM EDTA, 1 mM DTT, 0.1 mg/ml BSA, and 20 mM $MgCl_2$. Poly(A) (10 pmol) was used as a non-specific competitor. Following a 30-min incubation at 4°C, 2 μl of agarose loading dye were added, and the samples were subjected to non-denaturing 5% polyacrylamide gel electrophoresis (49:1, acrylamide:bis-acrylamide) at 40 volts in 1X Tris-

Borate-EDTA (TBE) buffer at room temperature. Bands were visualized by phosphorimaging.

Co-immunoprecipitation assay.

The co-immunoprecipitation experiments were carried out using the Protein G Immunoprecipitation Kit (Sigma) with a modified protocol. Radiolabeled R199G RNA (0.5 pmol) was incubated with 50 µg HeLa nuclear extract (Accurate Chemical and Scientific) in 1X RIPA buffer (50 mM Tris–HCl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl) for 30 min at 4°C. P11.60 RNA (50 pmol) or non-radiolabeled R199G RNA (50 pmol) were added to certain reactions, where indicated. Following the 30-min incubation, 5 µg of either α-eEF1A1 (Abnova), α-ASF/SF2 (Zymed), α-p54^{nrb} (Upstate), α-hnRNP-L (Abcam), α-GAPDH (Santa Cruz), or α-SMN (BD Transduction Laboratories) were added, and the reaction mixtures were incubated for 1 h at 4°C before the addition of 30 µl pre-washed protein G agarose beads. After an overnight incubation at 4°C, the beads were washed 5 times with 1X RIPA buffer and the RNA was eluted with 40 µl of acrylamide loading dye. RNA was resolved on denaturing 5% polyacrylamide gel and visualized by autoradiography.

Ribonucleoprotein immunoprecipitation assay (RIPA).

The RIPA analyses were performed using the method of Niranjanakumari et al. (Niranjanakumari et al., 2002) with modifications described by Greco-Stewart et al. (2006; 2007). Briefly, HeLa cells were co-transfected with 10 µg of dimeric HDV genome of positive polarity and 6 µg of plasmid containing the HDAg-S under control of the high-expression CMV promoter (Lazinski and Taylor, 1993). Approximately 10^6 cells were cross-

linked by treatment with 1% formaldehyde for 5 minutes, collected by centrifugation, washed twice with 1X PBS, and preserved at -80°C prior to analysis.

Cells were resuspended in 2 mL of RIPA buffer (50 mM Tris–Cl, pH 7.5, 1% Nonidet P-40 (NP-40), 0.5% sodium deoxycholate, 0.05% SDS, 1 mM EDTA, 150 mM NaCl), disrupted by sonication, and the resulting lysate was collected by centrifugation at 13,000 rpm. Immunoprecipitation reactions were carried out using the Protein G Immunoprecipitation Kit (Sigma). The cleared lysate (500 µl) was incubated with 5 µg of each of the following antibodies: goat monoclonal α-mouse-IgG (Sigma-Aldrich), α-eEF1A1 (Abnova), α-ASF/SF2 (Zymed), α-p54^{nrb} (Upstate), α-hnRNP-L (Abcam), and α-GAPDH (Santa Cruz). Pre-washed protein G agarose beads (30 µl) were then added, and the reactions were incubated overnight at 4°C. Beads were washed 5 times with 600 µl of RIPA buffer containing 1 M urea, heated at 70 °C for 45 min in 50 mM Tris-HCl (pH 7.0), 5 mM EDTA, 10 mM DTT, and 1% SDS, and eluates were collected by centrifugation. Nucleic acids were extracted with phenol/chloroform, ethanol-precipitated in the presence of 100 µg of tRNA, and treated with DNase I (Promega). Following another phenol/chloroform extraction and ethanol precipitation, reverse transcription was performed as previously described (Pelchat et al., 2001) using AMV reverse transcriptase (Promega). Primers used for the RT reaction were those used to synthesize R199G cDNA (GGAATTCTAATACGACTCACTATAGGGACTGCTCGAGGATCTCTTCTCTCCC and GGAATTCACATCCCCTCTCGGGTGC; underlined nucleotides indicate the T7 promoter), and differentiate between the two polarities of the HDV genome. Following the RT reaction, samples were amplified by 40 cycles of PCR using Vent polymerase (New England Biolabs) with an annealing temperature of 50°C and an extension time of 1 minute. Samples were

visualized by electrophoresis on 2% agarose gel in the presence of SYBR green I (Invitrogen).

7.2.7 Acknowledgements

We thank Ms. Emilie Lemay and Ms. Catherine St-Laurent Thibault for technical assistance in both UV cross-linking and RNA affinity chromatography. GST-tagged splicing factors were a generous gift from Dr. Jocelyn Côté (University of Ottawa, Canada). V.S.G.-S. was supported by a Graduate Fellowship Award from the Canadian Blood Services. This work was supported by a discovery grant from the Natural Sciences and Engineering Research Council (NSERC) to M.P.

7.2.8 References

Abrahem, A., Pelchat, M., 2008. Formation of an RNA polymerase II preinitiation complex on an RNA promoter derived from the hepatitis *delta* virus RNA genome. Nucleic Acids Res. 36, 5201-5211.

Barbini, L., Rodriguez, J., Dominguez, F., Vega, F., 2007. Glyceraldehyde 3-phosphate dehydrogenase exerts different biologic activities in apoptotic and proliferating hepatocytes according to its subcellular localization. Mol. Cell. Biochem. 300, 19-28.

Beard, M.R., Macnaughton, T.B., Gowans, E.J., 1996. Identification and characterization of a hepatitis *delta* virus RNA transcriptional promoter. J. Virol. 70, 4986-4995.

Bichko, V.V., Taylor, J.M., 1996. Redistribution of the *delta* antigens in cells replicating the genome of hepatitis *delta* virus. J. Virol. 70, 8064-8070.

Blackwell, J.L., Brinton, M.A., 1997. Translation elongation factor-1 alpha interacts with the 3' stem-loop region of West Nile virus genomic RNA. J. Virol. 71, 6433-6444.

Castiglioni, F., Tagliabue, E., Campiglio, M., Pupa, S.M., Balsari, A., Menard, S., 2006. Role of exon-16-deleted HER2 in breast carcinomas. Endocr. Relat .Cancer 13, 221-232.

Chang, J., Nie, X., Chang, H.E., Han, Z., Taylor, J., 2008. Transcription of hepatitis *delta* virus RNA by RNA polymerase II. J. Viro.l 82, 1118-1127.

Chen, P.J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., Taylor, J., 1986. Structure and replication of the genome of the hepatitis *delta* virus. Proc. Natl. Acad. Sci 83, 8774-8778.

Circle, D.A., Neel, O.D., Robertson, H.D., Clarke, P.A., Mathews, M.B., 1997. Surprising specificity of PKR binding to delta agent genomic RNA. RNA 3, 438-348.

Cole, S.M., Gowans, E.J., Macnaughton, T.B., Hall, P.D., Burrell, C.J., 1991. Direct evidence for cytotoxicity associated with expression of hepatitis *delta* virus antigen. Hepatology 13, 845-851.

Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E., Kornblihtt, A.R., 1999. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol. Cell 4, 251-258.

Das, T., Mathur, M., Gupta, A.K., Janssen, G.M., Banerjee, A.K., 1998. RNA polymerase of vesicular stomatitis virus specifically associates with translation elongation factor-1 alphabetagamma for its activity. Proc. Natl. Acad. Sci. 95, 1449-1454.

De, B.P., Gupta, S., Zhao, H., Drazba, J.A., Banerjee, A.K., 1996. Specific interaction *in vitro* and *in vivo* of glyceraldehyde 3-phosphate dehydrogenase and LA protein with *cis*-acting RNAs of human parainfluenza virus type 3. J. Biol. Chem. 271, 24728-24735.

Dong, B., Horowitz, D.S., Kobayashi, R., Krainer, A.R., 1993. Purification and cDNA cloning of HeLa cell $p54^{nrb}$, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and *Drosophila* NONA/BJ6. Nucleic Acids Res. 21, 4085-4092.

Emili, A., Shales, M., McCracken, S., Xie, W., Tucker, P.W., Kobayashi, R., Blencowe, B.J., Ingles, C.J., 2002. Splicing and transcription-associated proteins PSF and $p54^{nrb}$/nonO bind to the RNA polymerase II CTD. RNA 8, 1102-1111.

Farci, P., 2003. *Delta* hepatitis: an update. J. Hepatol. 39 Suppl. 1, S212-9.

Fattovich, G., Stroffolini, T., Zagni, I., Donato, F., 2004. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. Gastroenterology 127, S35-50.

Filipovska, J., Konarska, M.M., 2000. Specific HDV RNA-templated transcription by pol II in vitro. RNA 6, 41-54.

Fukuhara, Y., Takeshima, T., Kashiwaya, Y., Shimoda, K., Ishitani, R., Nakashima, K., 2001. GAPDH knockdown rescues mesencephalic dopaminergic neurons from MPP+ -induced apoptosis. Neuroreport 12, 2049-2052.

Greco-Stewart, V.S., Thibault, C.S., Pelchat, M., 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology 356, 35-44.

Greco-Stewart, V.S., Miron, P., Abrahem, A., Pelchat, M., 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology 357, 68-78.

Greco-Stewart, V.S., Schissel, E., Pelchat, M., 2009. The hepatitis *delta* virus RNA genome interacts with the human RNA polymerases I and III. Virology 386, 12-15.

Gudima, S., Wu, S.Y., Chiang, C.M., Moraleda, G., Taylor, J., 2000. Origin of hepatitis *delta* virus mRNA. J Virol 74, 7204-7210.

Harris, K.S., Xiang, W., Alexander, L., Lane, W.S., Paul, A.V., Wimmer, E., 1994. Interaction of poliovirus polypeptide 3CDpro with the 5' and 3' termini of the poliovirus genome. Identification of viral and cellular cofactors needed for efficient binding. J. Biol. Chem. 269, 27004-27014.

Hartmann, A.M., Nayler, O., Schwaiger, F.W., Obermeier, A., Stamm, S., 1999. The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). Mol. Biol. Cell 10, 3909-3026.

Hastings, M.L., Krainer, A.R., 2001. Pre-mRNA splicing in the new millennium. Curr. Opin. Cell Biol. 13, 302-309.

Huang, Y., Steitz, J.A., 2001. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. Mol. Cell 7, 899-905.

Ishitani, R., Chuang, D.M., 1996. Glyceraldehyde 3-phosphate dehydrogenase antisense oligodeoxynucleotides protect against cytosine arabinonucleoside-induced apoptosis in cultured cerebellar neurons. Proc. Natl. Acad. Sci. 93, 9937-9941.

Ishitani, R., Tanaka, M., Sunaga, K., Katsube, N., Chuang, D.M., 1998. Nuclear localization of overexpressed glyceraldehyde 3-phosphate dehydrogenase in cultured cerebellar neurons undergoing apoptosis. Mol. Pharmacol. 53, 701-707.

Joshi, R.L., Ravel, J.M., Haenni, A.L., 1986. Interaction of turnip yellow mosaic virus Val-RNA with eukaryotic elongation factor EF-1 alpha. Search for a function. EMBO J. 5, 1143-1148.

Kanai, Y., Dohmae, N., Hirokawa, N., 2004. Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. Neuron 43, 513-525.

Kuo, M.Y., Goldberg, J., Coates, L., Mason, W., Gerin, J., Taylor, J., 1988. Molecular cloning of hepatitis *delta* virus RNA from an infected woodchuck liver: sequence, structure, and applications. J. Virol. 62, 1855-1861.

Kuo, M.Y., Chao, M., Taylor, J., 1989. Initiation of replication of the human hepatitis *delta* virus genome from cloned DNA: role of *delta* antigen. J. Virol. 63, 1945-1950.

Lai, M.M., 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. Virology 244, 1-12.

Lazinski, D.W., Taylor, J.M., 1993. Relating structure to function in the hepatitis *delta* virus antigen. J. Virol. 67, 2672-2680.

Lee, E.J., Jo, M., Park, J., Zhang, W., Lee, J.H., 2006. Alternative splicing variants of IRF-1 lacking exons 7, 8, and 9 in cervical cancer. Biochem. Biophys. Res. Commun. 347, 882-888.

Li, X., Manley, J.L., 2005. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell 122, 365-378.

Li, X., Wang, J., Manley, J.L., 2005. Loss of splicing factor ASF/SF2 induces G2 cell cycle arrest and apoptosis, but inhibits internucleosomal DNA fragmentation. Genes Dev. 19, 2705-2714.

Li, Y.J., Macnaughton, T., Gao, L., Lai, M.M., 2006. RNA-templated replication of hepatitis *delta* virus: genomic and antigenomic RNAs associate with different nuclear bodies. J. Virol. 80, 6478-6486.

Lin, S.S., Chang, S.C., Wang, Y.H., Sun, C.Y., Chang, M.F., 2000. Specific interaction between the hepatitis *delta* virus RNA and glyceraldehyde 3-phosphate dehydrogenase: an enhancement on ribozyme catalysis. Virology 271, 46-57.

Liu, X., Mertz, J.E., 1995. HnRNP L binds a *cis*-acting RNA sequence element that enables intron-dependent gene expression. Genes Dev. 9, 1766-1780.

Macnaughton, T.B., Shi, S.T., Modahl, L.E., Lai, M. M., 2002. Rolling circle replication of hepatitis *delta* virus RNA is carried out by two different cellular RNA polymerases. J. Virol. 76, 3920-3927.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., Bentley, D.L., 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature 385, 357-361.

Mitsuzawa, H., Kimura, M., Kanda, E., Ishihama, A., 2005. Glyceraldehyde 3-phosphate dehydrogenase and actin associate with RNA polymerase II and interact with its Rpb7 subunit. FEBS Lett. 579, 48-52.

Modahl, L.E., Macnaughton, T.B., Zhu, N., Johnson, D.L., Lai, M.M., 2000. RNA-Dependent replication and transcription of hepatitis *delta* virus RNA involve distinct cellular RNA polymerases. Mol. Cell. Biol. 20, 6030-6039.

Moraleda, G., Taylor, J., 2001. Host RNA polymerase requirements for transcription of the human hepatitis *delta* virus genome. J. Virol. 75, 10161-10169.

Morgenegg, G., Winkler, G.C., Hubscher, U., Heizmann, C.W., Mous, J., Kuenzle, C.C., 1986. Glyceraldehyde 3-phosphate dehydrogenase is a nonhistone protein and a possible activator of transcription in neurons. J. Neurochem. 47, 54-62.

Mota, S., Mendes, M., Penque, D., Coelho, A.V., Cunha, C., 2008. Changes in the proteome of Huh7 cells induced by transient expression of hepatitis D virus RNA and antigens. J. Proteomics 71, 71-79.

Niranjanakumari, S., Lasda, E., Brazas, R., Garcia-Blanco, M.A., 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions *in vivo*. Methods 26, 182-190.

Pelchat, M., Cote, F., Perreault, J.P., 2001. Study of the polymerization step of the rolling circle replication of peach latent mosaic viroid. Arch. Virol. 146, 1753-1763.

Pelchat, M., Perreault, J.P., 2004. Binding site of *Escherichia coli* RNA polymerase to an RNA promoter. Biochem. Biophys. Res. Commun. 319, 636-642.

Ponting, C.P., 1997. Tudor domains in proteins that interact with RNA. Trends Biochem. Sci. 22, 51-52.

Pospisil, H., Herrmann, A., Butherus, K., Pirson, S., Reich, J.G., Kemmner, W., 2006. Verification of predicted alternatively spliced Wnt genes reveals two new splice variants (CTNNB1 and LRP5) and altered Axin-1 expression during tumour progression. BMC Genomics 7, 148.

Romeo, R., Ninno, E.D., Rumi, M., Russo, A., Sangiovanni, A., Franchis, R.D., Ronchi, G., Columbo, M., 2009. A 28-year study of the course of hepatitis *delta* infection: a risk factor for cirrhosis and hepatocellular carcinoma. Gastroenterology 136, 1629-1638.

Schultz, D.E., Hardin, C.C., Lemon, S.M., 1996. Specific interaction of glyceraldehyde 3-phosphate dehydrogenase with the 5'-nontranslated RNA of hepatitis A virus. J. Biol. Chem. 271, 14134-14142.

Song, X., Wang, B., Bromberg, M., Hu, Z., Konigsberg, W., Garen, A., 2002. Retroviral-mediated transmission of a mouse VL30 RNA to human melanoma cells promotes

metastasis in an immunodeficient mouse model. Proc. Natl. Acad. Sci. 99, 6269-6273.

Song, X., Sui, A., Garen, A., 2004. Binding of mouse VL30 retrotransposon RNA to PSF protein induces genes repressed by PSF: effects on steroidogenesis and oncogenesis. Proc. Natl. Acad. Sci. 101, 621-626.

Song, X., Sun, Y., Garen, A., 2005. Roles of PSF protein and VL30 RNA in reversible gene regulation. Proc. Natl. Acad. Sci. 102, 12189-12193.

Su, C.W., Huang, Y.H., Huo, T.I., Shih, H.H., Sheen, I.J., Chen, S.W., Lee, P.C., Lee, S.D., Wu, J. C., 2006. Genotypes and viremia of hepatitis B and D viruses are associated with outcomes of chronic hepatitis D patients. Gastroenterology 130, 1625-1635.

Taylor, J. M., 2006. Hepatitis *delta* virus. Virology 344, 71-76.

Vioque, A., Altman, S., 1986. Affinity chromatography with an immobilized RNA enzyme. Proc. Natl. Acad. Sci. 83, 5904-5908.

Wang, D., Pearlberg, J., Liu, Y.T., Ganem, D., 2001. Deleterious effects of hepatitis *delta* virus replication on host cell proliferation. J. Virol. 75, 3600-3604.

Wong, S.K., Lazinski, D.W., 2002. Replicating hepatitis *delta* virus RNA is edited in the nucleus by the small form of ADAR1. Proc. Natl. Acad. Sci. 99, 15118-15123.

Yamaguchi, Y., Delehouzee, S., Handa, H., 2002. HIV and hepatitis *delta* virus: evolution takes different paths to relieve blocks in transcriptional elongation. Microbes Infect. 4, 1169-1175.

CHAPTER 8:

CURRICULUM VITAE

Valerie Suzanne Greco-Stewart

Education

2004 – 2009 Ph.D. (Biochemistry), University of Ottawa.
Thesis title: Characterization of the interaction of human DNA-dependent RNA polymerases and transcription factors with RNA from the hepatitis *delta* virus: Novel perspectives in HDV biology.
Thesis supervisor: Dr. M. Pelchat.

2001 – 2004 M.Sc. (Microbiology, and Immunology), University of Ottawa.
Thesis title: Structural and functional analysis of the division inhibitor MinC from *Neisseria gonorrhoeae*.
Thesis supervisor: Dr. J. R. Dillon.

1997 – 2001 B.Sc. (Biology) with Honours Biotechnology option (Suma Cum Laude) University of Ottawa.
Honours project: Site-directed mutagenesis of essential residues of MinC from *Neisseria gonorrhoeae* results in loss of protein functionality.
Thesis supervisor: Dr. J. R. Dillon.

Work Experience

2004
Research technician
Dr. J. R. Dillon, University of Ottawa; Dr. S. Ramirez-Arcos, Canadian Blood Services

- Completed functional analysis of the cell division inhibitor MinC from *Neisseria gonorrhoeae* initiated during M.Sc. research.
- Optimized a bacterial two-hybrid assay for use in cell division research and established interactions among MinC and other gonococcal cell division proteins.

2001
Undergraduate NSERC summer scholarship recipient
Dr. S. Evans and Dr. J. R. Dillon, University of Ottawa

- Performed research involving the cell division protein FtsZ from *Neisseria gonorrhoeae.*
- Studies involved construction of an FtsZ-polyhistidine containing construct, optimizing a protein purification protocol, SDS-PAGE electrophoresis, Western blot analysis, and preliminary crystallization techniques.

2000
Undergraduate NSERC summer scholarship recipient
Dr. V. L. Trudeau, University of Ottawa

- Performed research involving the identification of novel glutamic acid decarboxylase (GAD) sequences from elasmobranches to elucidate GAD phylogeny.
- Research included basic molecular biology techniques to locate novel GAD sequences and intense bioinformatic analysis for sequence identification and phylogenetic tree construction.

1999 – 2000 (Part time)
Laboratory technician
Synthaid Biotechnology Institute
Dr. G. Alvarado

- Synthesized custom oligonucleotide primers and assisted with the synthesis of peptide sequences.
- Responsible for laboratory maintenance and product purchasing.
- Performed office duties including reception, processing and shipping/delivering orders, and account management.

1999 – 2000 (Part time)
Laboratory technician
Dr. I. Altosaar, University of Ottawa

- Responsible for general laboratory maintenance, product purchasing, and media and reagent preparation.
- Assisted with research of graduate students in various areas of plant molecular biology.

1999
Careers Canada summer research award recipient
Dr. I. Altosaar, University of Ottawa

- Learned basic molecular biology principals and assisted graduate student and post-doctoral fellows with various research projects involving transgenic tobacco and rice.
- Responsible for general laboratory maintenance, journal club presentations, and product purchasing.

Related Work Experience

Winter 2008 and 2009
Teaching assistant of a general biochemistry undergraduate course
University of Ottawa, Faculty of Science

- Proctored and corrected examinations.
- Presented 12 discussion sessions (2 hours each) and review sessions explaining course materials.
- Hosted weekly online discussion group to answer student questions regarding course content.

Fall 2007
Teaching assistant of a macromolecules undergraduate course
University of Ottawa, Faculty of Science

- Corrected examinations and provided office hours for discussion of course materials.

Fall 2001 and 2002
Teaching assistant of a general microbiology undergraduate course
University of Ottawa, Faculty of Science

- Assisted lesson and test preparation, maintained course web page, and answered student questions.
- Evaluated tests, entered and submitted grades, and proctored examinations.

1997-1998
Co-operative education student and hospital technician
Navan Veterinary Hospital

- Responsible for general hospital maintenance, animal care, and customer service/administration.
- Laboratory technician with duties including blood, skin, and urinalysis, bacterial staining and culture, microscopy of biological samples, serological analysis, automated endocrinological analysis, and antimicrobial susceptibility testing.

Other Skills

- ***Microbiology:*** culture, isolation, and identification of bacterial species using biochemical and growth-dependant assays as well as commercially-available products, preparation of culture media, antimicrobial susceptibility testing, light, fluorescence, and phase-contrast microscopy.
- ***Molecular biology:*** DNA and RNA isolation, purification, and electrophoresis (agarose and polyacrylamide gel), polymerase chain reaction (PCR), reverse transcriptase PCR, inverse PCR, quantitative PCR, molecular cloning and plasmid

construction, *in vitro* transcription, electrophoretic mobility shift assay (EMSA), protein SDS-PAGE electrophoresis, Northern, Southern, and Western blot analysis, antibody purification, protein purification (IMAC, FPLC), oligonucleotide and peptide synthesis, yeast and bacterial two-hybrid analyses, flow cytometry, immunoprecipitation (*in vitro* and *in vivo*), basic cell culture.

- ***Computer:*** Windows operating systems, Microsoft Word, Powerpoint, Excel, Wordperfect, Corel Draw, Netscape Navigator and Internet Explorer, web-based bioinformatic tools (BLAST, sequence analysis, restriction mapping, multiple alignment, phylogenetic tree construction, mFold), protein structural analysis (CN3D, Rasmol, Protein Explorer, Swiss PDB), molecular biology software (Primer Design).

Memberships and Affiliations

- Golden Key Society (for academic excellence)
- Canadian Society of Microbiologists
- American Society for Microbiology
- American Society for Virology
- RNA Society

Related Academic Activities

- Biochemistry, Microbiology, and Immunology Graduate Students' Association Vice President of Internal Affairs (2007)
- Biochemistry, Microbiology, and Immunology Graduate Students' Association Vice President of Academic Affairs (2006)
- Biochemistry, Microbiology, and Immunology bi-monthly bulletin contributor and temporary Editor-in-Chief (2007)
- Participant in the RNA Club of the University of Ottawa

Peer-Reviewed Publications

Sikora, D., **Greco-Stewart, V.**, Miron, P., and M. Pelchat. 2009. The hepatitis *delta* virus RNA genome interacts with eEF1A1, $p54^{nrb}$, hnRNP-L, GAPDH and ASF/SF2. Virology **390**: 71-78.

Greco-Stewart, V., Schissel, E., and M. Pelchat. 2009. The hepatitis *delta* virus RNA genome interacts with the human RNA polymerases I and III. Virology **386**: 12-15.

Greco-Stewart, V., Ramirez-Arcos, S., Liao, M., and J. R. Dillon. 2007. N-terminal residues of MinC from *Neisseria gonorrhoeae* are required for cell division function and mediate interaction with FtsZ. Arch. Microbiol. **187**: 451-458.

Greco-Stewart, V., Miron, P., Abrahem, A., and M. Pelchat. 2007. The human RNA polymerase II interacts with the terminal stem-loop regions of the hepatitis *delta* virus RNA genome. Virology **357**: 68-78.

Greco-Stewart, V., St-Laurent Thibault, C., and M. Pelchat. 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to the hepatitis *delta* virus RNA. Virology **356**: 35-44.

Ramirez-Arcos, S., **Greco, V.,** Douglas, H., and Dillon, J.R. 2004. Conserved glycines in the C-terminus of MinC proteins are implicated in their functionality as cell division inhibitors. J. Bacteriol. **186**: 2841-2855.

Greco, V., Ng, L.K., Catana, R., Li, H., and Dillon, J.R. 2003. Molecular Epidemiology of *Neisseria gonorrhoeae* Isolates with Plasmid-mediated Tetracycline Resistance (TRNG) in Canada: Temporal and Geographical Trends (1986-1997). Microb. Drug Resist. 9: **353**-360.

Lariviere, K., MacEachern, L., **Greco, V.,** Majchrzak, G., Chiu, S., Drouin, G., and Trudeau, V. 2002. GAD_{65} and GAD_{67} isoforms of the glutamic acid decarboxylase gene originated before the divergence of cartilaginous fish. Mol. Biol. Evol. **19**: 2325-2329.

Manuscripts Submitted and in Preparation

Greco-Stewart, V. and M. Pelchat. 2009. RNA-templated transcription by DNA-dependent RNA polymerases. Review. *In preparation.*

Greco-Stewart, V. and M. Pelchat. 2009. HDV: host-pathogen interactions. Review. *In preparation.*

Conference Posters and Presentations

Sikora, D., **Greco-Stewart, V.,** Miron, P., and M. Pelchat. 2008. Interactions of the hepatitis *delta* virus RNA with various host proteins and potential roles of these interactions in viral pathogenesis. 9th Annual Ribo-Club Opening Session. Mont-Orford, Québec, Canada. (Poster)

Al-Ali, Y., **Greco-Stewart, V.,** and M. Pelchat. 2008. Replication of the hepatitis *delta* virus requires the binding of both the RNA polymerase II and polypyrimidine tract-binding protein-associated splicing factor. 9th Annual Ribo-Club Opening Session. Mont-Orford, Québec, Canada. (Poster)

Greco-Stewart, V. and M. Pelchat. 2007. RNA from the hepatitis *delta* virus interacts specifically with components of the host transcriptional machinery. 8th Annual Ribo-Club Opening Session. Mont-Orford, Québec, Canada. (Poster)

Greco-Stewart, V. and M. Pelchat. 2007. The multifunctional nuclear protein PSF mediates binding of host DNA-dependent RNA polymerase II to hepatitis *delta* virus RNA. American Society for Microbiology 107th Annual General Meeting. Toronto, Canada. (Poster)

Greco-Stewart, V. and M. Pelchat. 2006. RNA polymerase II and the nuclear protein PSF bind specifically to the terminal stem-loop domains of both polarities of hepatitis *delta* virus RNA. University of Ottawa RNA Club. (Oral presentation)

Greco-Stewart, V. and M. Pelchat. 2006. RNA polymerase II and the nuclear protein PSF bind specifically to the terminal stem-loop domains of both polarities of hepatitis *delta* virus RNA. 7th Annual Ribo-Club Opening Session. Mont-Orford, Québec, Canada. (Poster)

Greco-Stewart, V. and M. Pelchat. 2006. Binding of the polypyrimidine tract-binding protein-associated splicing factor (PSF) to hepatitis *delta* virus RNA. RNA Society Annual Meeting. Seattle, U.S.A. (Poster)

Greco-Stewart, V. and M. Pelchat. 2006. RNA from the hepatitis *delta* virus interacts with the multifunctional nuclear protein PSF. Canadian Society of Microbiology 56th Annual Conference. London, Canada. (Poster)

Greco-Stewart, V. and M. Pelchat. 2006. RNA promoters from the hepatitis *delta* virus interact with RNA polymerase II and the multifunctional nuclear protein PSF. University of Ottawa RNA Club. (Oral presentation)

Greco-Stewart, V. and M. Pelchat. 2005. RNA promoters from the hepatitis *delta* virus interact with RNA polymerase II and the multifunctional nuclear protein PSF. 6th Annual Ribo-Club Opening Session. Mont-Orford, Québec, Canada. (Poster)

Greco, V., L. K. Ng, R. Catana, H. Li, and J. R. Dillon. 2003. Molecular epidemiology of *Neisseria gonorrhoeae* isolates with plasmid-mediated tetracycline resistance (TRNG) in Canada: Temporal and geographical trends (1986-1997). 2003 ISSTDR Congress. Ottawa, Canada. (Poster)

Ramirez-Arcos, S., **V. Greco**, H. Douglas, and J. R. Dillon. 2002. Four conserved glycines in the C-terminus of MinC are essential for protein functionality. International Pathogenic *Neisseria* Conference. Madrid, Spain. (Poster)

Lariviere, K., L. MacEachern, **V. Greco**, and V. L. Trudeau. 2000. Multiple forms of GAD likely arose prior to the divergence of Chondrichtheys. 20th Conference of European Comparative Endocrinology. Faro, Portugal. (Poster)

Greco, V. and I. Altosaar. 1999. Expression of antimicrobial peptides from *Sarcophaga peregrina* in transgenic rice. International Rice Research Institute Annual Conference. Ottawa, Canada. (Oral presentation).

Awards, Bursaries, and Scholarships

Name of award	Value	Level of award	Location of tenure	Period held
Canadian Blood Services Scholarship	44 000	National	U. of Ottawa	05/2007 – 04/2009
Faculty of Medicine Excellence Award	500[a]	Institutional	U. of Ottawa	12/2006
University of Ottawa Excellence Award	5 535[b]	Institutional	U. of Ottawa	05/2005 – 04/2009
NSERC CGS Award	70 000	National	U. of Ottawa	05/2005 - 04/2007
Ontario Graduate Scholarship	15 000	Provincial	U. of Ottawa	Declined (05/2005)
FGPS Admission Scholarship	5 535[b]	Institutional	U. of Ottawa	09/2004 - 04/2005
Ontario Graduate Scholarship	15 000	Provincial	U. of Ottawa	Declined (09/2003)
NSERC PGS-A Award	34 600	National	U. of Ottawa	09/2001 - 12/2003
Strategic Areas of Development Award	6 000	Institutional	U. of Ottawa	09/2001
University of Ottawa Excellence Award	5 535[b]	Institutional	U. of Ottawa	09/2001 - 12/2003
Ontario Graduate Scholarship	15 000	Provincial	U. of Ottawa	Declined (09/2001)
NSERC Undergraduate Award	5 500	National	U. of Ottawa	05/2001 - 08/2001
Dean's Merit Scholarship	2 000	Institutional	U. of Ottawa	09/2000 - 04/2001
NSERC Undergraduate Award	5 000	National	U. of Ottawa	05/2000 - 08/2000
Dean's Merit Scholarship	1 000	Institutional	U. of Ottawa	09/1999 - 04/2000
Careers Canada Scholarship	5 000	National	U. of Ottawa	05/1999 - 08/1999
Admission Scholarship	2 000	Institutional	U. of Ottawa	09/1997 - 04/1998

a) Award for most outstanding student in the Ph.D. Biochemistry program
b) Value indicated is per annum for tuition.

Additional notes:

- Nominated for best M.Sc. thesis (2004), University of Ottawa, Faculty of Medicine.
- Eureka! Award for innovation and discovery as administered by the departmental chairperson for the poster entitled "Characterization of an antigenomic promoter from the hepatitis *delta* virus" (2005). Graduate Student Poster Presentation, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa.
- Contributed to an episode of American Society for Microbiology's Microbe World web video series (filmed at ASM's 107th Annual General Meeting 2007).
- Nominated for best Ph.D. thesis (2009), University of Ottawa, Faculty of Medicine.