

2000

An Alternative View of the Templates and Functions of RNA Polymerase II

Julija Filipovska

Follow this and additional works at: http://digitalcommons.rockefeller.edu/student_theses_and_dissertations



Part of the [Life Sciences Commons](#)

Recommended Citation

Filipovska, Julija, "An Alternative View of the Templates and Functions of RNA Polymerase II" (2000). *Student Theses and Dissertations*. 355.
http://digitalcommons.rockefeller.edu/student_theses_and_dissertations/355

This Thesis is brought to you for free and open access by Digital Commons @ RU. It has been accepted for inclusion in Student Theses and Dissertations by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.



An Alternative View of the Templates and Functions of RNA Polymerase II

**A thesis presented to the faculty of The Rockefeller University
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy**

**by
Julija Filipovska**

**October 2000
The Rockefeller University
New York, New York**

To the memory of my grandparents

ACKNOWLEDGMENTS:

I would like to thank Dr. Magda Konarska, for giving me the opportunity to study a very intriguing and peculiar problem in her lab. I have valued her expertise and supervision, which pointed out directions and at the same time nurtured the space for personal expression. Her scientific approach has been, and will surely continue to be, an inexhaustible source of inspiration.

I am also grateful to the past and present members of the Konarska lab, for the inspirational discussions, practical help, and for the colorful memories of the time we spent together in the lab: Aurel Betz, Marco Gottardo, Hilary Gustafson, Naima Ismaili, Marek Juszczuk, Pavol Kois, Sei Kameoka, Boyana Konforti, Tao Levy, Lyle Najita, Beata Nowakowska, Jose ("Pepe") Reyes, Berthold Rutz, Ralf Schmauder, Ma Sha and Mirka Siatecka.

Special thanks to the members of the "Bronk 7th floor" for generously sharing their equipment and knowledge, and creating an excellent working environment that I will always miss.

I am indebted to the members of my thesis committee Drs. Titia de Lange, Peter Model, David Thaler, and Jean-Pierre Perreault for their help and guidance. Their genuine interest in my work has been invaluable for the progress of my scientific learning.

I also want to thank The Rockefeller University for a financial support that has allowed me the freedom to explore my own choice of paths and The Deans' office for their patience and help at numerous occasions during my studies.

With great joy, I would like to say **"Thanks !!! Gracias !!! Hvala !!! Blagodaram !!! OM Shanti"** to an exceptional 'bunch' of friends for recognizing their many names in these few words, and helping create the energy essential for me to finalize this work and much, much more.

Lastly, but mostly, I want to thank my parents, Jovan and Stojka, and all my family, who with their love and constant support make things for me possible in so many ways.

Table of contents

List of Figures	viii
List of Abbreviations	x
<u>Abstract</u>	1
1. INTRODUCTION	3
1.1 REGULATION OF POL II ACTIVITY ON THE ‘CLASSICAL’ DNA TEMPLATE	5
1.1.1 Preinitiation and initiation steps	5
1.1.2 Post-initiation steps	8
<i>Examples of elongation control consistent with the RNA-dependent pol II function</i>	10
1.2 RNA PATHOGENS THOUGHT TO BE REPLICATED BY THE HOST RNA-POLYMERASE II	13
1.2.1 Plant Viroids	13
1.2.2 Human Hepatitis Delta Virus	19
1.2.3 Studies HDV RNA synthesis in vivo	23
<i>Protein factors requirements</i>	24
<i>RNA template requirements</i>	28
1.2.4 Studies HDV RNA synthesis in vitro	23
2. RESULTS	33
2.1 HDV RNA –TEMPLATED TRANSCRIPTION BY POL II IN VITRO	34
2.1.1 Description of a system to study HDV RNA-templated transcription in vitro	34
<i>Deletion analysis of the AG HDV RNA template</i>	39

2.1.2	Sequence analysis of the product of HDV RNA-templated transcription in vitro	4 4
	<i>RNase digestion analysis</i>	4 5
	<i>Analysis of the 5' end of the HDV RNA pol II transcript</i>	5 2
	<i>RNase H analysis: Demonstrating that NE RNA product is a chimeric molecule</i>	6 1
	<i>Sequence determination of the template/transcript junction</i>	6 4
	<i>Analysis of the 3' end of the pol II transcript</i>	7 6
2.1.3	Effect of hepatitis delta antigen (HDAg) on transcription in vitro	8 5
2.1.4	Template requirements for the HDV RNA-templated pol II transcription	9 2
	<i>Secondary structure of the RNA template determines functional pol II templates</i>	9 2
	<i>Mapping of the secondary structure of the HDV RNA templates</i>	1 0 2
2.2	RNA SECONDARY STRUCTURE REQUIREMENTS FOR HDV REPLICATION IN VIVO	1 1 0
2.2.1	Description of a system to study RNA template requirements for HDV replication in vivo	1 1 0
2.2.2	In vivo replication of HDV mutants containing RNA sequence/structure alterations that modulate pol II transcription in vitro	1 1 9
	<i>Replication of wild type HDV RNA in the COS7 cells</i>	1 1 9
	<i>Replication of the HDV15, HDV18 and HDV63 mutants in COS7 cells</i>	1 2 0
	<i>Replication of wild type, and HDV15, HDV18, and HDV63 mutants in HeLa cells</i>	1 2 6
	<i>Replication of HDV73, HDV75, HDV65, HDV14, and HDV20 mutants in COS7 cells</i>	1 3 0
	<i>Replication of HDV73, HDV75, HDV65, HDV14 and HDV20 mutants in HeLa cells</i>	1 3 6

3.	DISCUSSION	138
3.1	Does HDV RNA represent a specific template for pol II-mediated transcription?	139
3.2	Initiation of pol II-mediated transcription on the HDV RNA template in vitro	142
3.3	Fidelity and processivity of the pol II-mediated HDV RNA transcription in vitro and the effect of delta antigen	146
3.4	Specific secondary structure of the HDV RNA template determines the specificity and efficiency of the HDV RNA-templated pol II transcription in vitro	160
3.5	Effects of alterations of the RNA secondary structure on HDV replication in vivo- a template switching model	159
4.	CONCLUSIONS	173
5.	MATERIALS AND METHODS	178
5.1	Recombinant DNA technology	179
5.2	Plasmids and construction of mutants for in vitro analysis	179
5.3	Plasmids and construction of mutants for in vivo analysis	180
5.4	Preparation of RNA templates for HeLa NE transcription	181
5.5	Preparation of control RNA for the RNase digestion analyses	182
5.6	Preparation of 5' end labeled RNA for secondary structure mapping	183
5.7	Transcription reactions in HeLa NE	183
5.8	Phosphatase and β-elimination treatment of the RNA	184
5.9	RNase T1 and RNase A digestion	185
5.10	RNase H analysis	185
5.11	RNA secondary structure predictions	186
5.12	Enzymatic mapping of RNA secondary structure	186
5.13	Lead cleavage mapping of RNA secondary structure	187
5.14	Cells and transfections	187
5.15	RNA isolation from transfected cells	188
5.16	Detection of replicating HDV RNA by hyb-shift assay	188
5.17	Quantification of the relative replication efficiency of HDV mutants	189
6.	BIBLIOGRAPHY	190

List of Figures

Chapter 1

Figure 1	RNA pathogens thought to be replicated by pol II.	17
-----------------	---	-----------

Chapter 2

Figure 2.	Antigenomic (AG) HDV RNA represents a template for pol II-mediated RNA synthesis.	36
Figure 3.	Deletion analysis of HDV RNA segment required as a template for NE transcription.	41
Figure 4.	Transcription using the smallest HDV RNA templates encompassing the left-hand terminal hairpin is also sensitive to α -amanitin.	43
Figure 5.	The HDV RNA-templated pol II transcription is precise—RNase T1 analysis of the transcription products.	47
Figure 6.	The HDV RNA-templated pol II transcription is precise—RNase A analysis of the transcription products.	49
Figure 7.	Absence of a triphosphate or a cap structure at the 5' end of the RNA-templated pol II transcript.	54
Figure 8.	Detection of a 5' cap structure on a synthetic RNA transcript.	56
Figure 9.	Detection of a hydroxyl group at the 5' end of the NE RNA transcript.	58
Figure 10.	The NE product of the pol II-mediated HDV RNA transcription represents a chimeric template/transcript molecule.	63
Figure 11.	Determination of the sequence at the template/transcript junction in NE RNA products.	68
Figure 12.	The template/transcript junction point is identical in all mutant NE products.	70
Figure 13.	The adenosine residues within the template/transcript junction sequence are not incorporated by the pol II transcription.	72
Figure 14.	Pol II initiates HDV RNA-templated transcription at a unique site.	74
Figure 15.	Mapping the 3' end of the pol II transcript by RNase T1 digestion.	78

Figure 16.	Mapping the 3' end of the pol II transcript by RNase A digestion.	80
Figure 17.	Full sequences of the AG103 template and the product generated by the pol II-mediated transcription in HeLa NE.	84
Figure 18.	Delta antigen stimulates the HDV RNA-templated pol II transcription in vitro.	88
Figure 19.	Delta antigen stimulates pol II elongation on the HDV RNA template.	90
Figure 20.	Specific secondary structure of HDV RNA, rather than its primary sequence, determines functional pol II templates.	97
Figure 21.	Compilation of secondary structures of AG HDV RNA templates.	99
Figure 22.	Transcription reactions in NE using structurally distinct RNA templates.	101
Figure 23.	Enzymatic mapping of the secondary structure in the terminal hairpin of the HDV RNA templates.	104
Figure 24.	Lead cleavage mapping of the secondary structure in the terminal hairpin of the HDV RNA templates.	106
Figure 25.	Description of the strategy used to analyze the effects of RNA secondary structure alterations on HDV replication in vivo.	112
Figure 26.	Specificity of the hyb-shift assay.	114
Figure 27.	Replication of wild type (wt) and genomic ribozyme mutant (GR ⁻) HDV in COS7 cells.	116
Figure 28.	Replication of HDV15, HDV18, and HDV63 in COS7 cells.	122
Figure 29.	Replication of the HDV15 and HDV18 in COS7 cells when delta antigen is overexpressed.	124
Figure 30.	Replication of the HDV15, HDV18, and HDV63 in HeLa cells.	128
Figure 31.	Replication of HDV73, HDV75, HDV65, HDV15, and HDV20 in COS7 cells.	132
Figure 32.	Replication of HDV73, HDV75, HDV65, HDV15, and HDV20 in HeLa cells.	134

Chapter 3

Figure 33.	Schematic representation of DNA- and HDV RNA-templated pol II transcription.	150
Figure 34.	Template switching model for the HDV RNA replication.	163
Figure 35.	Description of two different approaches designed to examine if pol II RNA template switching can occur in vitro.	170

List of Abbreviations

aa	amino acid
Ab	antibody
AG	antigenomic polarity
ATP	adenosine triphosphate
bp	base pair
BC100	buffer C, 100mM salt
°C	degrees Celsius
cDNA	complementary DNA
Ci	Curie (1Ci = 3.7×10^{10} Bequerel)
cm	centimeter
CTD	carboxy-terminal domain
CTP	cytidine triphosphate
CycT1	cyclin T1
ddH ₂ O	double distilled water
GFP	green fluorescent protein
dGTP	deoxyguanosine triphosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DRB	5,6-dichloro-1-b-D-ribofuranosylbenzimidazole
ds	double stranded
DTT	dithiothreitol
EDTA	ethylendiaminetetraacetic acid
EtBr	ethidium bromide
EtOH	ethanol
FA	formamide
FACS	Fluorescence Activated Cell Sorting
FACT	facilitates chromatin transcription
fmol	femtomole
FT	flow through
G	genomic polarity
GTF	general transcription factor
GTP	guanosine triphosphate
HDAg	hepatitis delta antigen

HDV	hepatitis delta virus
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HIV	Human Immunodeficiency Virus
hr	hour
kb	kilobase
KCl	potassium chloride
kDa	kilodalton
LiCl	lithium chloride
LTR	Long Terminal Repeat
M	molar
mAb	monoclonal antibody
μg	microgram
mg	milligram
Mg ²⁺	magnesium ion
MgCl ₂	magnesium chloride
μl	microliter
ml	milliliter
μM	micromolar
mM	millimolar
Mn ²⁺	manganese ion
mRNA	messenger RNA
NaAc	sodium acetate
NaCl	sodium chloride
NE	nuclear extract
NELF	negative elongation factor complex
ng	nanogram
NH ₄ Ac	ammonium acetate
NH ₄ OH	ammonium hydroxide
Ni ²⁺	nickel ion
nm	nanometer
nt	nucleotide
NTP	nucleoside triphosphate
P	product
PAGE	polyacrylamide gel electrophoresis
pmol	picomole
PMSF	phenylmethanesulfonylfluoride

RNA	ribonucleic acid
RNAse	ribonuclease
p-TEFb	positive transcription elongation factor b
rNTP	ribonucleoside triphosphate
rRNA	ribosomal RNA
SDS	sodium dodecyl sulfate
snRNA	small nuclear RNA
ss	single stranded
syn.	synthetic
sym.	symmetric
T	template
TAR	trans-activation response element
Tat	trans activator of transcription
TE	Tris/EDTA
TFII	pol II transcription factor
Tris	Tris-(hydroxymethyl)-aminomethane
tRNA	transfer RNA
txn	transcription
UTP	uridine triphosphate
UV	ultraviolet
wt	wild type
x	fold

... I say that it is not illogical to think that the world is infinite. Those who judge it to be limited postulate that in remote places the corridors and stairways and hexagons can conceivably come to an end—which is absurd. Those who imagine it to be without limit forget that the possible number of books does have such a limit. I venture to suggest this solution to the ancient problem: The Library is unlimited and cyclical. If an eternal traveler were to cross it in any direction, after centuries he would see that the same volumes were repeated in the same order (which, thus repeated, would be an order: the Order). My solitude is gladdened by this elegant hope.

from *The Library of Babel* by J. L. Borges
(translated by J. E. L.)

Abstract

RNA polymerase II (pol II) has been implicated in the RNA-templated RNA synthesis during replication of some Viroids and Hepatitis Delta Virus (HDV). In this study HDV RNA-templated pol II transcription was examined in vitro, using HeLa cell nuclear extracts (NE). Under standard conditions for the DNA-templated pol II transcription, a segment of the antigenomic (AG) HDV RNA encompassing the left-hand tip region of the HDV rod-like structure, serves as a template for efficient transcription. The transcription reaction is highly sensitive to α -amanitin in HeLa cell NE, but it is partially resistant to this toxin in NE from PMG cells that contain an α -amanitin resistant allele of the pol II largest subunit, strongly suggesting pol II involvement in the process. Direct RNase A, T1 and RNase H digestion analyses demonstrate that the product of the HDV RNA-dependent pol II transcription represents a chimeric molecule, in which the pol II transcript is covalently attached to the 5' half of the AG RNA template. Such a chimeric RNA product is generated by a cleavage of the template at a specific site followed by transcription that uses the new end as a primer. Pol II transcription of the RNA template proceeds with high fidelity but appears to pause after ~40 transcribed nucleotides. A viral protein, delta antigen (HDAg), stimulates pol II elongation on the RNA template. The initiation of HDV RNA transcription that involves cleavage of the RNA template may be mechanistically similar to the endonucleolytic cleavage of the nascent transcript that, under

certain conditions, occurs during the elongation phase of pol II transcription on a 'typical' DNA templateS.

Mutational and secondary structure mapping analyses of HDV RNA templates demonstrate that the secondary structure, rather than the primary sequence, specifies the functional pol II transcription templates. The secondary structure in the left-hand terminal hairpin of the HDV RNA also modulates HDV replication in vivo, in COS7 and HeLa cells. The correlation between the effects of secondary structure alterations on the efficiency of pol II transcription in vitro and HDV replication in vivo suggests that the observed RNA-templated pol II transcription in vitro reflects an important regulatory step of the HDV replication cycle. Based on these findings, a modified pol II-mediated rolling-circle model for HDV replication is proposed. It involves initiation of transcription by cleavage of one RNA template, followed by pol II switching to a new circular template. The model also suggests that, both the initiation and the template switching steps are regulated by specific secondary structure elements in the terminal hairpin of HDV RNA.

Chapter 1

INTRODUCTION

The initial steps in the complex process of gene expression that unravels the functional aspects of the genetic information involve RNA synthesis, or transcription, of the DNA or RNA genomes. In particular, RNA-templated RNA synthesis is essential to the replication of many viruses that contain RNA genomes and do not reverse transcribe. Thus, it is important to examine the components and the underlying regulatory mechanisms that control RNA synthesis and gene expression.

RNA synthesis in eukaryotes is carried out by three complex multisubunit enzymes, RNA polymerases I, II, and III (Roeder 1976). Each polymerase, in concert with a number of additional regulatory protein factors, recognizes specific promoter elements and transcribes functionally distinct DNA sequences. In particular, RNA polymerase II (pol II) responsible for transcription of the protein-coding pre-mRNAs, has been extensively studied both in terms of DNA template requirements and protein factors that aid and regulate its activity. Interestingly, pol II has been also implicated in replication of two classes of RNA pathogens: some plant viroids and human Hepatitis Delta Virus (HDV) that do not go through a DNA phase during their life cycle (Lai 1995). The concept of the RNA-templated pol II function may have a wide range of implications for the regulation of the RNA metabolism and gene expression in eukaryotic cells. Therefore, it is of a considerable importance to closely examine pol II involvement in the RNA-templated RNA synthesis and determine the template requirements and the protein factors involved in the process.

The work presented here was undertaken with the aim of addressing some aspects of the issues mentioned above. In particular, I have used HDV RNA as a model system to examine if pol II can precisely transcribe an RNA template, and, if so, what are the sequence and/or structural characteristics of an RNA molecule that represents a template for pol II transcription. These questions were addressed using an in vitro transcription system that supports pol II-mediated HDV RNA synthesis, developed in Dr. M. M. Konarska lab. Finally, I have examined the correlation between the RNA template requirements for efficient pol II-mediated RNA synthesis in vitro and HDV replication in the cells.

1.1 REGULATION OF POL II ACTIVITY ON THE 'CLASSICAL' DNA TEMPLATE

1.1.1 Preinitiation and initiation steps

RNA polymerase II, the enzyme implicated in the RNA-dependent RNA synthesis of viroid and HDV, has been extensively studied as a DNA-dependent RNA polymerase. It was first identified as the one of the three RNA polymerase activities able to support promoter independent transcription on nonspecific DNA templates that was the most sensitive to the toxin α -amanitin (Roeder and Rutter, 1969). The specificity and efficiency of pol II is determined by complex interactions of the multisubunit enzyme with specific DNA sequences, general

and gene-specific transcription factors, and the nascent RNA transcript. Such complexity provides multiple regulatory points for gene expression during each of the basic phases of transcription: promoter recognition or preinitiation, initiation, elongation, and termination. However, over the past 20 years, biochemical studies of eukaryotic mRNA synthesis have largely focused on the preinitiation and initiation stages of transcription, mostly because of the availability of in vitro systems in which pol II supplemented with the appropriate general transcription factors (GTFs) can support low levels of accurate (basal) transcription (Weil et al., 1979; Matsui et al., 1980; Roeder, 1996). The common core promoter elements that serve to nucleate the initiation complex in these systems are represented by the TATA box approximately located at position –30 to –25, and a pyrimidine-rich initiator (Inr) element located near the transcription start site. The general initiation factors include TFIIB, TFIID, TFIIE, TFIIIF, and TFIIH each of which functions in intimate association with RNA polymerase II, and is required for selective binding of pol II to the promoter, formation of the open complex, and synthesis of the first few phosphodiester bonds of nascent transcripts (Orphanides et al., 1996; Roeder, 1996). More specifically, recognition of the TATA box is accomplished by the TATA binding protein, TBP (Parker and Topol, 1984; Burley and Roeder, 1996; Nakajima et al., 1988; Horikoshi et al., 1989; Horikoshi et al., 1992). Binding of TBP to the minor groove of the TATA element induces a sharp bend in the DNA that results in partial unwinding of the TATA sequence (Kim et al., 1993a; Kim et al., 1993b; Kim and Burley, 1994; Burley and

Roeder, 1996; Nikolov et al., 1996). In vivo, TBP most likely operates in concert with a tightly associated set of transcription factors called TBP-associated factors (TAFs), that together with TBP form the TFIID complex (Roeder and Rutter, 1969; Orphanides et al., 1996). Formation of the preinitiation complex (PIC) proceeds by recruitment of RNA pol II /TFIIF, followed by the incorporation of TFIIE, which in turn recruits TFIIH (Buratowski et al., 1989; Orphanides et al., 1996; Roeder, 1996). TFIIA and TFIIB are thought to play an important role early during PIC formation by helping to recruit or stabilize the binding of TFIID to the core promoter. In contrast, TFIIE and TFIIH participate during the later stages of transcription, that include promoter-melting and modifications of the stable PIC that allow the promoter clearance and elongation (Orphanides et al., 1996; Roeder, 1996; Nikolov and Burley, 1997). In fact, the addition of TFIIH renders the polymerase competent to initiate transcription by phosphorylating the carboxy-terminal domain (CTD) of pol II, that consists of multiple copies of the heptapeptide repeats with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (Feaver et al., 1991; Lu et al., 1992; Dahmus, 1995; Roeder, 1996; Shiekhhattar et al., 1995; Orphanides et al., 1996). Phosphorylation of the pol II CTD tail by a variety of kinases appears to be important not only for the efficiency of the early and late steps of transcription (Conaway et al., 2000; Meininghaus et al., 2000), but also for coupling transcription with pre-mRNA processing (McCracken et al., 1997; McCracken et al., 1998; Misteli and Spector, 1999).

The CTD appears to be also involved in the assembly of a large multisubunit pol II holoenzyme complex containing a subset of the GTFs and a number of other protein factors (depending on the method of preparation and the transcriptional system used) (Kim et al., 1994; Koleske and Young, 1994; Maldonado et al., 1996). Such a preassembled complex can bind to a promoter in a single step and is able to support activated transcription in a variety of transcription systems (Koleske and Young, 1994; Bjorklund and Kim, 1996; Struhl, 1996), and may in fact represent the authentic mode of pol II initiation in vivo.

Regardless of the precise mechanism for transcription initiation in vivo, the in vitro systems used to study the step-wise assembly of the transcriptional machinery on the DNA promoter are still extremely useful for the study of fundamental aspects of the process that represents a prerequisite for understanding the superimposed regulatory mechanisms.

1.1.2 Post-initiation steps

The importance of simplified in vitro systems for studying complex processes that can involve multiple levels of control becomes especially evident when the understanding of the regulation of the post-initiation steps is considered. Although transcription elongation has been recognized as an important part of the control of gene expression (Yankulov et al., 1994; Krumm et al., 1995; Uptain et al., 1997), efforts to elucidate the mechanisms by which specific elongation

factors regulate the process have been for a long time hampered by the lack of suitable assays. This is largely due to the fact that, unlike most regulatory proteins involved during the initiation steps that recognize specific template sequences, positive and negative elongation factors appear to regulate transcription in a sequence-independent manner. For example, FACT and SWI/SNF protein family members promote efficient pol II elongation through the chromatin structures that negatively affect transcription. Other factors, like TFIIF, TFIIS, P-TEFb, Elongin, and Tat, suppress transient pausing and arrest by directly targeting pol II or protein factors that negatively regulate elongation (NELF and DSIF) (reviewed in Uptain, 1997; Conaway et al., 2000 and references therein). This kind of classification of the factors involved in the regulation of transcription elongation is based on the initial studies that led to their identification, and it may become less significant as new interactions and functional connections become apparent.

Unlike the promoter elements, the DNA template sequences responsible for pol II pausing and/or arrest are much less characterized, even though sites that intrinsically block the progression of the three eukaryotic polymerases have been described (Hawley and Roeder, 1985; Reines et al., 1987; Lis and Wu, 1993; Matsuzaki et al., 1994; Schnapp et al., 1996). The secondary structure of the nascent transcript has also been implicated in regulation of pol II elongation (Sharp and Marciniak, 1989; Reeder and Hawley, 1996; Bengal and Aloni, 1998). The involvement of various hairpin structures in the nascent transcript in the

regulation of transcription has been clearly demonstrated in prokaryotes (Lazinski et al., 1989; Mason et al., 1992; Landick et al., 1996; Richardson and Greenblatt, 1996). However, in eukaryotes, this aspect of transcriptional regulation remains to be analyzed in greater detail.

The identification of the protein factors that regulate pol II elongation and the availability of elongation-specific inhibitors (e.g. DRB or 5,6-dichloro-1- β -D-ribofuranosylbenzimidazole) (Yankulov et al., 1995; Marshall et al., 1996), facilitate further studies of this important stage of transcription regulation. The availability of both, the highly purified and recombinant elongation factors allows for the development of reconstituted systems that recapitulate the elongation stage of transcription and allow for biochemical analyses that help identification of rate-limiting steps in the process and formulation of models that can be tested *in vivo*.

Examples of elongation control consistent with the RNA-dependent pol II function

In the context of the work presented in this thesis, it is important to further discuss the mechanism of regulation of the early post-initiation steps by the NELF/DSIF-P-TEFb system, and the reactivation of pol II arrested complexes by TFIIS-stimulated transcript cleavage.

A successful promoter clearance phase requires phosphorylation of the pol II CTD tail by the GTFs and synthesis of short (~20-40 nt) transcripts. However, following this phase, pol II processivity appears to be blocked by the

action of the negative elongation factor, NELF, and the DRB sensitivity inducing factor, DSIF (Wada et al., 1998a; Yamaguchi et al., 1999). The negative effects of these factors is alleviated by a positive transcription elongation factor b, P-TEFb, that hyperphosphorylates the CTD domain of pol II and in that way stimulates transcription (Marshall et al., 1996; Wada et al., 1998b; Yamaguchi et al., 1999). The DRB sensitive P-TEFb kinase consist of two subunits, the CDK9 kinase and the regulatory cyclin T (CycT) subunit. The contribution of specific DNA template or RNA transcript sequences to the interplay of the negative and positive elongation factors is not clear, although differential sensitivity of transcription to the NELF has been observed using two different promoter systems (Yamaguchi et al., 1999).

Transcription from the HIV-1 LTR promoter is one of the well characterized examples of regulation of gene expression through control of post-initiation steps. It is also an example in which the importance of the secondary structure of the nascent transcript has been clearly demonstrated (Kao et al., 1987; Muesing et al., 1987; Ossipow et al., 1995 and reviewed in Garber and Jones 1999 and Karn 1999). In fact, it has been suggested that the 5' segment of the nascent transcript forms distinct hairpin structures that control precise interplay between the effects of negative and positive transcription factors. The model based on a collection of recent studies proposes that pol II complex is blocked after the promoter clearance due to the combined action of DSIF and NELF (Garber and Jones, 1999; Yamaguchi et al., 1999). The negative

elongation factors may be recruited through the interaction of a putative RNA binding motif in one of the subunits of the NELF with the nascent transcript (Yamaguchi et al., 1999). The purified pol II also pauses early post initiation, when a specific hairpin structure, a "pause" hairpin, can be formed in the 5' segment of the nascent transcript (Palangat et al., 1998). The pause signal is counteracted by the formation of a slightly longer but structurally distinct TAR (trans-activation response) hairpin that recruits the positive elongation factors, the viral trans-activating protein Tat, and the CycT1 component of the P-TEFb complex. As a result of the cooperative binding of these proteins to the TAR hairpin, the CDK9 kinase component of the P-TEFb complex is positioned in the proximity of the pol II CTD tail that allows for its efficient phosphorylation. Thus, the "pause" hairpin-induced pol II pausing may in fact contribute to the stimulation of elongation by fine tuning the timing of the CTD phosphorylation, which is important for the accessibility of pol II to the negative elongation factors (Palanga et al., 1998; Yamaguchi et al., 1999).

The general transcription factor TFIIIS stimulates pol II transcription at later stages of elongation (Wind and Reines, 2000). It was initially identified biochemically by its ability to promote synthesis of long transcripts by purified pol II (Reinberg et al., 1987; Reinberg and Roeder, 1987; Yoo et al., 1991). TFIIIS stimulates elongation by interacting with arrested pol II and activating cleavage of the nascent transcript by a latent endoribonuclease activity intrinsic to the polymerase (Izban and Luse, 1992; Rudd et al., 1994). The exact mechanisms

of pausing and arrest are not clearly established. However, it has been suggested that they result from an aberrant backward movement of pol II on the DNA template leading to displacement of the 3'-end of the transcript from the catalytic site. The TFIIS-mediated endonucleolytic cleavage of the nascent transcript may allow proper realignment of its 3'OH at the catalytic site of pol II and thus stimulate elongation (Izban and Luse, 1993; Reines, 1994; Gu and Reines, 1995; Nudler et al., 1995; Powell et al., 1996).

Interestingly, THIIIS-dependent RNA cleavage-extension has been also observed in binary complexes of yeast pol II with CPG79 RNA that encompasses a prokaryotic pause site (Johnson and Chamberlin, 1994). This finding demonstrates that pol II can bind the RNA transcript in a way that enables it to exhibit its intrinsic cleavage and polymerization properties in a functionally relevant manner, providing a simplified system for studying the details of the mechanism of pol II transcriptional elongation.

1.2 RNA PATHOGENS THOUGHT TO BE REPLICATED BY THE HOST RNA-POLYMERASE II

1.2.1 Plant Viroids

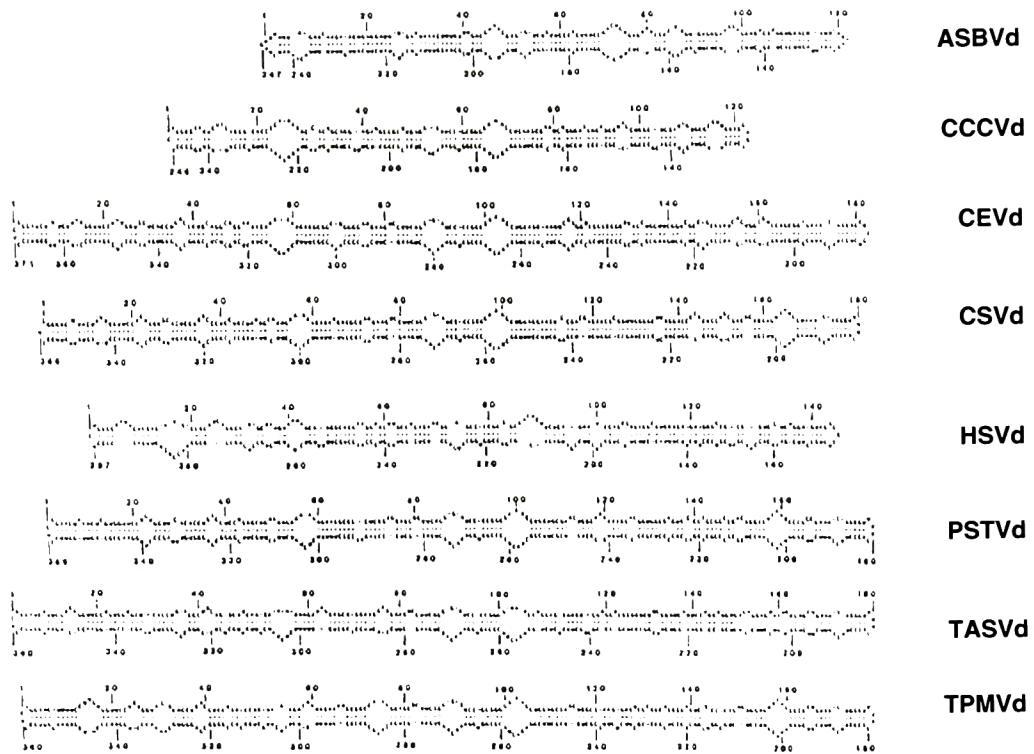
Viroids represent a group of infectious plant pathogens, whose genome consists of short (250 to 450 nt) single stranded, circular RNA, that due to the high degree

of intramolecular self-complementarity can be folded into an unbranched-rod like structure (Fig. 1A) (Gross et al., 1978; Diener, 1979; Gross et al., 1982), but can also contain regions with more complex interactions (Pelchat et al., 2000; Bussiere, 2000). As could be expected from their small size, there has been no evidence of any viroid-encoded translation products (reviewed in Riesner and Gross, 1985; Diener, 1987; Diener, 1993). Current viroid replication models are based on the presence of longer than unit length viroid RNA molecules in the infected plant tissues and propose different types of rolling circle mechanisms (Branch and Robertson, 1984; Daros et al., 1994). During the replication process the long multimeric, minus (-) sense RNAs synthesized on the circular genomic plus (+) RNA template would serve as templates for synthesis of new oligomeric (+) strand RNAs. The template for this process could be composed of either multimeric RNA, or circular monomers generated by processing from the multimeric RNA. In the final step, the multimeric (+) strand RNAs would be cleaved to unit-length molecules, which are then ligated into mature viroid circular RNAs. The self-cleaving ribozyme activity required for this mode of replication has been demonstrated in vitro only for a small subset of these pathogens (Hutchins, et al. 1986; Hernandez and Flores, 1992; Beaudry et al., 1995; Symons, 1997; Liu and Symons, 1998). For all other viroids, self-cleavage could not be shown despite repeated attempts. Processing of these viroids most likely involves a protein-dependent mechanism in which certain specific structural

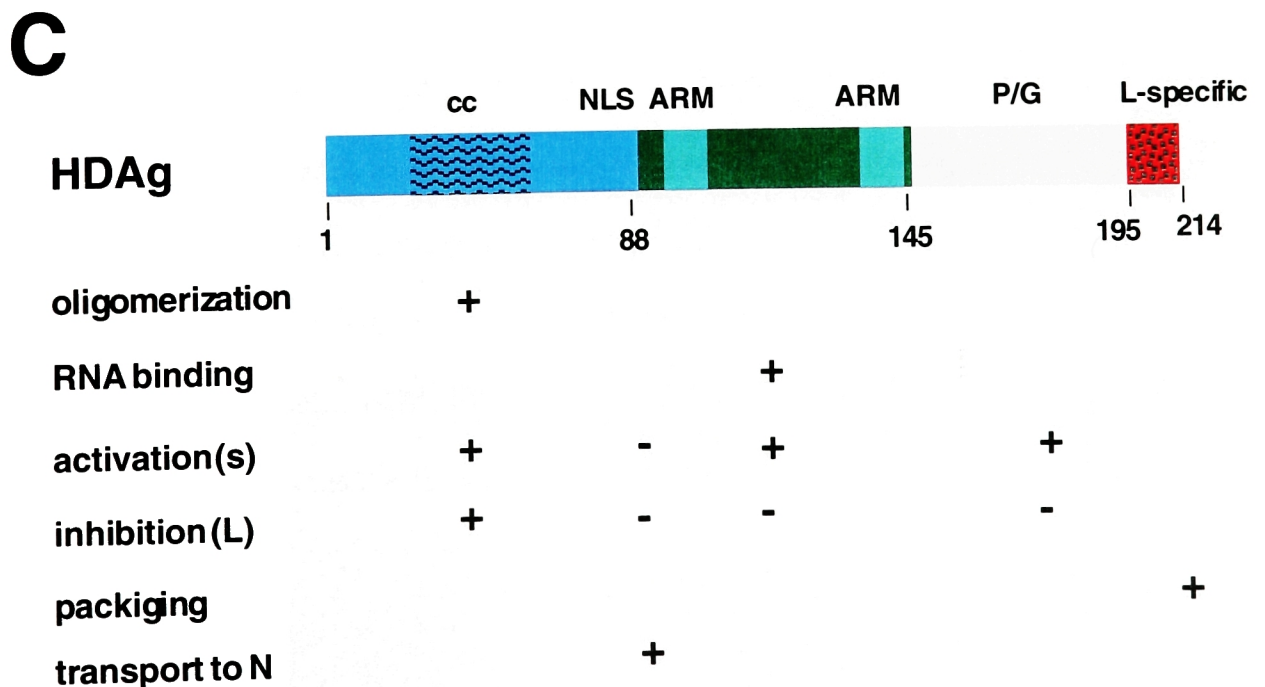
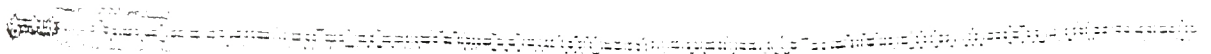
features of the viroid RNAs are required for functional interactions with the host cleavage and ligation enzymes.

Figure 1 RNA pathogens thought to be replicated by pol II. **A.** The rod-like structure of the genome of several plant viroid variants: ASBVd-avocado sunblotch; CCCVd-coconut cadang-cadang; CEVd-citrus exocortis; CSVd-chrysanthemum stunt; HSVd-hop stunt; PSTVd-potato spindle tuber; TASVd-tomato apical stunt; TPMVd-tomato planta macho. **B.** The rod-like structure of the human Hepatitis Delta Virus (HDV). **C.** A schematic representation of the HDV encoded protein, HDAg. N-terminal part: amino acids 1 to 88; central part, amino acids 88 to 145; C-terminal part, amino acids 145-195. The L-specific 19 a.a. long segment from aminoacid 195 to 214 is only present in one of the two forms of the antigen L-HDAg that is synthesized later in replication. cc-coiled-coil motif; NLS-nuclear localization signal; ARM-Arg rich motif; P/G-proline/glycine rich domain. The importance of the particular domains for the functions of HDAg as determined by mutation/deletion analyses is indicated by +. The - sign below a particular domain indicates that deletions or mutations in it did not affect the specific HDAg function.

A Plant Viroids



B Human Hepatitis Delta Virus (HDV)



Alternatively, certain viroid ribozymes may need host proteins to stabilize the RNA structure required for self-cleavage.

Since viroids do not encode for any proteins, the RNA synthesis during viroid replication must also entirely depend on the host enzyme systems. Initial attempts to uncover the identity of the host cell machinery involved in this process analyzed the effects of inhibitors of specific polymerases on the synthesis of viroid RNA in intact protoplasts or cell free nuclear homogenates. Inhibition of viroid RNA synthesis by α -amanitin concentrations (1 μ g/ml) known to inhibit RNA polymerase II function (Mühlbach and Sängner, 1979; Flores and Semancik, 1982; Semancik and Harper, 1984; Spiesmacher et al., 1985; Rivera-Bustamante and Semancik, 1989), and the ability of purified plant pol II to transcribe viroid RNA in vitro (Rackwitz et al., 1981; Boege et al., 1982), suggested the involvement of this activity in viroid replication. However, *E. coli* RNA polymerase, *E. coli* DNA polymerase I, Q β replicase, and RNA dependent RNA polymerase activity from tomato, were also able to copy viroid RNA in vitro (Owens and Diener, 1977; Rackwitz et al., 1981; Boege et al., 1982; Rohde et al., 1982). Efficient RNA synthesis in those in vitro systems depended on the presence of Mn²⁺ ions, conditions known to reduce the template specificity of the polymerases (Chamberlin, 1974; Semancik and Harper, 1984). Thus, the identity of the polymerase responsible for viroid replication in vivo remained unknown.

Most recently, an association between the host RNA polymerase II and citrus exocortis viroid (CEV) has been demonstrated in the transcriptionally active nucleoprotein complexes from infected tomato leaves (Warrilow and Symons, 1999). Plus- and minus-polarity CEV RNA were found to co-purify with the polymerase on columns containing a monoclonal antibody specific for the carboxy terminal (CTD) domain of the largest subunit of pol II (Warrilow and Symons, 1999). This result further supports the role of the host RNA polymerase II in replication of viroid RNAs.

It is important to note that all of the above mentioned studies concern group B viroids that replicate in the nucleus, but not group A viroids that replicate in the chloroplasts of the plant cell (Lafontaine et al., 1999). Other RNA polymerases have also been considered as possible candidates involved in the viroid replication. An RNA-dependent RNA polymerase from tomato has been purified and subsequently cloned (Schiebel et al., 1998). However, no viroid-specific polymerase activity has been demonstrated for this enzyme and the case for the viroid replicase still remains open.

1.2.2 Human Hepatitis Delta Virus

RNA polymerase II has been also implicated in replication of a human RNA pathogen, the Hepatitis Delta Virus (HDV). The initial recognition of HDV as a distinct infectious agent was achieved by the isolation of a novel antigen in hepatocytes of chronically infected hepatitis B patients (Rizzetto et al., 1977). In

the subsequent years it has become clear that this protein, termed hepatitis delta antigen (HDAg), is a product encoded by the novel infectious agent, Hepatitis Delta Virus. HDV particles (~36 nm) from infectious sera consist of an exterior HBV surface antigen and a lipid coat which encapsidates a ribonucleoprotein complex of the HDV RNA genome and two forms of the HDV-specific protein, the small and the large delta antigens (S-HDAg and L-HDAg) (Rizzetto et al., 1980; Bergmann and Gerin, 1986; Bonino et al., 1986). The HDV RNA genome represents a ~1.7 kb long, single stranded circular RNA, that similarly to the viroids, folds into rod-like structure, in which ~70% of the sequence is involved in the intramolecular base pairing (Fig. 1B) (Chen et al., 1986; Kos et al., 1986; Wang et al., 1986; Makino et al., 1987).

The structural similarities between viroid and HDV RNAs, together with the detection of longer than unit length HDV RNA molecules in infected cells, suggested that both classes of these pathogens may replicate via a similar rolling circle mechanism (Branch and Robertson, 1984). During the HDV rolling circle replication, multiple rounds of RNA synthesis on the circular genomic (G) RNA template would generate multimeric antigenomic (AG) RNA molecules, that can be subsequently resolved to the monomer unit length by the action of a self-cleaving ribozyme encoded by the HDV RNAs. The resulting linear monomers are thought to be ligated into circular AG RNAs, that serve as templates for multiple rounds of G HDV RNA synthesis. Consistent with this model, it was later demonstrated that both polarities of HDV RNA contain ribozyme domains

(Sharmeen et al., 1988; Wu and Lai, 1989; Wu et al., 1989; Branch and Robertson, 1991). Interestingly, the ribozyme domains are located in the left-hand tip of the HDV RNA rod structure (Fig. 1B), the region that shows a limited sequence homology to the viroids. Because of the sequence/structure and functional similarities to the viroids, this HDV RNA region is often referred to as the viroid-like domain (Wang et al., 1986; Taylor, 1992; Lai, 1995).

An RNA self-ligation activity that would allow the linear monomeric HDV RNAs to be ligated into circular RNA templates for replication, has not been experimentally demonstrated for HDV RNA, suggesting that circularization of HDV RNA *in vivo* may involve host cell proteins (Rosenstein and Been, 1990). Supporting this hypothesis, ligation *in vivo* of HDV RNAs containing the appropriate ends (5'-OH and 2',3'-cyclic phosphate) but lacking the ribozyme domains could be detected (Reid and Lazinski, 2000).

Although consistent with the proposed rolling-circle model for replication, the presence of multimeric HDV RNAs rises some questions concerning their functional significance. Since multimeric RNA could also fold into the rod-like structure that interferes with ribozyme activity (Wu et al., 1989; Matysiak et al., 1999; Isambert and Siggia, 2000), it is unclear whether *in vivo* they indeed act as the intermediates or they represent a dead-end products of the replication process.

In addition to the 1.7 kb and multimeric genomic and antigenomic RNAs, a third, ~800 nt long RNA species of AG polarity can be detected in HDV

replicating cells. This RNA encodes for the delta antigen (HDAg) and, as expected for a typical pol II transcript, it is polyadenylated (Hsieh et al., 1990). The 5' end of the HDAg mRNA has been mapped by primer extension analysis, and shown to correspond mostly to position 1631 in the right-hand terminal hairpin of AG HDV RNA (Hsieh et al., 1990; Gudima et al., 1999; Gudima et al., 2000). However, convincing evidence for the existence of a 5' cap structure on this mRNA has not been reported.

The two forms of the delta antigen, S-HDAg and L-HDAg are translated from the two, almost identical HDAg mRNA molecules synthesized during HDV replication (Bergmann and Gerin, 1986; Wang et al., 1986). The S-HDAg is a 24 kD nuclear phosphoprotein with RNA binding activity (Chang et al., 1988; Lin et al., 1990; Chao et al., 1991; Chou et al., 1998). It is translated from the HDV mRNA coding for 195 amino acids that is generated at the early steps of the replication cycle. Later in the replication cycle, the L-HDAg is translated from similar mRNA that instead of the UAG stop codon contains the UGG Trp codon at the same position, allowing for the translation of additional 19 amino acids and synthesis of a 27 kD L-HDAg protein (Luo et al., 1990). The A to G editing of the stop codon occurs on the AG HDV RNA (Casey and Gerin, 1995; Polson et al., 1998) and is most likely mediated by dsRNA-dependent adenosine deaminase (Polson et al., 1996).

Although S-HDAg and L-HDAg are structurally very similar, their functions differ significantly. The S-HDAg is essential for and stimulates HDV replication

(Kuo et al., 1989), while L-HDAg inhibits replication and is required instead for the packaging of the HDV RNA into virion particles (Chang et al., 1991; Glenn and White, 1991).

In infected patients HDV is found as a satellite of HBV. HDV uses the surface antigens of the HBV helper virus to infect hepatocytes. However, the replication of HDV RNA does not depend on any proteins encoded by the helper virus since HDV can replicate in a variety of animal cells lines in the absence of any HBV functions (Kuo et al., 1989; Glenn et al., 1990). Although essential for replication, the HDAg does not provide the replicase activity either. It shares no homology with any of the known RNA polymerases and it is not required for HDV RNA synthesis observed in vitro, using nuclear homogenates from uninfected cells (MacNaughton et al., 1991; Fu and Taylor, 1993). Thus, HDV, like the viroids, must rely on host cell machinery for the replication of its RNA.

1.2.3 Studies of HDV RNA replication in vivo

Some aspects of the requirements for HDV RNA synthesis have been resolved using in vivo systems that employ transfection of eukaryotic expression vectors that carry HDV cDNA into a hepatic carcinoma or fibroblast cell lines (Kuo et al., 1989), and monitoring HDV RNA synthesis at different time points post transfection.

Protein factors requirements

The essential role of the delta antigen for HDV replication was first demonstrated using a specifically modified HDV cDNA transfection system. Introducing a point mutation in the HDV sequence that disrupts the ORF for HDAg fully abolished HDV RNA replication (Kuo et al., 1989). However, viral replication could be restored if the wild type form of the protein was provided in trans, from a separate expression vector containing HDAg cDNA (Kuo et al., 1989). Similarly, altering the sequence of the polyadenylation signal from AAUAAA to UUUAAA also abolished HDV RNA replication that could be reconstituted by providing of wild type HDAg in trans (Hsieh et al., 1990).

The mechanism by which delta antigen regulates HDV RNA replication is not clear. However, several functional properties that may contribute to the stimulatory effect of the S-HDAg have been demonstrated in vivo. For example, replication-competent, as well as replication-incompetent HDV RNAs generated from HDV deletion mutants in a cDNA transfection system were stabilized in the presence of S-HDAg provided in trans (Lazinski and Taylor, 1994). However, the L-HDAg, that inhibits HDV replication, had a similar effect (Lazinski and Taylor, 1994). Thus, it is likely that stabilization of HDV RNA by S-HDAg may play a role, but is not the sole mechanism, by which this protein stimulates replication in vivo.

The roles of the individual domains of the delta antigen in the regulation of HDV replication have also been investigated in the cDNA transfection systems.

The N-terminal portion of HDAg (Fig. 1C) contains leucine heptad repeats that have a potential to form coiled-coil protein interaction domains. In fact, the ability of the delta antigen proteins to engage in protein-protein interaction and form both, homo- and hetero-dimers has been demonstrated in vitro (Xia and Lai, 1992). Interestingly, deletions and specific mutations in this region of the protein interfere with the ability of S-HDAg to homodimerize in vitro and in vivo (Xia and Lai, 1992; Lazinski and Taylor, 1993), and abolish the stimulatory effect of the protein on HDV replication in vivo (Chang et al., 1993). This result demonstrates that homodimerization is required for the S-HDAg stimulatory effect on HDV RNA replication. Furthermore, mutations in the N-terminal region in the context of L-HDAg abolish its negative effect on HDV RNA replication, as well as its ability to homodimerize and to interact with the S-HDAg in vitro and in vivo (Lazinski and Taylor, 1993). These results are consistent with the hypothesis that the L-HDAg exerts its trans-dominant inhibitory effect on HDV RNA replication by interfering with homodimerization of S-HDAg.

Interestingly, the N-terminal coiled-coil domain of HDAg was also shown to be required for the interaction of HDAg with its cellular homolog, the Delta Interacting Protein A (DIPA), that like L-HDAg, inhibits HDV RNA replication in vivo (Brazas and Ganem, 1996). The host cell is likely to provide some additional regulatory protein functions for the HDV RNA replication. However, the identification of the HDAg cellular homolog is especially significant in the light of the possible evolutionary links of HDV to the viroids and the mechanism of

HDV RNA replication itself. These points are discussed in greater detail later, in chapter 3.

The N-terminus of the delta antigens contains also a nuclear localization signal (NLS) shown to be required for the transport of HDV RNA to the nucleus in permeabilized HeLa cells (Chou et al., 1998). Deletion in this region partially inhibits the transport of HDAg to the nucleus in cells transfected HDV cDNA (Lazinski and Taylor, 1993). In addition to the NLS, two Arg-rich RNA binding motifs in the central region of the delta antigen are involved in the transport of HDV RNA to the nucleus (Chou et al., 1998). Various sequence changes in these motifs completely abolish the stimulatory effect of the S-HDAg on HDV RNA replication and the S-HDAg RNA binding activity in vitro (Chang et al., 1993; Lazinski and Taylor, 1993; Lee et al., 1993). Interestingly, equivalent mutations in the context of L-HDAg did not interfere with its trans-inhibitory effect on HDV RNA replication in vivo. This result demonstrates that S-HDAg binding to the HDV RNA contributes to the mechanism of stimulation of HDV RNA replication, while the L-HDAg most likely exerts its inhibitory effect solely by binding to the S-HDAg and interfering with its homodimerization.

Similarly, deletions in the C-terminal region abolish stimulation of HDV RNA replication by S-HDAg, but not the inhibitory effect of L-HDAg (Lazinski and Taylor, 1993). This Pro-Gly rich domain of S-HDAg may be involved in protein-protein interactions with other cellular factors and in that way stimulate the replication. It is not clear if the equivalent region in the context of L-HDAg is

involved in any functionally significant protein-protein interactions with host cell proteins since deletion of this region, which also includes the RNA binding domain and a portion of the L-HDAg specific region, did not affect its trans-inhibitory function (Lazinski and Taylor, 1993). The C-terminal L-HDAg specific region is required for the HDV RNA packaging, as deletion of these 19 amino acids abolish L-HDAg interactions with the HBV surface proteins and formation of virion particles (Chang et al., 1991; Lazinski and Taylor, 1993),

The HDV cDNA-based transfection systems are very convenient since they bypass the complications related to the synthesis of large amounts of long multimeric or circular RNA templates of consistent quality. In fact, attempts were made early on to study HDV replication requirements using in vitro transcribed HDV RNAs for transfection in cell lines. HDV RNA replication in such systems occurs only in cell lines that steadily expressed the delta antigen (Glenn et al., 1990). Recently however, a different RNA transfection system was developed that involves cotransfection of linear 1.9 kb HDV RNA molecules of either polarity with capped mRNA encoding for the delta antigen (Modahl and Lai, 2000). Despite the fact that the G HDV RNA template did not contain the two ribozyme domains and therefore, most likely, could not have been properly processed to circular RNA templates, the synthesis of the complementary, 1.7 kb long AG HDV RNA and low amounts of HDAg mRNA could be observed (Modahl and Lai, 2000; Modahl et al., 2000). In this system, the accumulation of the presumably newly synthesized HDAg mRNA was sensitive to α -amanitin (Modahl et al.,

2000). In contrast, the high levels of 1.7 kb AG HDV RNA synthesized from the transfected G RNA remained unaffected in a wide range of α -amanitin concentrations, unless the inhibitor was added early post-transfection. This effect was interpreted to be secondary to the HDAg mRNA inhibition, even though apparently sufficient amounts of capped mRNA were transfected to initiate HDV RNA replication (Modahl et al., 2000). The authors suggested that the observed differences in the effect of α -amanitin demonstrates that distinct cellular RNA polymerases are involved in the synthesis of 1.7 kb HDV and 0.8 kb HDAg mRNA. The effect of α -amanitin on G HDV RNA synthesis from an AG RNA template was not examined in this system. The RNA transfection system that employs AG RNA templates was used by the same authors to demonstrate the preferential sensitivity of G HDV RNA synthesis to the inhibitory effect of the L-HDAg (Modahl and Lai, 2000). Therefore, the mechanism for the observed differential sensitivity of G and AG HDV RNA synthesis to different factors as well as the identity of the polymerase(s) involved remains to be further investigated and established.

RNA template requirements

The RNA template requirements for efficient HDV RNA replication have also been investigated in vivo. Consistent with the proposed rolling circle model, mutations that interfere with the self-cleavage activity in vitro, abolish HDV RNA

replication in vivo in cDNA-transfected cells (Perrotta and Been, 1990; Wu and Huang, 1992; MacNaughton et al., 1993).

Subsequently, a deletion analysis was used to assess the HDV RNA requirements for efficient replication and packaging (Lazinski and Taylor, 1994). Although as little as 348 nt long G HDV RNA encompassing the left-hand terminal region of the HDV rod-like structure was capable of efficient self-cleavage, ligation, and copackaging with the HDV and HBV-specific proteins, no replication or synthesis of the opposite polarity RNA was observed with the shorter HDV constructs (Lazinski and Taylor, 1994). In fact, all of the tested deletions in the HDV RNA sequence, abolished the replication, despite the fact that the self-cleavage activities were not affected. Although this study did not define the RNA promoter element, it excluded the possibility that any region of HDV can be recognized as a template for the polymerase to initiate replication. The experiments could not exclude the possibility that two distinct promoters direct initiation of synthesis of the two complementary HDV RNAs generated during HDV replication. In fact, a more recent study that employed a linker-scanning mutagenesis instead of the deletion approach, demonstrated that specifically positioned mutations differentially affected the synthesis of full length AG and G HDV RNAs and HDAg mRNA (Wang et al., 1997b). Similarly, an independent mutagenesis study pointed out that different regions of the HDV RNA genome affect replication to a different extent (Wu et al., 1997). However, as in the previous studies, no structural or functional basis for this effect was

established. Thus, the RNA template requirements for efficient synthesis of the HDV genomic, antigenomic and HDAg mRNA remain to be further characterized at a molecular level and considered in a functional context of the RNA replication of this virus.

1.2.4 Studies HDV RNA synthesis in vitro

RNA polymerase II was first suggested to be the host cell activity responsible for HDV RNA replication because of the apparent structural and functional similarities of HDV to the plant pathogens, viroids. The initial studies in vitro actually addressed the role of the delta antigen in HDV RNA synthesis. In vitro transcribed trimeric length HDV RNA templates were incubated in the presence of ^{32}P labeled NTP with nuclear homogenates from either regular HepG2 hepatoma cell line or HepG2 that steadily expresses the HDAg protein (MacNaughton et al., 1991). The ^{32}P labeled HDV RNA from both types of homogenates could be detected to hybridize to the HDV cDNA probe. Thus it was concluded that HDV RNA replication occurred in HDAg independent manner. In addition, it was suggested that the cellular pol II activity is involved in the observed HDV RNA synthesis even though the reaction was not α -amanitin sensitive. In similar run-on experiments carried out with nuclear homogenates from a cell line containing the integrated trimeric length HDV cDNA, HDV RNA synthesis was inhibited by α -amnitin and pol II involvement was proposed (MacNaughton et al., 1991).

In a later study, using again a variety of nuclear homogenates and trimeric length RNA templates, α -amanitin sensitive HDV RNA synthesis was detected using a nested RT-PCR assay (Fu and Taylor, 1993). Again, the reaction was not dependent on the presence of HDAg protein in the nuclear homogenates, confirming that the viral protein is not required. Low levels of 1.7 kb long HDV RNA complementary to the template, could also be detected by Northern blot analysis (Fu and Taylor, 1993). However, the template requirements for the reaction were not examined in that study. In addition, considering the length and extensive degradation of the input RNA template, a more precise product analysis that could assess the template region initially recognized by the polymerase was not possible, and a nonspecific end-to-end RNA copying could not be excluded.

More recently, a ~200 nt G polarity RNA corresponding to the right-hand tip of the HDV rod-like structure was used by Beard and colleagues as a template for transcription in vitro in HeLa nuclear extracts (NE) (Beard et al., 1996). The detection of a specific HDV RNA product in this system implied the presence of a G HDV RNA promoter located in this short RNA segment. However, in the absence of a direct analysis of the RNA product, the detailed mechanism of this transcription reaction remains unclear. Furthermore, no specific sequence or structural requirements could be recognized for such an RNA promoter since almost all of the introduced mutations completely abolished

the reaction in vitro, while only a few had any effect on HDV replication in vivo and none interfered with HDAg synthesis (Beard et al., 1996).

In summary, although many attempts have been made to determine the polymerase activity responsible for HDV RNA replication, there is mostly indirect evidence suggesting pol II involvement, but a role for other, unidentified host cell polymerases can not be excluded (MacNaughton et al., 1991; Fu and Taylor, 1993; Modahl et al., 2000). In addition, the RNA template requirements for the process still remain enigmatic. Therefore, further studies are needed to directly test if RNA polymerase II can use HDV RNA as a template for transcription, and determine the specific characteristics of HDV RNA that can redirect the DNA-dependent polymerase into the RNA-templated transcription. Studies of the pol II function on this 'atypical' RNA template may also provide an additional point of view on certain aspects of the pol II-mediated transcription and thus, contribute to the better understanding of the process of gene expression.

Chapter 2

RESULTS

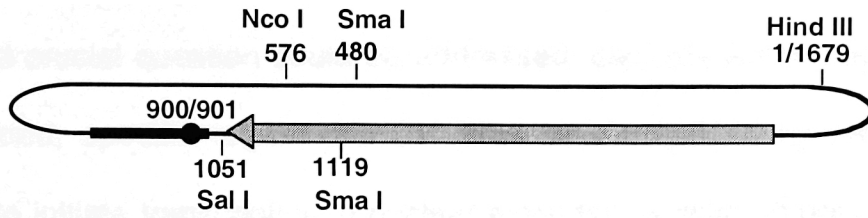
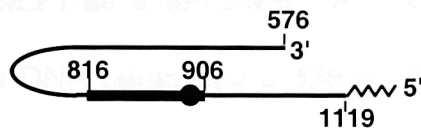
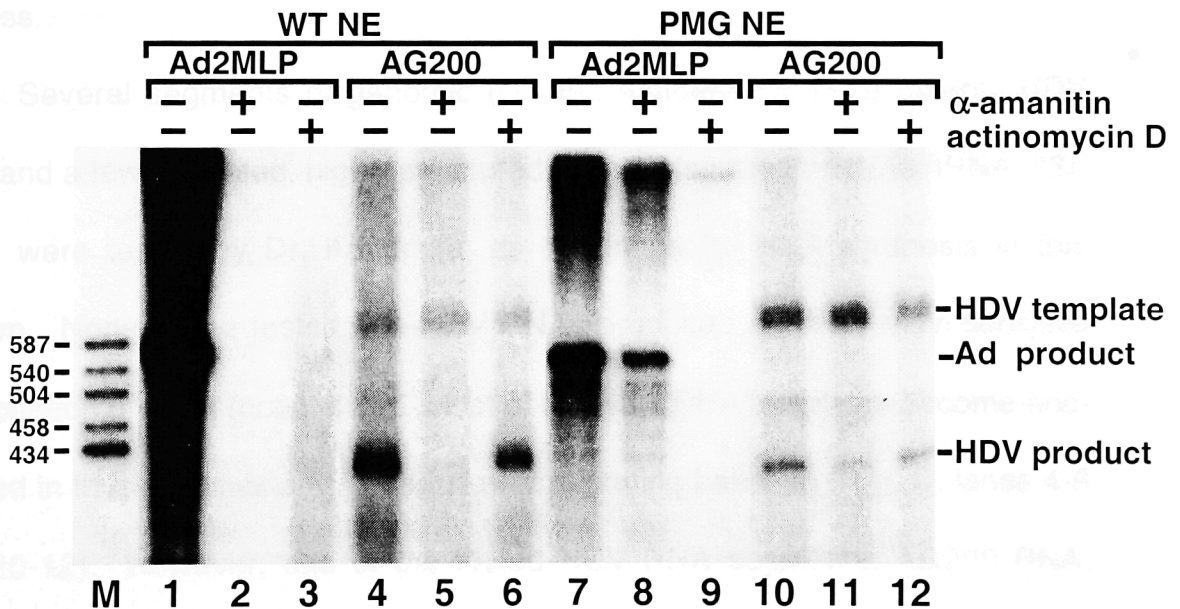
2.1 HDV RNA –TEMPLATED TRANSCRIPTION BY POL II IN VITRO

2.1.1 Description of a system to study HDV RNA-templated transcription in vitro

Previous in vitro studies of RNA-templated transcription using viroid or HDV RNA templates (described in the Introduction) were crucial for introducing the idea that pol II is involved in replication of these pathogens. However, the α -amanitin inhibition studies only implicated pol II involvement in the process, without directly demonstrating pol II's ability to specifically use these RNAs as templates for transcription. Furthermore, those in vitro systems suffered from inability to specifically and directly address very important questions regarding the RNA species involved in the process. In particular, the length of the RNA templates used (full unit length or multimers), together with the low efficiency of the observed transcription, did not allow for the comprehensive, direct analysis of the transcribed RNA. In addition, the indirect assays (RT-PCR, Primer Extension) used to analyze the transcription products could give misleading results, as both viroid and HDV represent highly structured RNA molecules. Similarly, Northern hybridization could not provide a precise information about the fidelity of pol II transcription using the RNA templates. Specific template requirements for the RNA-templated pol II transcription are also not easily accessible in such systems.

Because of the complexity of this problem, we pursued different approach in our laboratory. A simplified, in vitro system was considered, in

Figure 2. Antigenomic (AG) HDV RNA represents a template for pol II-mediated RNA synthesis. **A.** Schematic representation of the full-length single-stranded circular AG HDV RNA folded into an unbranched rod-like structure. The AG ribozyme domain (rectangle) and its site of cleavage (circle) are indicated. The arrow depicts the open reading frame for HDAg. Some of the characteristic restriction sites in the HDV cDNA and their relative positions in the AG HDV RNA are also depicted. **B.** AG200, the segment of AG HDV RNA used as a template in NE transcription reactions in C. **C.** RNA synthesis in wt (lanes 1-6) and PMG (lanes 7-12) HeLa NE using Ad2MLP DNA (lanes 1-3 and 7-9) or the AG200 RNA (lanes 4-6 and 10-12) as templates. Transcription reactions monitored by incorporation of α -³²P-GTP into newly synthesized RNAs were carried out under standard conditions (lanes 1, 4, 7 and 10), in the presence of 1 μ g/ml α -amanitin (lanes 2, 5, 8 and 11), or 20 μ g/ml actinomycin D (lanes 3, 6, 9 and 12). Products were resolved in a 5% polyacrylamide/8M urea gel and detected by autoradiography. Lane M contains pBR322xHaeIII DNA size marker. Positions of Ad2MLP and HDV RNA transcription products and HDV RNA templates end-labeled in the presence of NE are indicated. The experiment was performed by Dr. M. M. Konarska.

A**AG HDV RNA****B****AG200****C**

which shorter segments of HDV RNA were tested for their ability to serve as templates for pol II transcription in nuclear extracts of HeLa cells (NE). In this way the first crucial question could be addressed; can any single segment, or only a defined, specific region of HDV RNA be recognized by the cellular machinery to initiate transcription in nuclear extracts? Activity of pol II in these extracts was confirmed in standard reactions using the Adenovirus 2 major late promoter DNA (Ad2MLP) as a template. As expected, accurate initiation of transcription from this DNA generated a 536 nt RNA product (Fig. 2C, lane 1). RNA synthesis in these reactions is completely inhibited by α -amanitin (1 μ g/ml) (Fig. 2C, lane 2), consistent with the involvement of pol II, and by actinomycin D (20 μ g/ml) (Fig. 2C, lane 3), consistent with the use of a DNA template in this process.

Several segments of genomic (G) and antigenomic (AG) polarity HDV RNA and a few unrelated, highly structured RNA molecules (snRNAs, tRNA, 7SL RNA) were tested by Dr. Konarska, as templates for RNA synthesis in this system. None of the tested non-HDV RNAs produced an α -amanitin-sensitive transcription product (not shown). Most of the HDV RNA templates become end-labeled in these reactions (e.g. see slower migrating bands in Fig. 2C, lanes 4-6 and 10-12). However, one of the tested HDV RNA segments, AG200 RNA, yielded in addition a unique, faster migrating product (Fig. 2C, lane 4). The AG200 RNA template (Fig. 2B), contains 533 nt of AG HDV RNA from position 576 to 1109 (Wang et al., 1986) flanked at its 5' end by a 59 nt polylinker

sequence. Synthesis of the faster migrating product (~400 nt) exhibits only a modest inhibition of transcription in the presence of actinomycin D (Fig. 2C, lane 6), consistent with an RNA-, rather than DNA-templated process. Formation of this RNA product is strongly inhibited by α -amanitin (Fig. 2C, lane 5), suggesting that it depends on pol II. In contrast, template labeling is not significantly affected by either of these toxins (Fig. 2C, lanes 5,6).

As mentioned earlier, α -amanitin sensitivity of transcription in vitro only implicates pol II in the process and the involvement of another, as yet uncharacterized polymerase activity with a similar α -amanitin sensitivity can not be excluded. To confirm pol II involvement in the HDV RNA-templated transcription observed in vitro, NE was prepared from PMG HeLa cell line. PMG cell line contains a copy of the mouse RPII215 gene, encoding an α -amanitin resistant variant of the largest subunit of pol II, in addition to the endogenous α -amanitin-sensitive pol II alleles [Bartolomei and Corden, 1987; J. Corden, personal communication). As expected, in PMG NE, transcription from the Ad2MLP DNA template is partially resistant to α -amanitin (Fig. 2C, compare lanes 7, 8 with 1,2). Similar resistance is observed when AG200 HDV RNA template is used (Fig. 2C, lanes 10, 11). These results firmly demonstrate that pol II in the HeLa NE can use HDV RNA as a template. However, pol II purified from NE (gift of A. Hoffmann and R. G. Roeder) did not support RNA transcription under these conditions, suggesting that additional NE components must also be required. One such crucial protein factor was identified by M.

Gottardo as PC4 (Gottardo, 1998), a protein earlier described by M. Kretzschmar and colleagues as a co-activator of pol II transcription (Kretzschmar et al., 1994).

Deletion analysis of the AG HDV RNA template

To identify the smallest HDV RNA template used by pol II for transcription in vitro, a deletion analysis was performed. Progressive deletions of HDV sequences from either end of the AG200 template do not significantly affect the efficiency of transcription. In fact, a 280 nt AG128 RNA (Fig. 3B) is used even more efficiently than the 592 nt AG200 RNA template (Fig. 3A, lane 1 and data not shown). A number of RNAs containing internal deletions: AG130, AG131, AG129, and AG103 (see Fig.2B) serve as efficient templates in vitro (Fig. 3A, lanes 2, 3, 6 and 7, respectively) while efficiency of RNA synthesis using other templates, e.g. AG132 and AG138 RNA, is greatly diminished (Fig. 3A lanes 4, 5). Interestingly, in cases where ~30 nt deletions on either side of the terminal hairpin disrupt the overall unbranched rod-like structure, the efficiency of transcription is strongly inhibited (AG132 and AG138, Fig. 3A, lanes 4, 5). A combination of these two deletions in AG103 restores the unbranched hairpin structure and results in efficient transcription (Fig. 3A, lane 7). These results suggest that the characteristic structure of the RNA template in proximity of the hairpin tip is important for efficient transcription. Similar ~30 nt deletions introduced more distally (AG130 and AG131) do not have a significant effect (Fig. 3A, lanes 2,3).

Figure 3. Deletion analysis of HDV RNA segment required as a template for NE transcription. **A.** Transcription reactions using templates represented in panel B. In addition to the labeled NE RNA products (~125-155 nt), some end labeling of the full-length RNA templates was frequently observed. Products were resolved in a 5% polyacrylamide/8M urea gel. Positions of pBR322xHaeIII DNA marker are indicated (lane M). **B.** Schematic representation of the RNA templates. AG128 RNA contains HDV sequence (pos. 912-642) spanning the left-hand terminal portion of the rod-like structure and includes the AG ribozyme domain. Subsequent internal deletions of this template generated AG130 (Δ pos. 863-890), AG131 (Δ pos. 693-718), AG132 (Δ pos. 718-768), and AG138 (Δ pos. 822-858). To generate AG129 and AG103, deletions of AG130+AG131 and AG132+AG138, respectively, were combined. Positions of the cleavage site of G (circle) or AG ribozyme (rectangle) are indicated. Active ribozymes are represented by closed symbols, while those inactivated by the introduced deletions are shown as opened rectangles. The experiment was performed by Dr. M. M. Konarska.

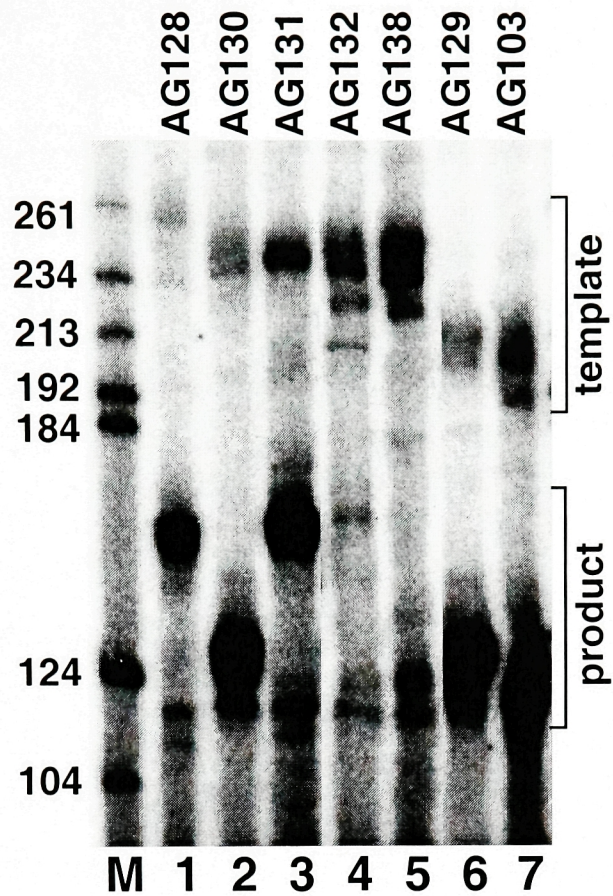
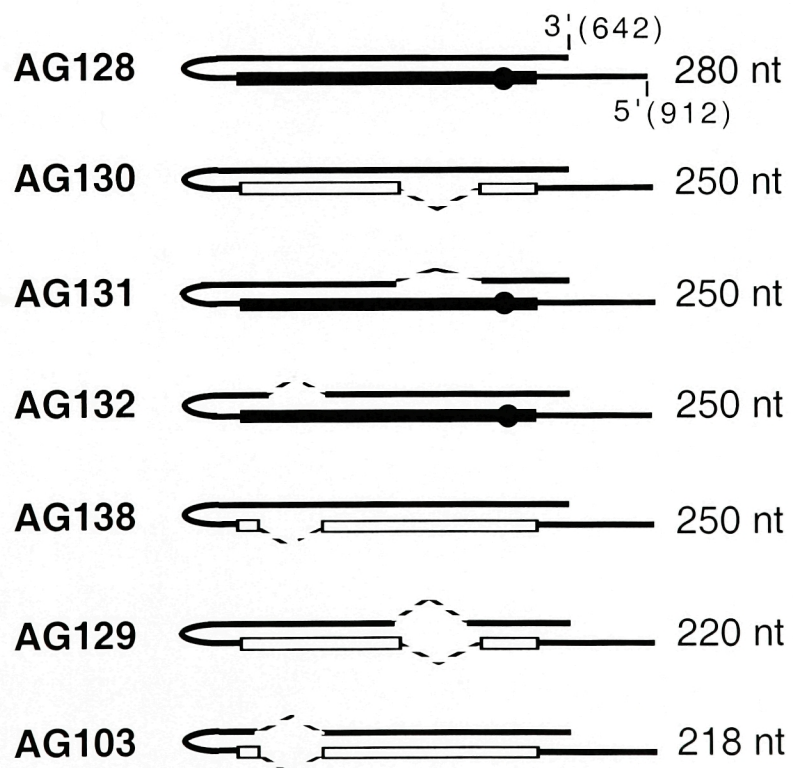
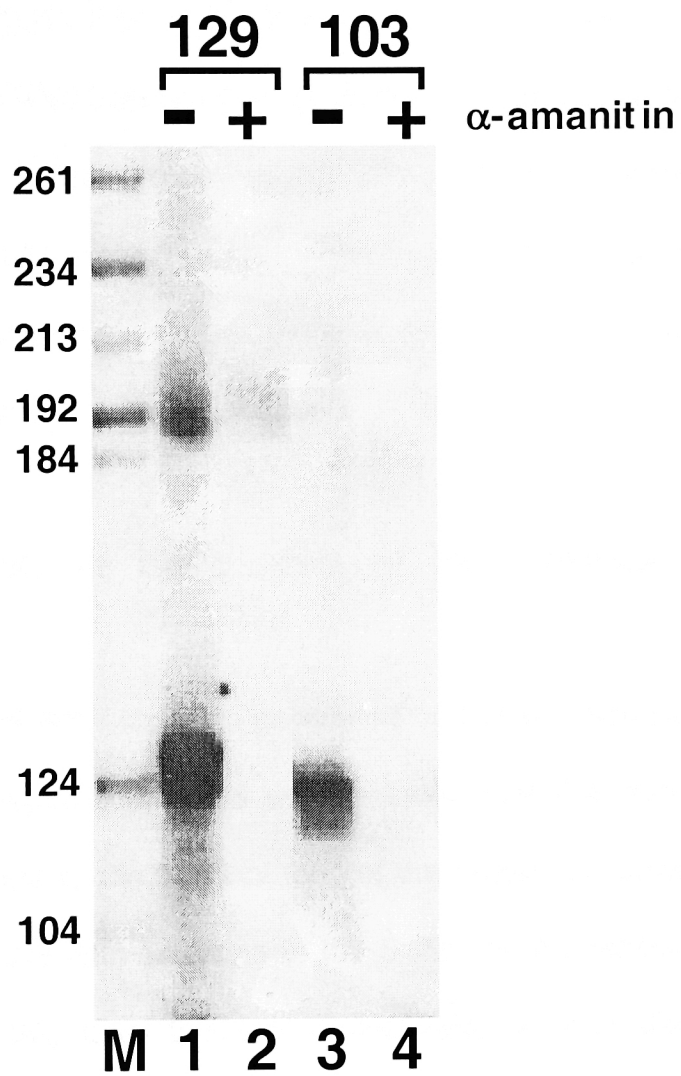
A**B**

Figure 4. Transcription using the smallest HDV RNA templates encompassing the left-hand terminal hairpin is also sensitive to α -amanitin. Transcription reactions using the templates described in Figure 3, AG129 (lanes 1 and 2) or AG103 (lanes 3 and 4) were carried out in standard HeLa NE (lanes 1 and 3) or the presence of 1 μ g/ml α -amanitin (lanes 2 and 4). Products were resolved in a 5% polyacrylamide/8M urea gel and detected by autoradiography. Lane M contains pBR322xHaeIII DNA size marker.



RNA synthesis using all templates, including the shortest AG129 and AG103 RNAs (Fig. 3B), is sensitive to α -amanitin (Fig. 4, lanes 2 and 4 respectively), consistent with the involvement of pol II in these reactions. Further deletions of the 218 nt AG103 RNA from either 5' or 3' end inactivated these templates and thus AG103 RNA was used as a standard, minimal template in all subsequent experiments. Identification of the short AG103 template that generates high levels of ~120 nt long product allows for a more precise analysis of the template sequence and/or structure requirements for efficient transcription and for the direct sequence analysis of the product.

2.1.2 Sequence analysis of the product of HDV RNA-templated transcription in vitro

From the apparent sizes of RNA products generated by transcription using the templates described in Figure 3, it could be concluded that the start site of transcription is located close to the terminal loop of the hairpin structure of the AG103 template. Templates with deletions in the 5' half of the molecule, e.g. AG130 and AG138 (Fig. 3A, lanes 2 and 5 respectively), yield products that migrate faster than the product of AG128 RNA. The change in product mobility correlates well with the size of the deletion (~30 nt), suggesting that transcription starts 3' of the deletion in AG138 template. In addition, a number of deletions in the 3' portion of the template did not affect gel mobility of the corresponding products (compare AG128 with AG131 and AG132, Fig. 3A, lanes 1, 3 and 4),

suggesting that the start site of transcription is located 5' of the region deleted in AG132 RNA. Together, these results indicate that pol II starts transcription near the terminal loop, and copies the 5' half of the template. Consistent with this notion, double deletions in both 5' and 3' regions of the template (e.g. AG129 and AG103, Fig. 3A, lanes 6, 7) affect the mobility of products to the same extent as single deletions in the 5' region only (e.g. AG130 and AG138, Fig. 3A, lanes 2 and 5).

RNase digestion analysis

To determine the precise start site of pol II transcription on the HDV RNA template in vitro and to evaluate the fidelity of this process, a detailed RNA sequence analysis of the NE product was performed. Gel-purified transcription products were subjected to exhaustive digestion with RNase T1 (cleaves at G↓) or RNase A (cleaves at Py↓) and the resulting oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel. Oligonucleotide markers were generated by an identical treatment of synthetic RNAs. For example, the marker shown in Figure 5D, lane 1, was generated by digestion of SP6-G103 RNA (Fig. 5B) that is complementary to the 5' half of the template (Fig. 5A). RNAs complementary to the full-length RNA templates containing specific mutations, e.g. T3-G103-77, T3-G129 (only the sequence spanning the mutated region is shown in Fig. 5C) were also included as markers (Fig. 5D, lanes 7 and 8, respectively).

Figure 5. The HDV RNA-templated pol II transcription is precise—RNase T1 analysis of the transcription products. **A.** Sequence of the minimal AG103 HDV RNA template. The numbering system corresponds to the positions in the RNA template. **B.** SP6-G103 represents a genomic (G) polarity RNA, a copy of the bottom strand of the AG103 template (except for 2 nt at its 5' end). RNase T1 cleavage sites are represented by vertical bars. The length (in nt) of the characteristic RNase T1 oligonucleotides used as size markers are indicated. **C.** Genomic polarity sequences of the mutant HDV RNAs analyzed in D. Only RNase T1 oligonucleotide fragments changed by the introduced mutations are indicated. The boxed regions in B and C encompass the segments of G RNA detected by the RNase T1 digestion analysis of the corresponding NE RNA products. **D.** RNase T1 digestion patterns of gel purified NE RNA products 129, 103, 103-12, 103-16, and 103-77 (lanes 2 to 6, respectively) are compared with patterns of SP6-transcribed G103 RNA (lane 1), and T3-transcribed G103-77 and G129 RNAs (lanes 7 and 8, respectively). The digestion fragments were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography.

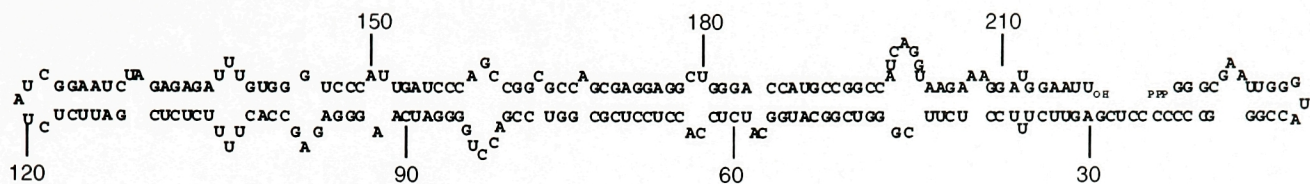
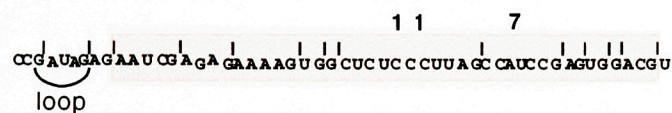
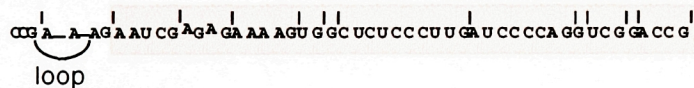
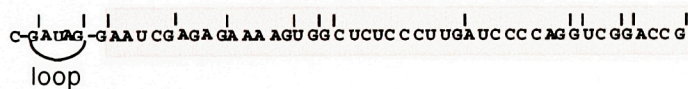
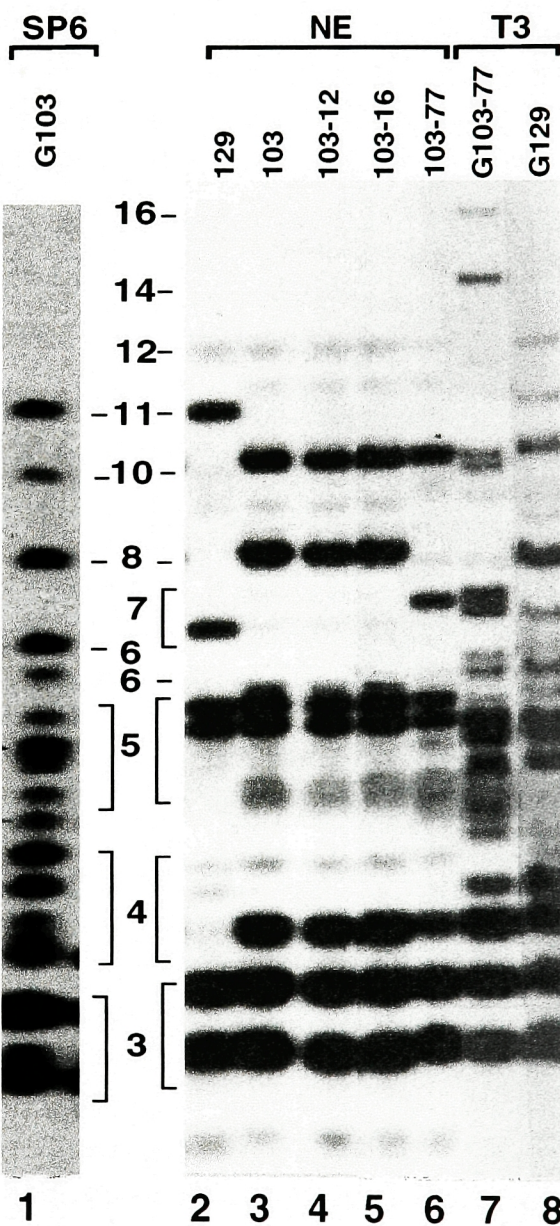
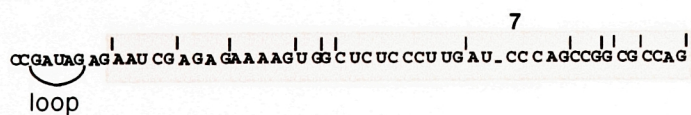
A**AG103****B****SP6-G103****C****D****129****103-12****103-16****103-77**

Figure 6. The HDV RNA-templated pol II transcription is precise— RNase A analysis of the transcription products. A. SP6-G103 as in Figure 5B except that RNase A cleavage sites are represented by vertical bars. **B.** Genomic polarity sequences of the mutant HDV RNAs analyzed in C. Only RNase A oligonucleotide fragments changed by the introduced mutations are indicated. The boxed regions in A and B encompass the segments of G RNA detected by the RNase A digestion analysis of the corresponding NE RNA products. **C.** RNase A digestion patterns of gel purified NE RNA products 103, 103-12, 103-16, 103-18 and 103-73 (lanes 2 to 6, respectively) are compared with patterns of SP6-transcribed G103 RNA (lane 1). The digestion fragments were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography.

The predicted RNase T1 oligonucleotides for SP6-G103 are shown in Figure 5B, and the actual RNase T1 digestion pattern in Figure 5D, lane 1. The pattern consists of a set of 11 nt, 10 nt, 8 nt, and 6 nt-long oligonucleotide fragments, four 5 nt-long fragments, one 4 nt, one 3 nt, and a number of smaller fragments (compare Fig. 5B and D, lane 1). The RNase T1 pattern of NE103 product contains one 10 nt, 8 nt, 4 nt, two 5 nt-long, and a number of shorter oligonucleotide fragments (Fig. 5D, lane 3). They all correspond to a region of ~40 nt in the control SP6-G103 RNA (Fig. 5B, shaded area), suggesting that only a portion of the AG template is copied in this reaction. The transcription process is precise, as demonstrated by the comparison of the RNase digestion patterns of several different mutant NE products (Fig. 5C and D). For example, RNase T1 digestion of NE129 RNA product (Fig. 5D, lane 2) generates a set of fragments distinct from those characteristic of NE103 RNA (Fig. 5D, lane 3). The differences include the absence of a 4 nt fragment in the NE129 pattern and the presence of 11 nt and 7 nt fragments instead of the 10 nt and 8 nt fragments characteristic of the NE103 product (compare lanes 2 and 3 in Fig. 5D). The observed differences are expected, considering the sequences of G129 and G103 RNAs (compare Fig. 5B and C) and demonstrate that pol II precisely copies the AG RNA template. Similarly, digestion of the NE103-77 RNA product yields the expected 7 nt fragment that replaces an 8 nt oligo found in NE103 RNA (Fig. 5D, lane 6 and 3, respectively). In contrast, RNase T1 patterns of the NE103-12 and 103-16 RNA products are not changed (Fig. 5D, lanes 4 and 5

respectively) as expected, since the corresponding mutations in their RNA templates do not affect any RNase T1 sites in the copied region (Fig. 5C). This result additionally demonstrates that pol II precisely copies the RNA template and does not introduce any non-templated nucleotides during transcription.

The results of the RNase A digestion analysis, an example of which is presented in Figure 6, are also consistent with this finding. The predicted RNase A oligonucleotides from SP6-G103 RNA are presented in Fig. 6A and the actual RNase A digestion pattern in Fig. 6C, lane 1. It consists of a 14 nt, 11 nt, 10 nt, 8 nt, 7 nt, 6 nt and a number of shorter oligonucleotides. The RNase A pattern of the NE103 RNA product (Fig. 6C, lane 2) contains the 11 nt long and a number of shorter fragments that correspond to the region copied by pol II, as demonstrated by the RNase T1 analysis. Mutations in the AG template e.g. in AG103-18 and AG103-73 that change the RNase A sites in the complementary RNAs (see Fig. 6B for sequences) result in distinct digestion patterns of the corresponding NE products. For example, the characteristic 11-mer is replaced by a 9 nt-long oligo in the digestion pattern of NE103-18 (Fig. 6C, lane 5). This is expected, since the AG103-18 template contains a 2 nt deletion in the region encompassing the wt 11 nt fragment (Fig. 6C). Similarly, eliminating the RNase A sites upstream of the 11 nt fragment in the AG103-73 mutant (Fig. 6B) change the digestion pattern of the corresponding NE RNA product (Fig. 6C, lane 6). The 11-mer is replaced by a 17 nt long fragment (Fig. 6C, lane 6). This fragment is 4 nt shorter than predicted if transcription copied the entire region from which

RNase A sites were eliminated (Fig. 6B), suggesting that transcription initiates within the 21 nt fragment.

Analysis of the 5' end of the HDV RNA pol II transcript

To determine the exact start site of transcription, RNase T1 digestion analysis was performed on NE products generated from a number of mutant templates.

In RNase T1 pattern of the NE103 product (Fig. 7A, lane 1) the 5 nt AAUCG represents the 5'-most ³²P-labeled oligonucleotide of the transcript (Fig. 7C). In digestions of NE103-63 RNA, the AAUCG 5-mer is replaced by a new 12 nt fragment (Fig.7A, lane 5), indicating that the RNA product contains at least 7 nt of an upstream sequence. Similarly, in digestions of NE103-75 RNA, a new 20 nt fragment appears instead of AAUCG and AAAAG 5-mers (Fig.7A, lane 9), consistent with the introduced mutations of RNase T1 sites located between these two fragments (Fig.7C). The new 12 nt and 20 nt long oligos are 1 nt shorter than the fragments generated from the corresponding region in synthetic G polarity RNAs containing the same mutations (Fig.7A, lanes 13 and 14 and Fig. 7C for sequences), indicating that they span the initiation site of the transcript. Consistent with these results, RNase A digestion pattern of NE103-73 contains a 17 nt long oligo (Fig. 7B, lane 2) that is 4 nt shorter than the 21-mer predicted if transcription copied the entire mutated region (Fig. 7D and 7B, lane 6 for digestion pattern of synthetic G103-75 RNA).

Figure 7. Absence of a triphosphate or a cap structure at the 5' end of the RNA-templated pol II transcript. **A.** RNase T1 digestion patterns of wild type, NE103 transcription product (lanes 1-4) or mutant NE products (103-63, 103-75 and 103-73, as indicated). Prior to RNase T1 digestion the NE products were subjected to phosphatase, CIP (lanes 2, 6 and 10), β -elimination (lanes 3, 7 and 11) or β -elimination followed by CIP treatment (lanes 4, 8 and 12). For comparison, RNase T1 digestion patterns of T3-transcribed G103-75 and G103-63 RNAs are included (lanes 13, 14). The 12-mer and 20-mer that represent the 5'-most digestion fragments in the transcribed segments of NE103-63 and -75 respectively, are indicated by asterisks. **B.** RNase A digestion pattern of wild type, NE103 transcription product (lane 1) or mutant NE103-73 (lanes 2-5). The NE products were digested by RNase A directly (lanes 1 and 2) or after pretreatment with phosphatase, CIP (lane 3), β -elimination (lanes 4) or β -elimination followed by CIP treatment (lanes 5). The 17-mer that represents the 5'-most oligonucleotide of the NE103-73 product is indicated by asterisk. RNase A digestion pattern of control, T3-transcribed G103-73 RNA is also included (lane 6). The digestion fragments were resolved in a 25% polyacrylamide/8M urea gel. **C. and D.** Genomic polarity sequences of the wt and mutant HDV RNAs analyzed in A and B, respectively. RNase T1 (in C) and RNase A (in D) sites are designated by vertical bars. The length (in nt) of the digestion fragments is indicated above the corresponding sequence. The boxed regions encompass fragments detected by the RNase analysis.

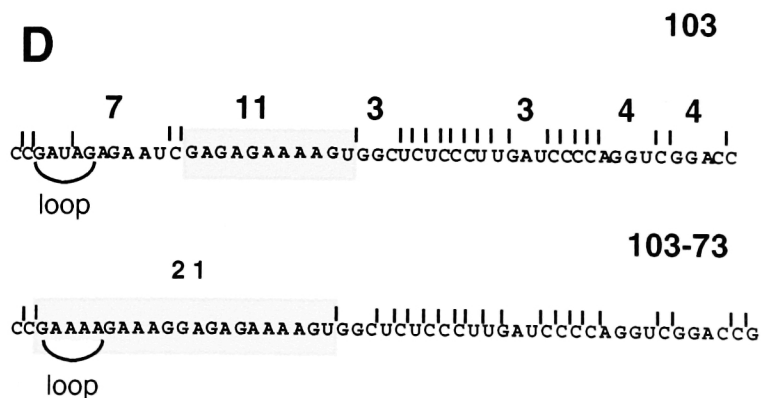
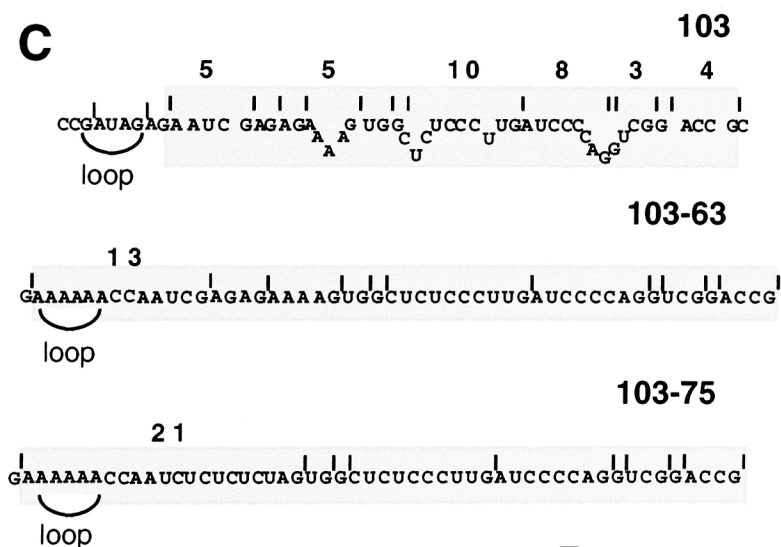
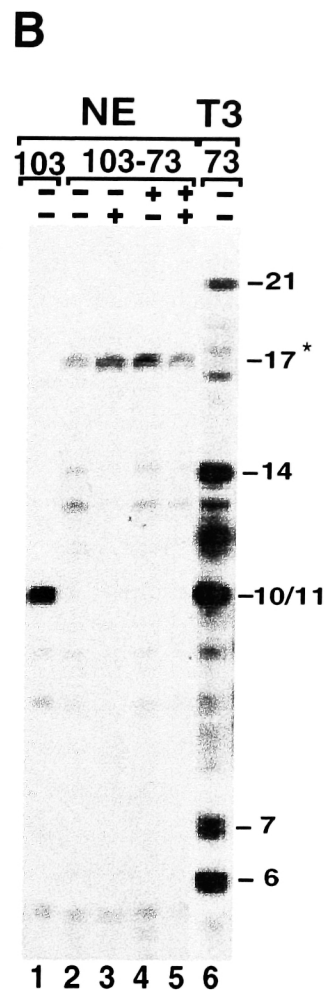
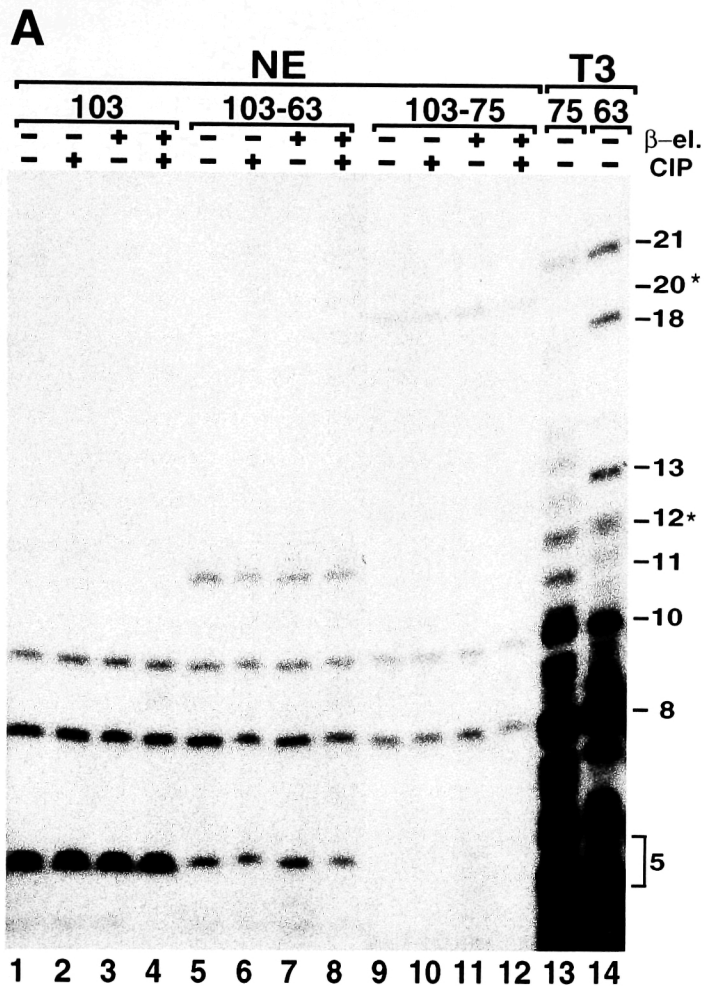


Figure 8. Detection of a 5' cap structure on a synthetic RNA.

Uncapped (P) or GpppG capped (C) oligonucleotides were subjected to phosphatase, CIP (lanes 2, 4 and 6), β -elimination (lane 7) or β -elimination followed by CIP treatment (lane 8) and then digested (lanes 3-8) or not (lane 2) with RNase A. The RNase A sites in the oligonucleotides P and C are indicated by vertical bars. The digestion fragments were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography.

pppGAAGGUAAGUAU
P

GpppGAAGGUAAGUAU
C

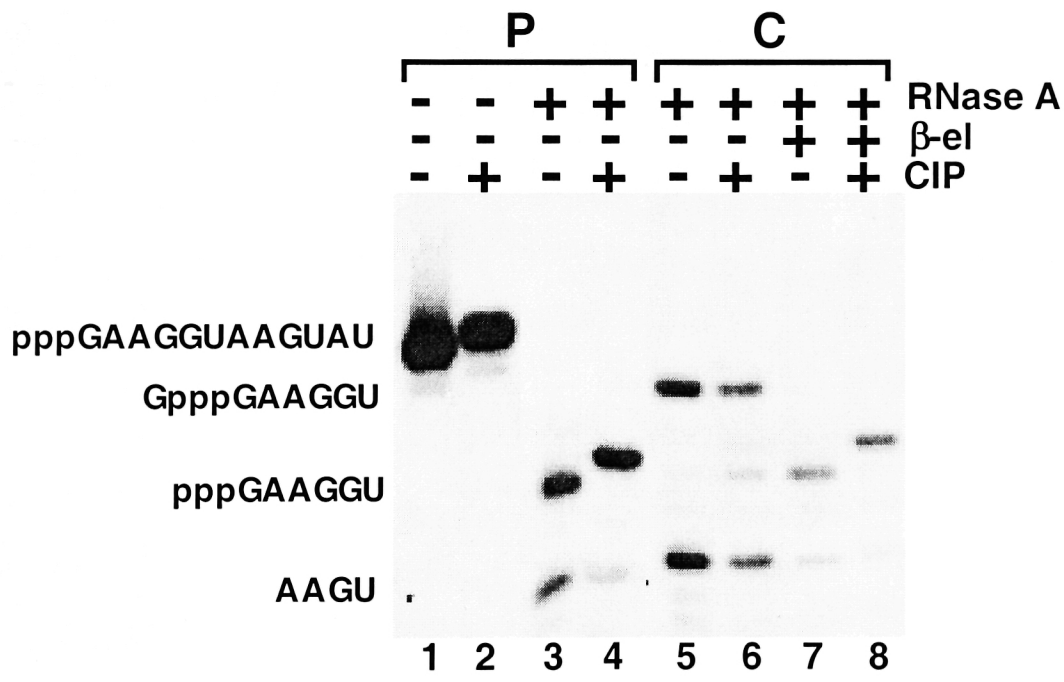
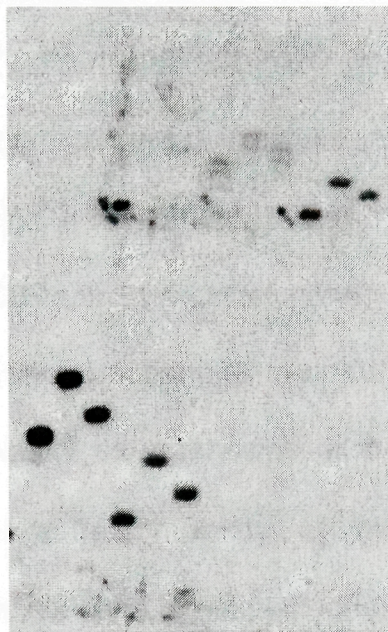


Figure 9. Detection of a hydroxyl group at the 5' end of the NE RNA transcript. Selected oligonucleotides generated by the RNase T1 digestion represented in Figure 7A were gel purified and further treated to demonstrate that they all contain 5'-OH group accessible to phosphorylation. The 13- and 21-mers excised from G103-63 and G103-75, respectively, or 12- and 20-mers from NE103-63 and NE103-75, respectively, were subjected to phosphatase, CIP (lanes 2, 5, 8 and 11) or CIP followed by T4 polynucleotide kinase treatment (lanes 3, 6, 9 and 12) and resolved in a 25% polyacrylamide/8M urea gel.

13			12*			21			20*			
-	+	+	-	+	+	-	+	+	-	+	+	CIP
-	-	+	-	-	+	-	-	+	-	-	+	T4 Kinase



1 2 3 4 5 6 8 9 10 11 12

Typically, the newly transcribed RNAs contain either a 5'-terminal triphosphate group (pppN) or a cap structure (m⁷GpppN) characteristic for the pol II transcripts. To determine the presence of either, the triphosphate or cap structures at the 5' end of the NE RNA products, they were subjected to Calf Intestinal Phosphatase (CIP) or β -elimination treatment prior to RNase digestion (Fig. 7A and B). As demonstrated by the control experiment shown in Figure 8, removal of the 5' terminal phosphate groups by CIP treatment changes the electrophoretic mobility of the T7 RNA transcript (P), and of its 5'-terminal GAAGGU oligo generated by the subsequent RNase A digestion (Fig. 8, compare lanes 2 and 4 to 1 and 3). A cap structure at the 5' end of a GpppG-primed T7 transcript (C) blocks CIP activity. Consequently, the mobility of the capped GpppGAAGGU oligo remains unchanged after CIP treatment (Fig. 8, compare lanes 6 and 5). Removal of the guanosine of the GpppG cap by β -elimination, results in the increased mobility of the decapped RNA oligonucleotide (Fig. 8, compare lane 7 to 5), as expected for a shorter molecule. Subsequent phosphatase treatment of the decapped RNA oligo additionally changes its mobility, consistent with the removal of 5' phosphate groups exposed by the β -elimination treatment (Fig. 8, compare lane 8 to 7).

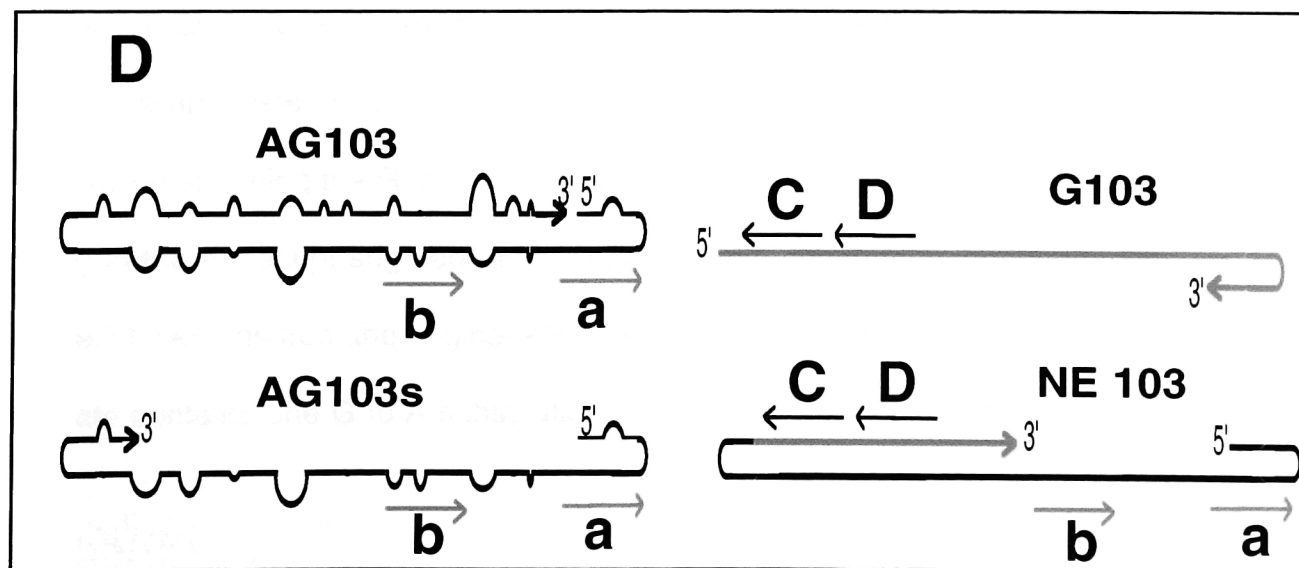
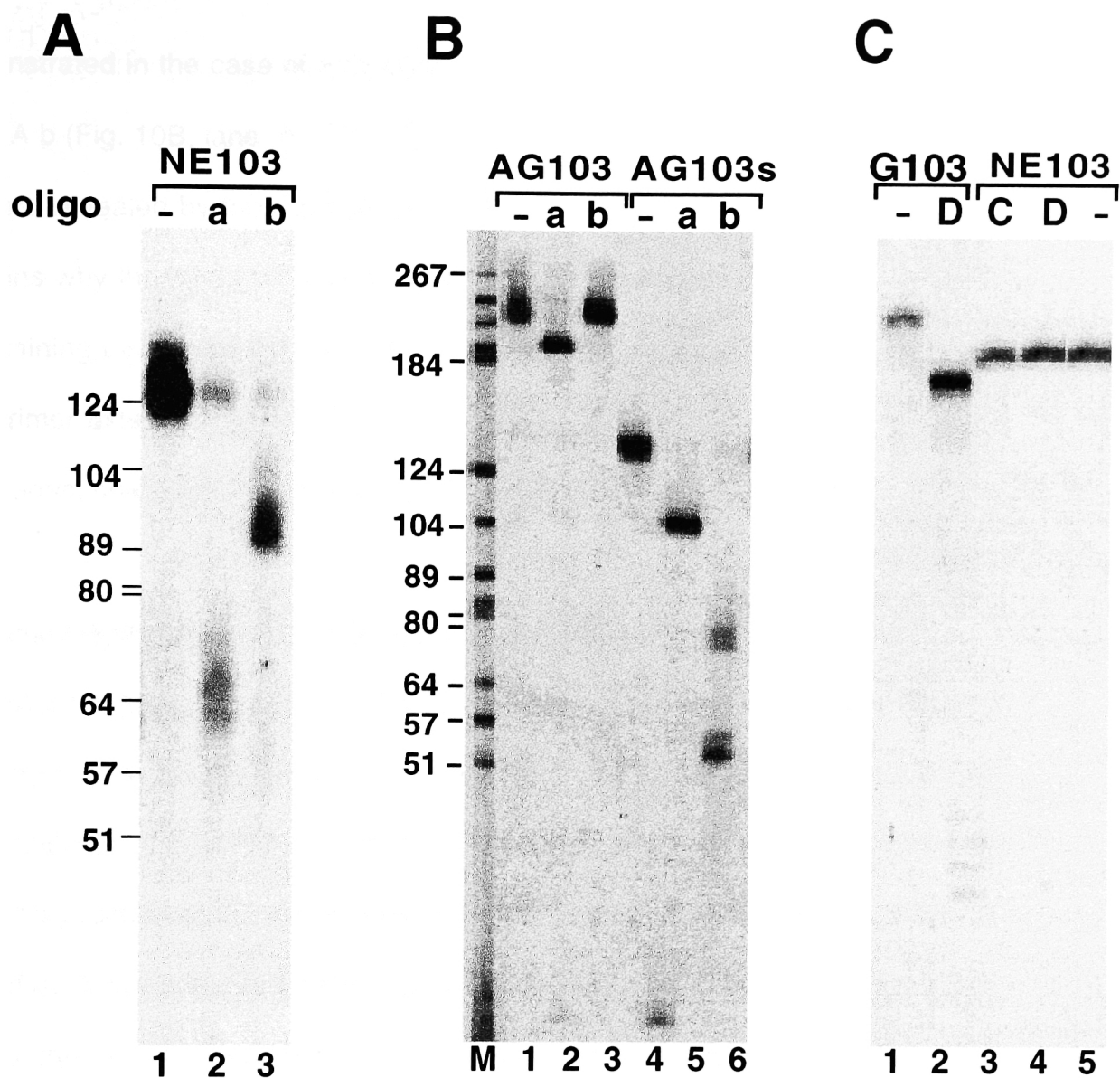
When similar CIP and β -elimination treatments were performed prior to RNase digestion on NE103, NE103-63, NE103-75 products (Fig 7A, lanes 2-4, 6-8 and 10-12 respectively), and NE103-73 (Fig. 7B, lanes 3-5), the electrophoretic mobility of all RNase fragments that span the start site of transcription remained

unchanged (oligonucleotides designated by asterisk). To exclude the possibility that the 5' end of the transcript is blocked by an alternative cap structure that is resistant to the described treatments, it was necessary to demonstrate the presence of a hydroxyl group at the 5' end of the 12 nt and 20 nt fragments. Therefore, these oligonucleotides were gel-purified, together with the control 13 and 21 nt fragments generated by RNase T1 from longer T7 G103-63 and G103-75 RNAs that contain a 5'-OH group accessible to phosphorylation. All of these oligonucleotides were first CIP-treated to remove the RNase-generated 3' phosphate group that could interfere with the detection of the additional 5'-terminal phosphate. As expected, CIP treatment changes the mobility of all the tested oligonucleotides due to the removal of the 3'-terminal phosphate groups, (Fig. 9, lanes 2, 5, 8 and 11). Subsequent treatment with T4 polynucleotide kinase additionally changes their electrophoretic mobility (Fig. 9, lanes 3, 6, 9 and 12 compared to 2, 5, 8 and 11 respectively), demonstrating that a phosphate group is added onto the 5' end of all the tested oligonucleotides. This result clearly demonstrates that all of the RNase T1 oligonucleotides that span the start site of HDV RNA-templated pol II transcription, the AAUCG 5-mer originated from NE103 and the 12 nt and 20 nt fragments from NE103-63 and NE103-75, respectively, contain the 5'-OH groups. Therefore, these oligonucleotides must be excised by the RNase from a longer RNA molecule.

RNase H analysis: Demonstrating that NE RNA product is a chimeric molecule consisting of segment of AG template and pol II transcript

The presence of the hydroxyl group at 5' end of the pol II transcript, together with the detection of only ~40 transcribed nucleotides within the apparently ~125 nt-long NE RNA product, suggest that the pol II transcript is covalently linked to another RNA species that remains unlabeled during NE transcription. The AG template itself appears to be the most likely RNA molecule to be attached to the pol II transcript. To examine this possibility, RNase H analysis of NE RNA products was performed using DNA oligonucleotides complementary to various segments of the G or AG polarity HDV RNA (Fig. 10). The NE103 product is cleaved by RNase H in the presence of DNAs a and b (Fig. 10A, lanes 2 and 3, respectively) that are complementary to regions in AG103 (Fig. 10D), as is the control AG103s (Fig. 10B, lanes 5 and 6, respectively). This result demonstrates that the 5' half of the AG template is covalently linked to the pol II transcript, generating a chimeric template/transcript NE product. In such a product, the pol II transcript is involved in the intramolecular base-pairing with the copied template sequence (Fig. 10D), preventing RNase H cleavage in the presence of DNAs C and D (Fig. 10C, lanes 4 and 5, respectively). DNAs C and D are complementary to the transcribed region and hybridize to control G103 RNA (Fig. 10D), as demonstrated by efficient RNase H cleavage of this RNA (Fig. 10C, lane 2). Resistance to RNase H cleavage in the presence of DNA complementary to a region involved in the intramolecular base-pairing is also

Figure 10. The NE product of the pol II-mediated HDV RNA transcription represents a chimeric template/transcript molecule. Gel purified NE103 transcription product (**A and C**, lanes 3-5) T7 synthesized, control AG103, AG103s (**B**) and T3 synthesized G103 (**C**, lanes 1-2) RNAs were subjected to RNase H digestion in the absence (lanes designated with the -) or in the presence of DNA oligonucleotides a, b, C or D as indicated, and resolved in polyacrylamide/8M urea gels. Positions of pBR322xHaeIII DNA marker are also indicated. **D.** Schematic representation of the control RNAs and the chimeric NE transcription product with the relative positions of the DNA oligonucleotides used (see Material and Methods for sequence information). Antigenomic, AG sense RNAs and DNA oligonucleotides (C and D) are represented in black and the genomic, G, sense RNAs and DNA oligonucleotides (a and b) are represented in red.



demonstrated in the case of a synthetic, full-length AG103 RNA in the presence of DNA b (Fig. 10B, lane 2). The highly structured nature of the NE transcription product, revealed by the combination of RNase T1, A, and H digestion assays, explains why the direct RNA sequence analysis presented above was essential in obtaining conclusive results. Other, probably more commonly used assays, like primer extension or RT-PCR analysis, were pursued at the initial stages of the product analysis but produced inconclusive results (data not shown).

Sequence determination of the template/transcript junction

To precisely determine the template/pol II-transcript junction sequence, RNase T1 digestion was performed on NE products from mutant templates prepared specifically for this purpose. The used mutations included the replacement of selected guanosines (G) with adenosine (A) or uridine (U) residues (Fig. 11C). It should be noted that despite the sequence changes, these mutant templates are fully active in NE transcription (Fig. 11A). The strategy considered in this digestion analysis is that RNase T1 pattern of the mutant NE product will be changed if the mutated G residues (RNase T1 sites) in the AG RNA are not transcribed and instead belong to the template segment of the NE RNA product. If the region spanning the G mutation is transcribed, the mutated positions will be copied in the transcript segment as U or A residues, causing no change in the RNase T1 patterns from the original and the mutant NE products. The AG103-91 template contains one G to A substitution in the base of the terminal loop of the

template AG103-63 (Fig. 11C). The AG103-92 template is derived from AG103-12. In addition to replacing a G residue 3' of the terminal loop with a U, the opposite cytosine is replaced by adenosine to retain base pairing in the stem (Fig. 11C).

Figure 11B represents the RNase T1 digestion analysis of the mutant NE products. A unique 12-mer, that represents the 5'-most ³²P-labeled fragment in the digestion pattern of NE103-63 RNA (lane 2) is replaced by a 13-mer in digestion of the NE103-91 product (lane 3). This change is possible only if the mutated G residue is not transcribed, but rather is covalently attached to the transcribed segment. Similarly, substitution of the G:C base pair in the AG103-12 RNA with the U:A pair in the AG103-92 template (Fig. 11C) results in the replacement of the characteristic AAUCG 5-mer with a UAAUCG 6-mer in RNase T1 digestions of the corresponding NE products (Fig.11B, lanes 4,5). This alteration is also possible only if the nucleotide adjacent to the AAUCG 5-mer (G in AG103-12 and U in AG103-92) is not transcribed but rather originates from the template itself.

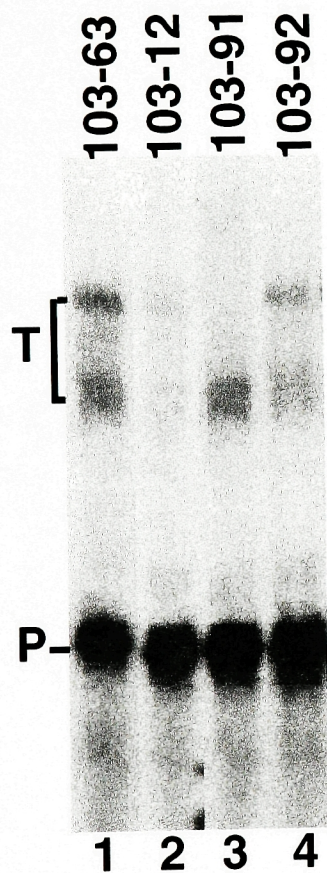
The RNase T1 digestion pattern of another mutant product, NE103-96 (Fig. 12B, lane 3), is fully consistent with these results. The AG103-96 template has the same sequence as AG103-63, except that the G:C pair in the stem following the terminal loop is replaced with a U:G pair (Fig. 12C). The AG103-96 RNA is used for transcription as efficiently as the AG103 and AG103-63 templates (Fig. 12A, lanes 1-3). The RNase T1 digestion of the corresponding

NE products demonstrates that the nucleotide 5' to the AAUCG 5-mer in NE103-96 is not transcribed. Transcription of this nucleotide would incorporate an A residue, eliminating the RNase T1 site and replacing the faster migrating 5-mer with a longer, 12 nt fragment (Fig. 12C) similar to that present in the digestion pattern of the NE103-63 product (Fig. 12B, lane 1). However, such an oligonucleotide is not observed in the digestion pattern of NE103-96 (Fig. 12B, lane 3). Instead, its digestion pattern is identical to that of NE103 (Fig 12B, compare lanes 2 and 3), demonstrating that in both products, the G residue 5' of the AAUCG 5-mer represents the template G. Thus, the initiating nucleotide in the HDV RNA-templated pol II transcription must be located within the AAUCG 5-mer.

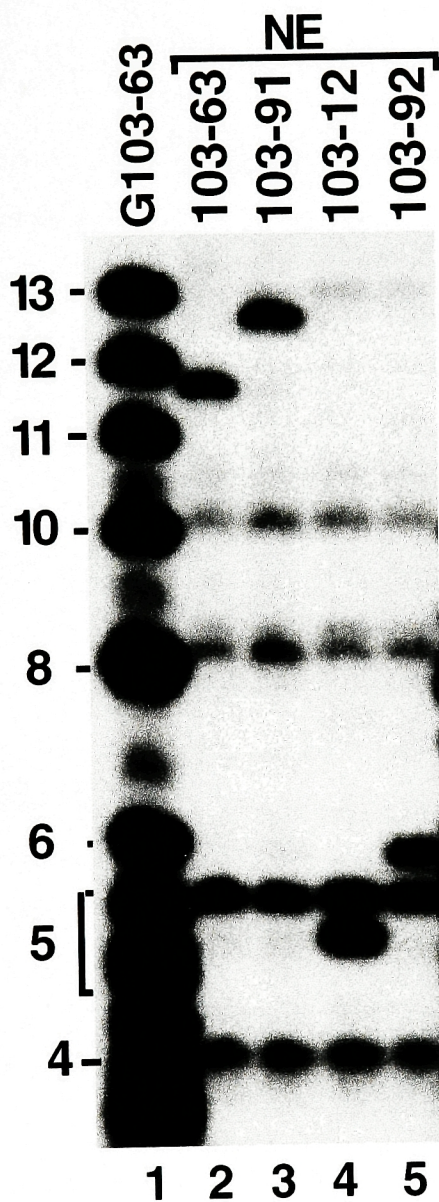
To determine if the A residues within the AAUCG 5-mer of NE103-96 or the 12-mer of NE103-63, are incorporated by pol II transcription, RNase T1 digestion patterns of products synthesized using α - ^{32}P -ATP were compared with those of products synthesized as usual, using α - ^{32}P -GTP (Fig. 13A). The 12 nt fragments generated from the α - ^{32}P -GTP and α - ^{32}P -ATP-synthesized NE103-63 have similar intensities (Fig. 13A, lanes 2, 3), indicating that they both contain one ^{32}P labeled phosphate group. In the case of the α - ^{32}P -GTP 12-mer, the ^{32}P is incorporated by the terminal G residue, while in the α - ^{32}P -ATP 12-mer, the ^{32}P is transferred onto the 12-mer from the 3' adjacent A residue (see sequences in Fig. 13C). This result indicates that the A residues within the NE103-63 characteristic 12-mer are not incorporated by the pol II, but they rather belong to

Figure 11. Determination of the sequence at the template/transcript junction in NE RNA products. **A.** Standard transcription reactions using the templates AG103-63, AG103-12, AG103-91 and AG103-92 (lanes 1-4 respectively). Products were resolved in a 5% polyacrylamide/8M urea gel and detected by autoradiography. **B.** Gel purified NE RNA products from A, 103-63, 103-91, 103-12 and 103-92 (lanes 2-5, respectively) were subjected to RNaseT1 digestion and oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel. Marker RNase T1 oligonucleotides generated by digestion of SP6-G103-63 RNA (lane1). For SP6-G103-63 sequence see Figure 15. **C.** Sequence at the template/transcript junction in the analyzed NE RNA products. Only the terminal hairpin region that contains the mutated nucleotides (highlighted) is presented. For entire sequence of a NE103 product see Figure 17B. The arrows point to RNaseT1 cleavage sites that generate characteristic oligonucleotides at the junction.

A



B



C

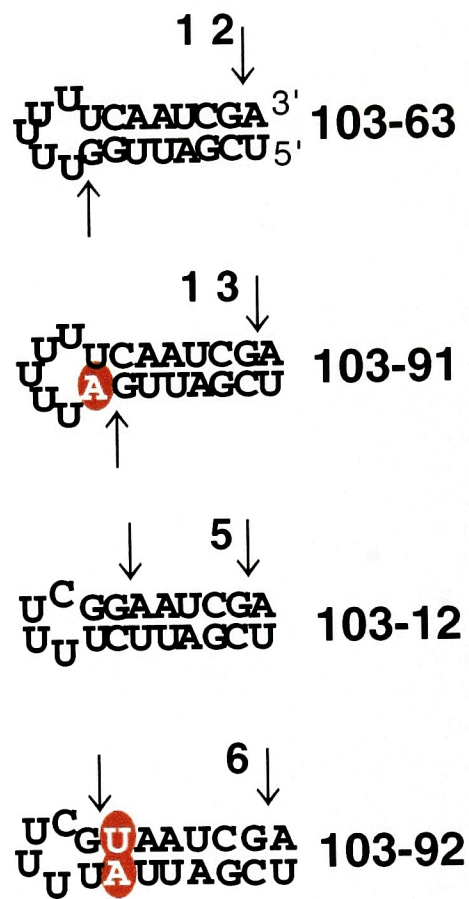
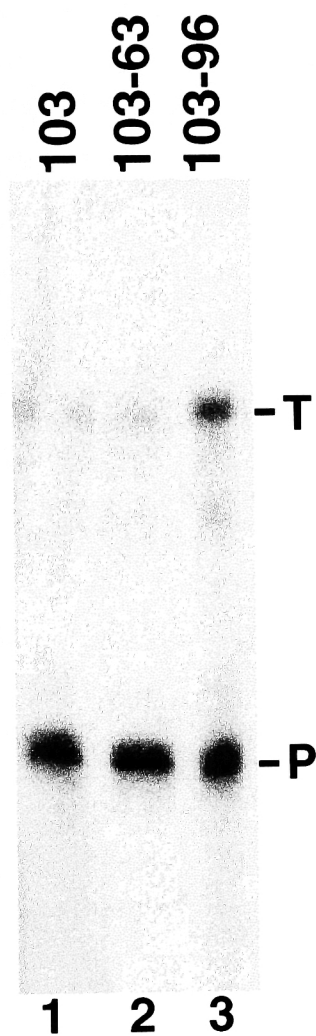
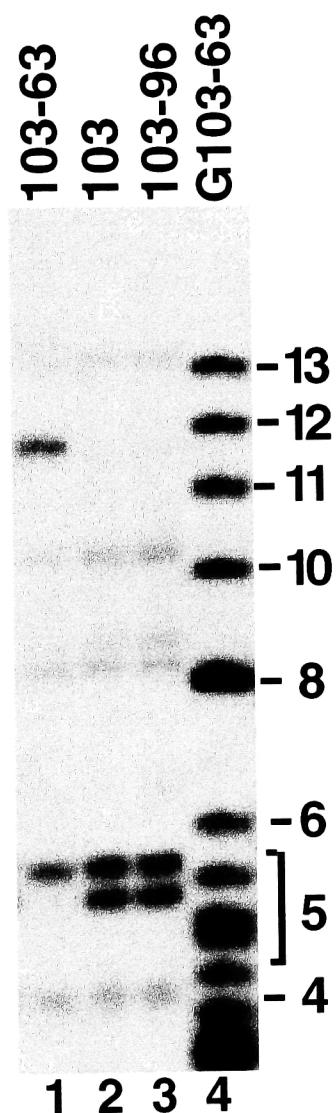


Figure 12. The template/transcript junction point is identical in all mutant NE products. **A.** Standard transcription reactions using the templates AG103, AG103-, AG103-91 and AG103-96 (lanes 1-3 respectively). Products were resolved in a 5% polyacrylamide/8M urea gel and detected by autoradiography. **B.** Gel purified NE RNA products from A, 103-63, 103, 103-96 (lanes 1-3, respectively) were subjected to RNaseT1 digestion and oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel. Marker RNase T1 oligonucleotides generated by digestion of SP6-G103-63 RNA (lane5). For SP6-G103-63 sequence see Figure 15. **C.** Sequences at the template/transcript junction in NE RNA products. Only the terminal hairpin region that contains the mutated nucleotides (highlighted)is presented. For entire sequence of a NE103 product see Figure 17B. The arrows point to RNaseT1 cleavage sites that generate characteristic oligonucleotides at the junction.

A



B



C

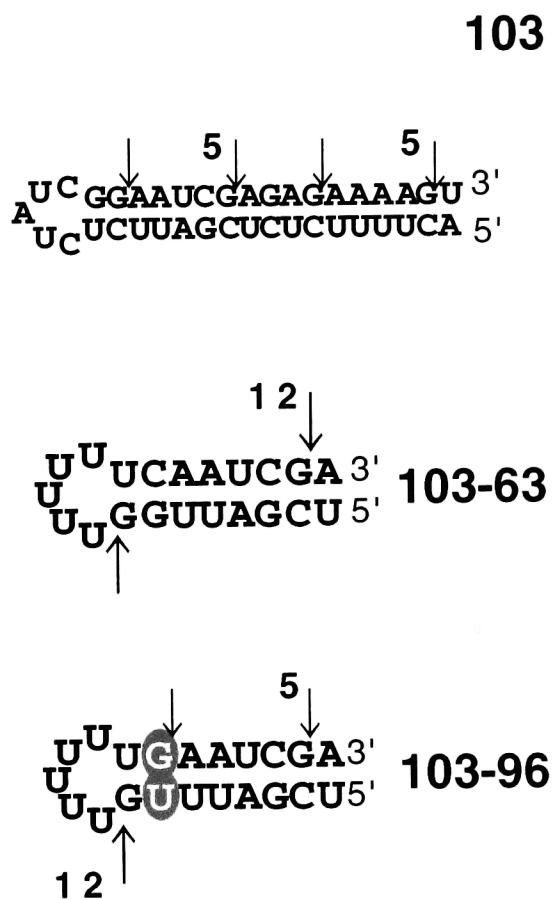


Figure 13. The adenosine residues within the template/transcript junction sequence are not incorporated by the pol II transcription. A. Gel purified NE transcription products, 103-63 (lanes 2 and 3) and 103-96 (lanes 4 and 5), synthesized in the presence of α - 32 P-GTP (lanes 2 and 4) or α - 32 P-ATP (lanes 3 and 5), were subjected to RNaseT1 digestion and resulting oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel. RNase T1 oligonucleotides of α - 32 P-GTP labeled SP6-G103-63 RNA (lane1) are used as size markers. **B.** The internal AAAAG oligonucleotide (lanes 1-4) and the 5'-most AAUCG 5-mers (lanes 5-8) were excised from the gel represented in A and separated again in a 25% polyacrylamide/8M urea gel untreated (lanes 1, 3, 5 and 7) or after phosphatase (CIP) treatment (lanes 2, 4, 6 and 8). The arrows point to RNaseT1 cleavage sites relative to the terminal phosphates of the 5-mers. The positions of the untreated and CIP treated α - 32 P-GTP or α - 32 P-ATP labeled 5-mers are indicated. The 32 P phosphates within them are depicted in red. **C.** Sequences at the template/transcript junction in the NE103 and NE103-96 transcription products. As in the previous figures, only the terminal hairpin is presented. For entire sequence of a NE103 product see Figure 17B. The arrows point to RNaseT1 cleavage sites in the corresponding product sequences.

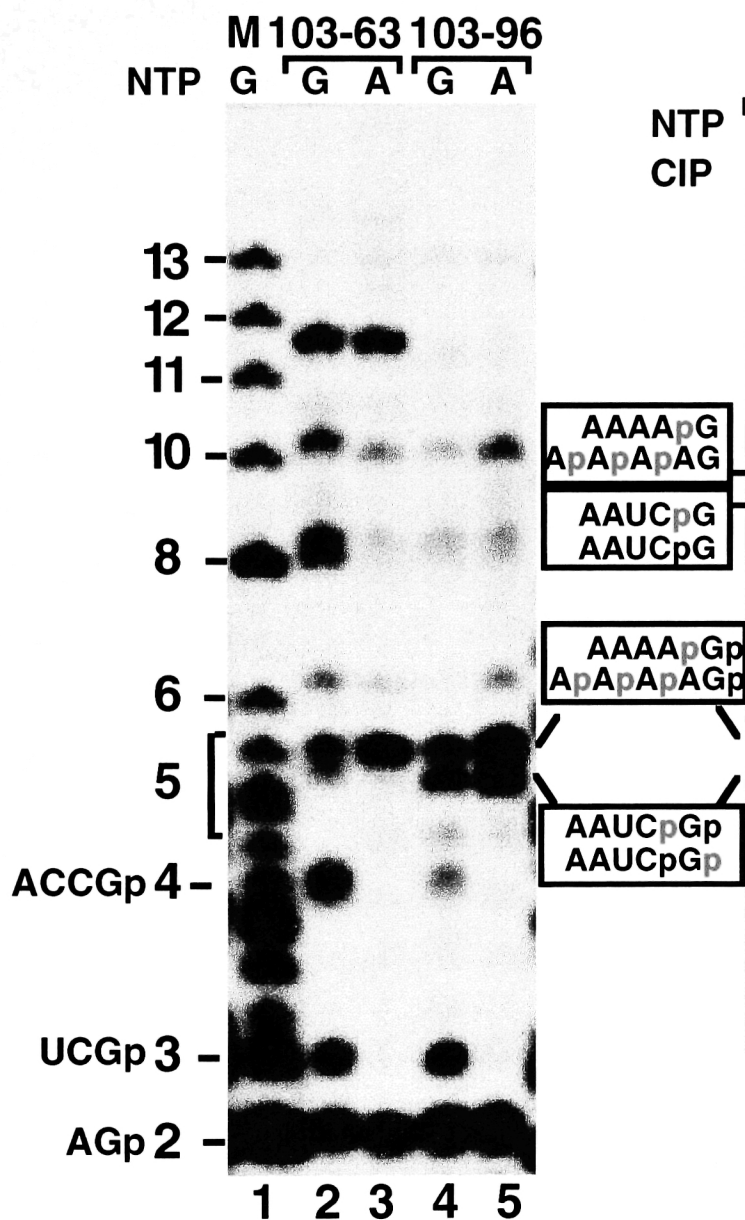
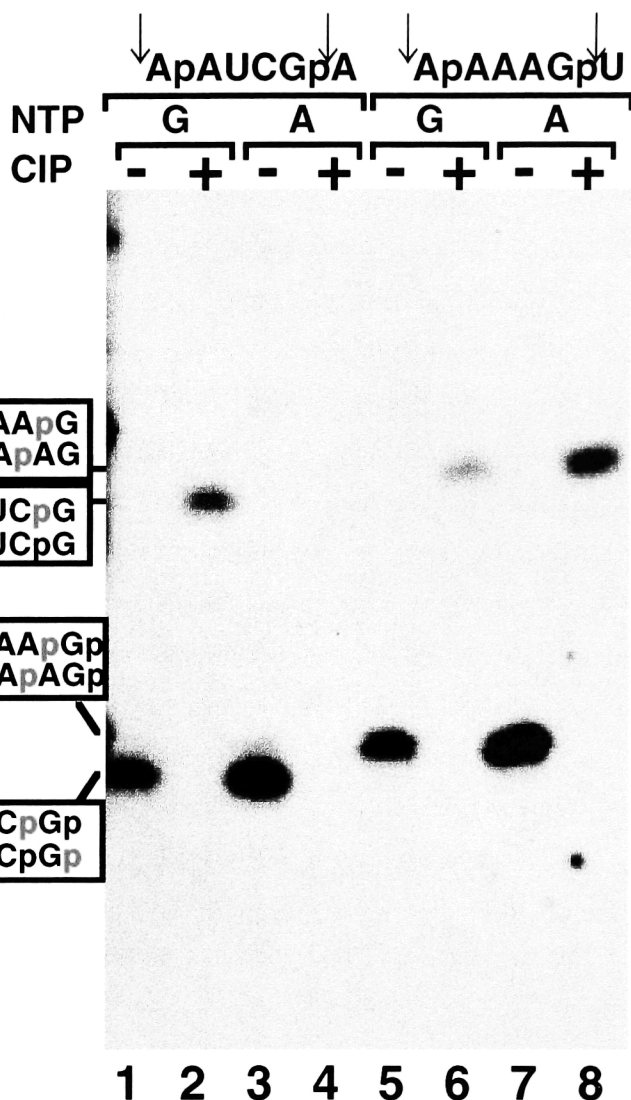
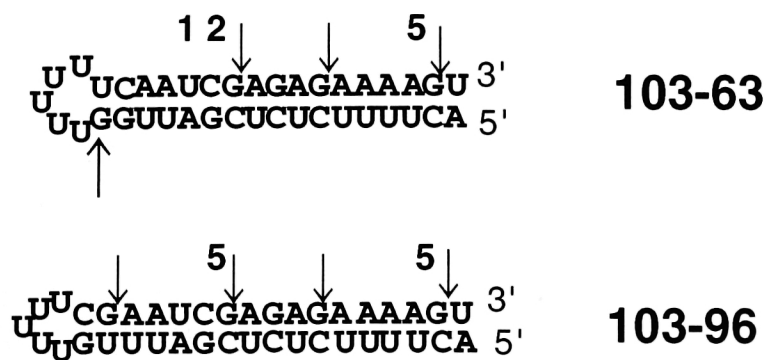
A**B****C**

Figure 14. Pol II initiates HDV RNA-templated transcription at a unique site. **A.** RNase T1 digestion of NE transcription products, 103 (lanes 1, 3 and 5) or 103-96 (lanes 2, 4 and 6), generated in the presence of α -³²P-GTP (lanes 1 and 2), α -³²P-CTP (lanes 3 and 4) or, α -³²P-UTP (lanes 5 and 6). Only the region of 5-mers is shown. **B.** Sequences at the template/transcript junction in the NE103 and NE103-96 transcription products. Only the terminal hairpin regions are presented. For entire sequence of a NE103 product see Figure 17B. The arrows point to RNaseT1 cleavage sites in the corresponding product sequences. The first nucleotide incorporated by the pol II transcription is highlighted by the yellow circle.

the template segment. This situation seems more complicated when RNase T1 digestion patterns of the NE103-96 products synthesized using α - ^{32}P -GTP or α - ^{32}P -ATP are compared (Fig. 13A lanes 4 and 5, respectively). Both oligonucleotides within the 5-mer region appear to be more strongly labeled in the pattern of α - ^{32}P -ATP compared to those of α - ^{32}P -GTP-synthesized NE103-96. Such a strong labeling is expected only for the slower migrating AAAAG 5-mer that may interfere with the appearance of the faster migrating AAUCG 5-mer. To examine if the A residues within the AAUCG 5-mer of NE103-96 product are incorporated by pol II transcription or if they originate from the template, all 5-mers were gel-purified and further analyzed. As expected, upon phosphatase treatment the α - ^{32}P -GTP labeled 5-mers change their mobility and retain the ^{32}P label incorporated during transcription (Fig. 13B, lanes 1, 2 and 5, 6). The phosphatase treated, AAAAG 5-mer generated from the α - ^{32}P -ATP-synthesized NE product also migrates more slowly and retains the ^{32}P label (Fig. 13B lanes 7 and 8), as expected if the ^{32}P A residues were incorporated by pol II during transcription. However, no labeled oligonucleotide could be detected after CIP treatment of the AAUCG 5-mer generated from the α - ^{32}P -ATP synthesized NE103-96 (Fig. 13B, lane 4). This result demonstrates that neither of the A residues in the AAUCG 5-mer, but only the A residue following this 5-mer is incorporated by transcription.

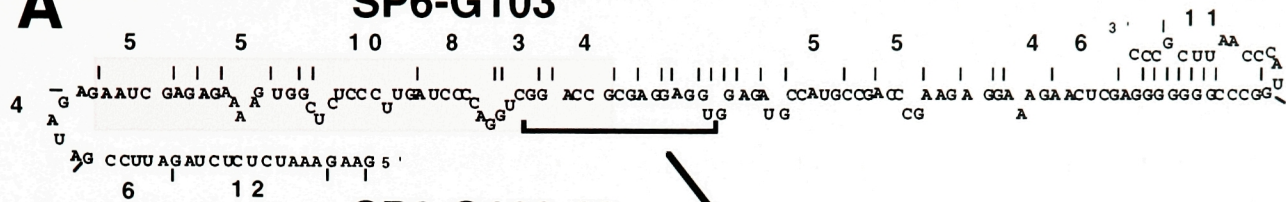
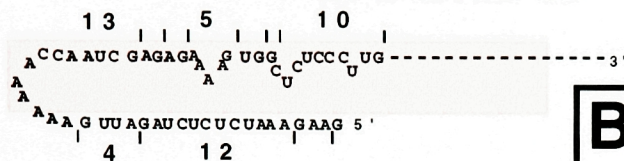
The first pol II-transcribed nucleotide was determined in the experiment presented in Figure 14, which compares RNase T1 patterns of the NE103 and

103-96 products (Fig. 14B) generated in transcription reactions containing α -³²P-GTP, α -³²P-CTP or α -³²P-UTP (Fig. 14A, only the region of 5-mers is shown). As demonstrated earlier for the NE103 and 103-96 products, two RNase T1-generated 5-mers are detected when α -³²P-GTP is used for NE transcription (Fig. 14A, lanes 1 and 2). In RNase T1 digestions of α -³²P-CTP-labeled products (Fig. 14A, lane 3 and 4) the AAUCG fragment is detected, demonstrating that the C residue is incorporated by pol II. The AAAAG 5-mer is not detected in this experiment since it cannot be labeled with α -³²P-CTP. Finally, in both NE products synthesized using α -³²P-UTP, the AAUCG fragment is not labeled (Fig. 14A, lanes 5 and 6), demonstrating that the U residue is not incorporated by pol II and thus must originate from the AG RNA template. The AAAAG 5-mer becomes labeled by ³²P of the following U in the sequence (Fig. 14B). This analysis establishes the C residue of the AAUCG 5-mer as the first nucleotide incorporated by pol II in the AG HDV RNA-templated reaction.

Analysis of the 3' end of the pol II transcript

The 3'-most oligonucleotide of the transcript detected by RNase T1 digestion analysis (Fig. 5) is the ACCG 4-mer presented also in Figure 15A. The sequence following this 4-mer is G-rich, and thus even if transcribed, it would not contain any long RNase T1 oligonucleotides. The RNase A digestion, on the other hand, would generate a characteristic 8 nt-long oligonucleotide, if pol II proceeded for at least 9 nt following the ACCG 4-mer (Fig. 16A). However, no

Figure 15. Mapping the 3' end of the pol II transcript by RNase T1 digestion. **A.** Sequence of control SP6 synthesized genomic 103 and 103-63 RNAs (the 3' segment of the 103-63 RNA that is not shown is identical to that of SP6-G103). RNase T1 cleavage sites are designated by vertical bars and the length (in nt) of the characteristic RNase T1 oligonucleotides are indicated. The boxed regions encompass the segments of G RNA sequence detected in the transcript segment of the corresponding NE products. **B.** The genomic polarity sequence of the mutant AG103^M and AG103-63^M templates. The mutations change the digestion pattern introducing a characteristic 15-mer instead of the wild type ACCG 4-mer. **C.** RNase T1 digestion patterns of transcription products NE103 (lanes 4 and 5), NE103^M (lane 6) or NE103-63^M (lane 7) generated in the presence of α -³²P-GTP (lane 4) or α -³²P-ATP (lanes 5-7). Oligonucleotide size markers were generated by RNase T1 digestion of the control SP6-G103 synthesized in the presence of α -³²P-UTP (lane 1) or the SP6-103^M (lane 2) and SP6-103-63^M (lane 3) synthesized in the presence of α -³²P-GTP. The oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography.

A**SP6-G103****SP6-G103-63****B**

<15

sequence near the
Txn end in both,
103^M and 63^M

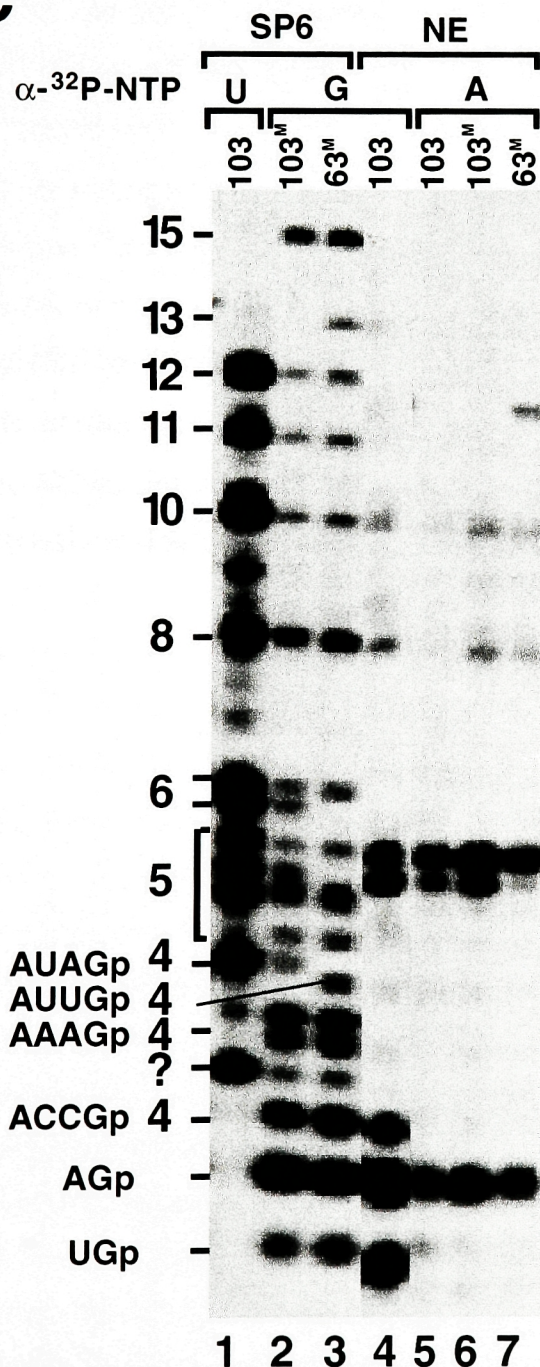
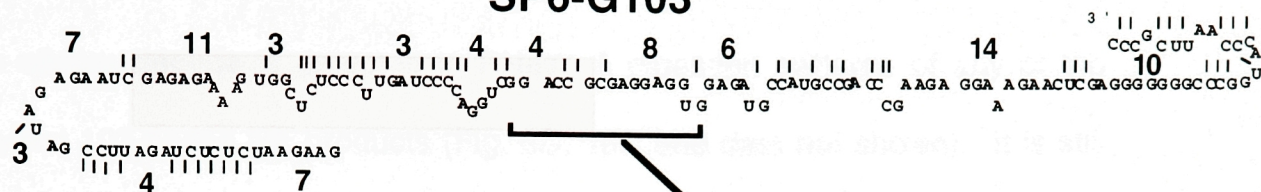
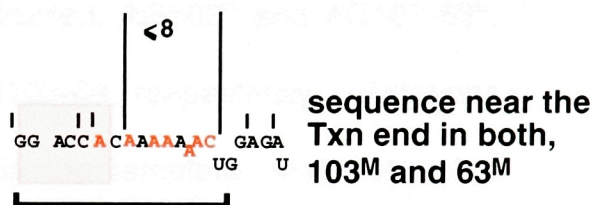
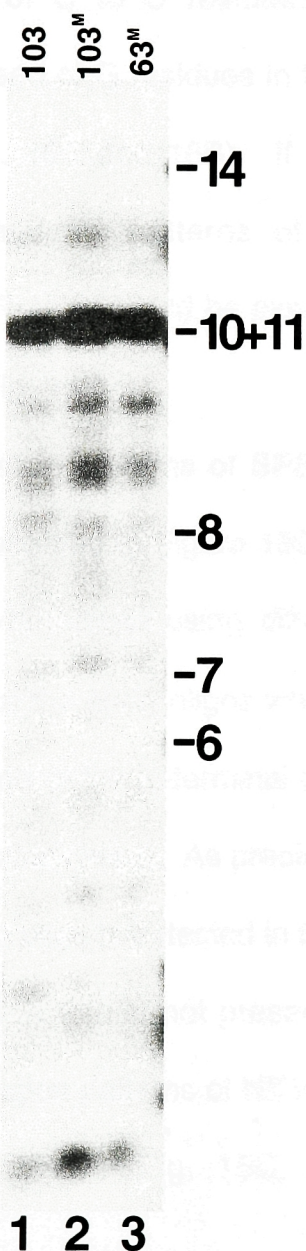
**C**

Figure 16. Mapping the 3' end of the pol II transcript by RNase A digestion. **A.** Sequence of control SP6 synthesized, genomic 103 RNA. RNase A cleavage sites are designated by vertical bars and the length (in nt) of the characteristic RNase A oligonucleotides are indicated. The boxed region encompasses the G sequence identical to the transcript segment of the NE products. **B.** The genomic polarity sequence of the mutant AG103^M and AG103-63^M templates. The mutations introduce five additional A residues in the RNase A 8-mer. **C.** RNase A digestion patterns of transcription products NE103 (lane 1), NE103^M (lane 2) or NE103-63^M (lane 3) generated in the presence of α -³²P-ATP. The oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography. The positions of oligonucleotide size markers are also indicated.

A**SP6-G103****B****C**

such oligonucleotide is detected in RNase A digestion patterns of any of the analyzed 103-based NE products (Fig. 6C, 16C and data not shown). It is still formally possible that transcription proceeds for only a couple of nucleotides downstream of the ACCG 4-mer. To determine the exact 3' end of the pol II transcript, two mutant templates were constructed, AG103^M and AG103-63^M, derived from the templates AG103 and AG103-63, respectively. Mutations include substitutions of C to U residues in the template, that determine incorporation of A instead of G residues in the transcript portion of NE103^M and 103-63^M products (Fig. 15B and 16B). If pol II copies any nucleotide of the mutated region, RNase T1 patterns of the corresponding NE products synthesized using α -³²P-ATP, would be expected to contain longer than 4 nt, A-rich oligonucleotides (Fig. 15B).

RNase T1 digestion patterns of SP6 generated 103, 103^M and 103-63^M genomic RNAs are presented in Figure 15C, lanes 1 to 3, respectively. These control RNAs were synthesized using different α -³²P-NTPs to help precisely determine the identity of the short oligos whose migration is strongly affected by the presence or absence of the 3'-terminal phosphate group, as well as by their sequence (Fig. 15C, as indicated). As predicted from the sequence (Fig. 15B), a characteristic 15 nt long oligo is detected in the digestion pattern of mutant RNAs (Fig. 15C, lanes 2 and 3) that is not present in the SP6-G103 RNA (Fig. 15C, lane 1). However, digestion patterns of NE103^M and 63^M products synthesized in the presence of α -³²P-ATP (Fig. 15C, lanes 6 and 7 respectively) are

indistinguishable from those of NE103 product synthesized under same conditions (Fig. 15C, lane 5), suggesting that transcription does not proceed beyond the ACCG 4-mer into the mutated region. Surprisingly, the ACCA 4-mer that corresponds to the ACCG 4-mer of wt NE103 (Fig. 15C, lane 4) is also not detected in the patterns of α - 32 P-ATP-synthesized NE103^M and 63^M products (Fig. 15C, lanes 6 and 7 receptively). The wt NE103 ACCG 4-mer cannot be labeled by α - 32 P-ATP and is therefore not detected (Fig. 15C, lane 5). This suggests that transcription most likely terminates just 3' of the ACCG sequence in wt NE103 products and within the mutated region in the NE103^M and 63^M products.

Consistent with this result, the RNase A digestion patterns of α - 32 P-ATP synthesized NE103^M and 63^M products are identical to those of NE103 (Fig. 16C, compare lanes 2 and 3 to 1). They do not contain an 8 nt-long or any shorter α - 32 P-ATP labeled oligonucleotides, confirming that pol II does not transcribe the A-rich region following the ACCA sequence.

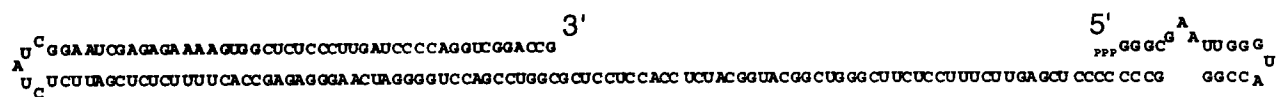
By combining the results from RNase T1, RNase A and RNase H analyses, the sequence of the product generated in the HDV RNA-templated pol II transcription in vitro can be precisely established (Fig. 17B, only the sequence of NE103 product using wt AG103 RNA template is presented). It represents a chimeric molecule that consists of a segment of the AG template covalently linked to the pol II transcript. The chimeric structure of the NE product suggests that when using HDV RNA as a template for transcription in vitro, RNA

Figure 17. Full sequences of the AG103 template and the product generated by the pol II-mediated transcription in HeLa NE. **A.** Sequence of the minimal AG103 HDV RNA template. The numbering system corresponds to nt positions in the RNA template and not to the positions in the full length AG HDV RNA. The arrow indicates the position at which the template is cleaved and pol II initiates transcription by extending on the new group. **B.** Sequence of the chimeric template/transcript product generated by pol II in the nuclear extract using the AG103 as a template. The template sequence (black) is covalently attached to the transcribed segment (red).

AG103



NE 103 product



polymerase II does not initiate de novo. Instead, it uses as a primer the 3'OH end generated by the cleavage of the RNA template at a unique position (designated by an arrow in Fig. 17A) adjacent to a characteristic bulge near the terminal loop. Transcription of the RNA template is precise and terminates after ~ 40 nt.

2.1.3 Effect of hepatitis delta antigen (HDAg) on transcription in vitro

The observed HDV RNA-templated transcription in vitro depends on pol II and other protein factors present in standard HeLa cell nuclear extracts. For efficient HDV RNA replication in the cell however, HDAg protein encoded by the virus is also required (Kuo et al., 1989). Therefore, determining the effect of HDAg on transcription in vitro was pertinent. As demonstrated in Figure 18A, increasing amounts of His-tagged recombinant HDAg protein (kindly provided by J. Doudna) stimulated HDV RNA-templated transcription in vitro. The strongest stimulation (~4 fold stimulation) was observed when the standard transcription reaction (15 μ l) was supplemented with 0.1 μ g HDAg (Fig. 18A, lane 5). Interestingly, the transcription product generated in the presence of HDAg (δ P) exhibits an altered electrophoretic mobility when compared to the NE103 product. In fact, the δ P product consists of two closely migrating RNA species, δ P_a and δ P_b, that can be clearly distinguished in transcription reactions (15 μ l) containing 1 μ g HDAg and standard (Fig 18B, lane 2) or limiting amounts of NE (Fig. 18C, lane 2). As expected, the RNase A pattern of the δ P product contains oligonucleotides

characteristic for both δP_b and δP_a products (Fig. 19B, lanes 2, 3 and 4, respectively). The RNase A digestion analysis of δP_a product that co-migrates with NE103 (Fig. 18B and C, lanes 2 and 1, respectively) demonstrates that both these RNAs contain the same 40 nt transcript, since they both have identical digestion patterns consisting of the characteristic 11 nt-long and a few shorter oligonucleotides (Fig. 19B, compare lanes 4 and 1). The same analysis of the faster migrating, δP_b product, demonstrates that it includes a longer transcript segment, since its digestion pattern contains additional 6 nt and 8 nt-long oligos (Fig. 19B, lane 3), expected only if pol II transcription proceeds beyond the 40 nt region characteristic for the NE103 product (Fig. 19A). The pattern of 4 nt-long RNase A fragments is also consistent with the extended pol II transcript in the δP_b product. The digestion patterns of both NE103 and δP_b products (Fig. 19B, lanes 1 and 3 respectively) contain two abundant 4 nt-long oligos. The 4-mer common to both products represents the AGGU oligonucleotide containing 3' phosphate group, consistent with it being excised by RNase A from an internal position in the transcript (Fig. 19A). The faster migrating GGAC 4-mer is also internally excised (Fig. 19A) and contains a 3' phosphate group (GGACp). The GGAC 4-mer in the NE103 digestion pattern migrates more slowly, consistent with the presence of the 3'OH rather than the phosphate group, demonstrating that the C residue is the ultimate nucleotide incorporated by transcription in this product (Fig. 19A). Transcription terminates at the same position in δP_a product, since the GGAC_{O_H} 4-mer represents the

Figure 18. Delta antigen stimulates the HDV RNA-templated pol II transcription in vitro. **A.** Transcription using the AG103 template in standard nuclear extract (lane 1) or supplemented with different amounts (as indicated) of His-tagged HDAg (lanes 2-8). Products were resolved in a 5% denaturing/8M urea gel and detected by autoradiography. Positions of the labeled template (T), NE RNA product (P) and HDAg supplemented NE product (δP) are indicated. **B.** As in A but only the region of transcription products is shown. The transcription reactions were supplemented with 0 (lane 1) 1 μ g (lane 2) or 0.1 μ g (lane 3) HDAg. Two differently migrating transcription products (δP_a and δP_b) can be distinguished in the reactions supplemented with 1 μ g HDAg. **C.** As in B only the transcription reactions contain limiting amount (6.7% instead of 40%) of nuclear extract.

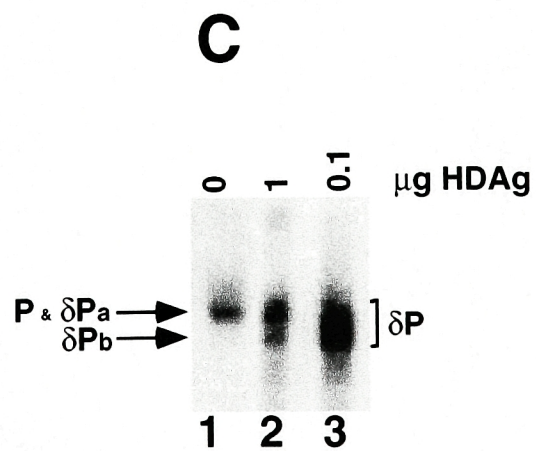
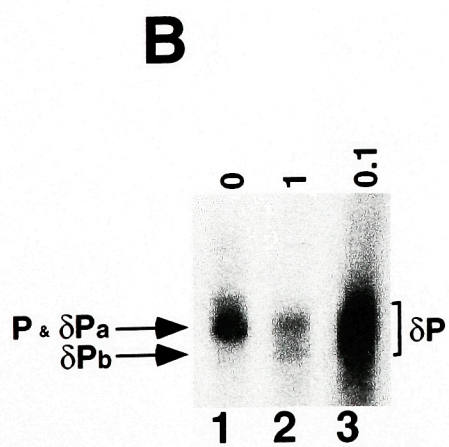
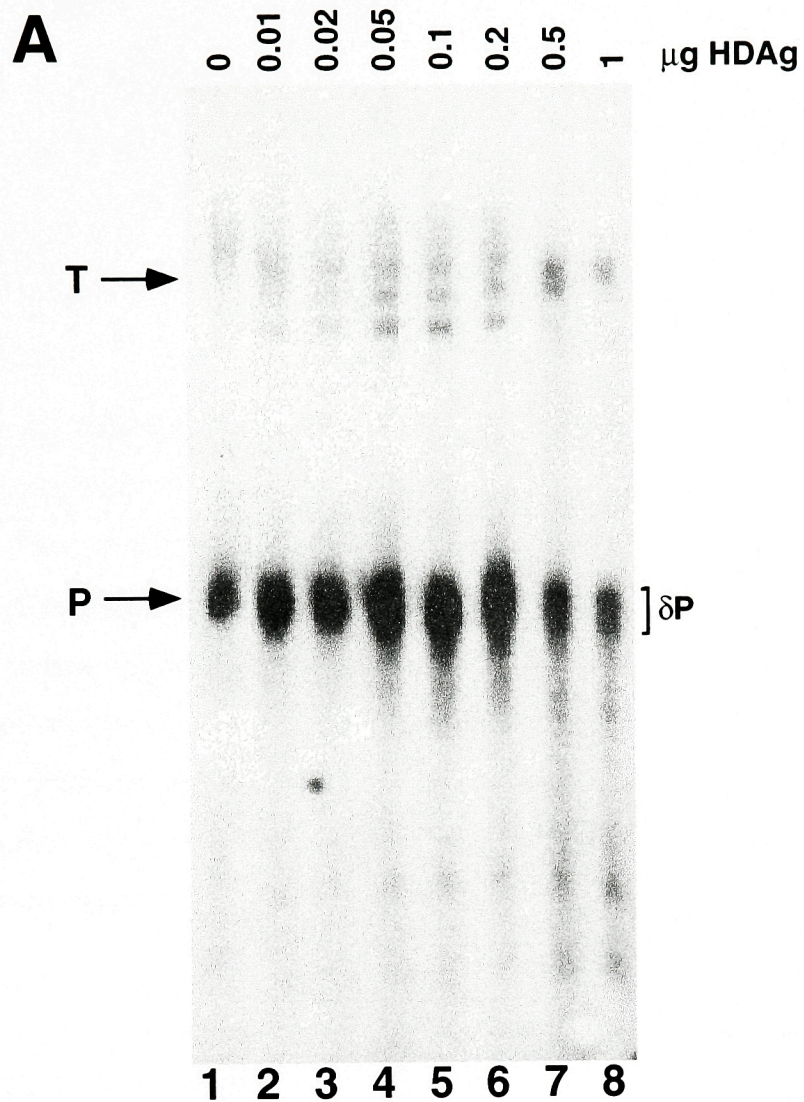
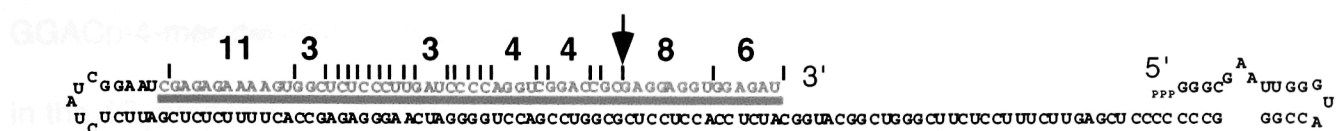


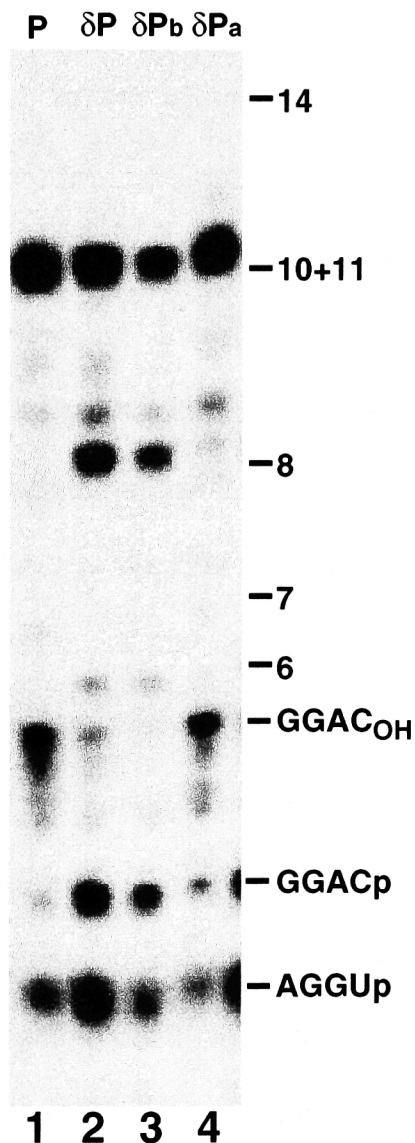
Figure 19. Delta antigen stimulates pol II elongation on the HDV RNA template. **A.** Sequence of the chimeric template/transcript product generated by pol II in the HDAg supplemented nuclear extract (δP_b and δP). The template sequence is represented in black and the transcribed segment is represented in red. The arrow indicates the 3' end of the NE product generated in the absence of HDAg. **B.** RNase T1 digestion of the transcription products from Figure 18, P (lane 1), δP (lane 2), δP_a (lane 4) and δP_b (lane 3). The resulting oligonucleotides were resolved in a 25% polyacrylamide/8M urea gel and detected by autoradiography. The positions of oligonucleotide size markers are also indicated.

A

NE + HDAg product (δP and δP_b)



B



predominant 4 nt oligo in its digestion pattern (Fig. 19B, lane 4). The low level of GGACp 4-mer detected is most probably a result of the presence of δP_b product in the δP_a preparation, as they migrate very close to each other (Fig. 18B and C, lane 2 in both). Alternatively, in a small subset of δP_a products transcription may extend for a few more nucleotides, which does not change the electrophoretic mobility of δP_a product but is reflected in its digestion pattern by the presence of GGACp 4-mer. Small amounts of GGACp 4-mer are also detected in the digestion pattern of NE103 product, demonstrating that while pol II predominantly terminates after incorporating the C residue of the 4-mer, it can with a low efficiency incorporate additional nucleotides. This suggests that rather than terminating at a specific nucleotide, pol II transcription is negatively affected in the region following the GGAC 4-mer. The presence of HDAg in the transcription reaction appears to stimulate pol II activity on the RNA template and allows for the synthesis of longer extended transcripts found in the δP_b and δP products.

The presence of 20 ^{32}P labeled G residues in δP , instead of 11 labeled residues in the NE103 product, (Fig. 19A) cannot by itself fully explain the observed ~4 fold stimulation of transcription observed in the presence of HDAg. Therefore, in addition to affecting pol II elongation, HDAg must also increase the efficiency of template recognition and initiation of transcription by the polymerase.

The faster electrophoretic mobility of δP_b product containing the longer transcript can be explained by the unusual structure of the chimeric

template/transcript product (Fig. 19A). The presence of a longer double stranded segment in the δP_b product, as compared to the NE103 and δP_a products, results in a more compact structure and an increased electrophoretic mobility.

2.1.4 Template requirements for the HDV RNA-templated pol II transcription

Secondary structure of the RNA template determines functional pol II templates

Remarkably, despite multiple mutations in the primary sequence, all of the RNA templates used in the product analysis are cleaved and extended at the same unique site near the terminal stem. The site of cleavage/extension is located immediately adjacent to the characteristic bulge near the terminal loop (Fig 17A and 20A for a schematic representation), suggesting a possible role of this structure in the start site selection. To more closely examine the structure/function relationship of the RNA templates, the AG103-20 template was constructed, in which the UA bulge positioned 3' of the terminal loop is deleted and a 2 nt insertion is introduced at the opposite side of the stem, creating a GC bulge positioned 5' of the terminal loop (see Fig. 20A for schematic representation and, Fig. 21C for sequence). These mutations are less extensive than the sequence changes found e.g. in the AG103-63 and -75 templates, but they significantly alter the local secondary structure of the RNA template, specifically affecting the characteristic bulge adjacent to the start site. Consistent

with the predicted importance of this bulge, the AG103-20 RNA template fails to support transcription in HeLa NE (Fig. 20B, lane 2). Furthermore, AG103-77 RNA, in which the 62 nt terminal stem-loop region (pos. 90-152) is inverted, and thus present in the genomic polarity, also fails to support transcription (Fig. 20B, lane 3). Remarkably, while the AG103-20 and AG103-77 templates have substantially different primary sequences in the terminal stem-loop segment, they both can be folded into structures similar to that of AG103 RNA, but containing a bulge at the opposite side of the terminal stem (Fig. 20A).

To further examine the importance of the specific positioning of the bulge, RNA templates containing the sequence complementary to that found in AG103, AG103-20 and AG103-77 RNAs were synthesized using T3 RNA polymerase (Fig. 20A) and tested for transcription in vitro, in HeLa cell NE (Fig. 20B, lanes 4-6). As expected from the predicted secondary structure, the T3-synthesized G103 RNA that has an identical terminal hairpin sequence and structure as AG103-77 template, fails to support specific transcription in vitro (Fig. 20B, lane 4). The signal detected in this reaction was too low to be extensively analyzed. However, RNase T1 digestion mapping (not shown) indicated that the signal does not correspond to the typical NE product, but rather to a degradation product of the 3' end-labeled template. Furthermore, T3-synthesized 103-20 and -77 RNA templates that can acquire the same secondary structure as the T7-synthesized AG103 RNA (Fig. 20A), efficiently support transcription in NE (Fig.

20B, lanes 5,6), confirming the importance of the characteristic terminal stem-loop structure in specifying functional pol II templates.

This is further demonstrated by the compilation of the predicted secondary structures within the terminal hairpin region of a number of mutant templates (Fig. 21), whose ability to serve as pol II templates was determined in the course of the NE product analysis (Fig. 11A, 12A and 22). The terminal hairpin region of all HDV RNAs that represent pol II templates can acquire a secondary structure that resembles that of the wt, AG103 template, although their primary sequences can be significantly different (Fig. 21B). The secondary structure of these 'functional' pol II templates is characterized by the presence of the terminal loop (TL), a bulge (B) and two internal loops (IL1 and IL2), whose sequence and size can vary significantly. For example, transcription is not inhibited by the sequence changes in the TL of mutant RNAs 12, 63, 73, 75, 91, 92 and 96 (Fig. 11A, 12A and 22, lanes as indicated and Fig. 21 for sequences). In addition, templates 73 and 75 have an altered sequence in the stem region between TL and internal loop 1 (IL1). Their predicted secondary structure, however, fully resembles that of the wt 103 template with the characteristic bulge (B) 3' of the TL (Fig. 21), and they represent functional pol II templates (Fig. 22, lanes 3 and 4, respectively). Similarly, shortening the stem between TL and B by deleting the G:U base pair at the base of TL in RNA16 does not affect transcription efficiency (Fig 22, lane 2). Furthermore, changing the sequence of the bulge (B) from UA to CC in RNA 75 does not significantly affect the efficiency of transcription (Fig. 22, lane 4). In

addition, inserting 4 nt at the position of the bulge in RNA13 has no negative effect, but rather slightly stimulates transcription from the AG103-13 template (Fig. 22, lane 8). Similar effect is observed with RNA 18 (Fig. 22, lane 7) that contains a 2 nt deletion in the stem opposite of the UA bulge. Interestingly, RNA18 can acquire a structure similar to that of RNA13 with a larger, 4 nt bulge (designated 18b in Fig. 21 and demonstrated by structure mapping in Fig. 23), in addition to the wt-like structure characterized by 2 nt UA bulge (designated 18a in Fig. 21C). The observed increased efficiency of RNAs 13 and 18 suggests that in addition to the typical stem-bulge structure that specifies functional pol II templates, other components of the overall geometry of the RNA template affect the efficiency of transcription. At this point we do not fully understand the mechanism by which the secondary structure affects pol II activity on the RNA template. It is tempting to speculate, however, that the larger bulge in RNAs 13 and 18 additionally “kinks” the stem between TL and IL1 that contains the start site of transcription, and in that way positively affects pol II recognition of the template and the efficiency of template cleavage at the transcription start site.

In RNA 18 and 75, the symmetry of the internal loop IL1 is disrupted (Fig. 21B). However, transcription from the AG103-18 and AG103-75 templates is not inhibited (Fig. 22, lanes 7 and 4, respectively), demonstrating that the structure of the IL1 is not crucial in determining the efficiency of pol II transcription from the HDV RNA template *in vitro*.

Figure 20. Specific secondary structure of HDV RNA, rather than its primary sequence, determines functional pol II templates. A. Schematic representation of the cDNA used to synthesize HDV RNA templates from T7 (green) or T3 (blue) promoter and the predicted secondary structures of RNA templates used in B. The asterisks indicate positions of the bulge important for template specificity. **B.** Two opposite polarities of RNA, synthesized by T7 (lanes 1-3) or T3 (lanes 4-6) were used as templates for transcription in HeLa NE. Products were resolved in a 5% denaturing/8M urea gel and detected by autoradiography. Positions of the labeled template (T), NE RNA product (P), and pBR322xHaeIII DNA marker (lane M) are indicated.

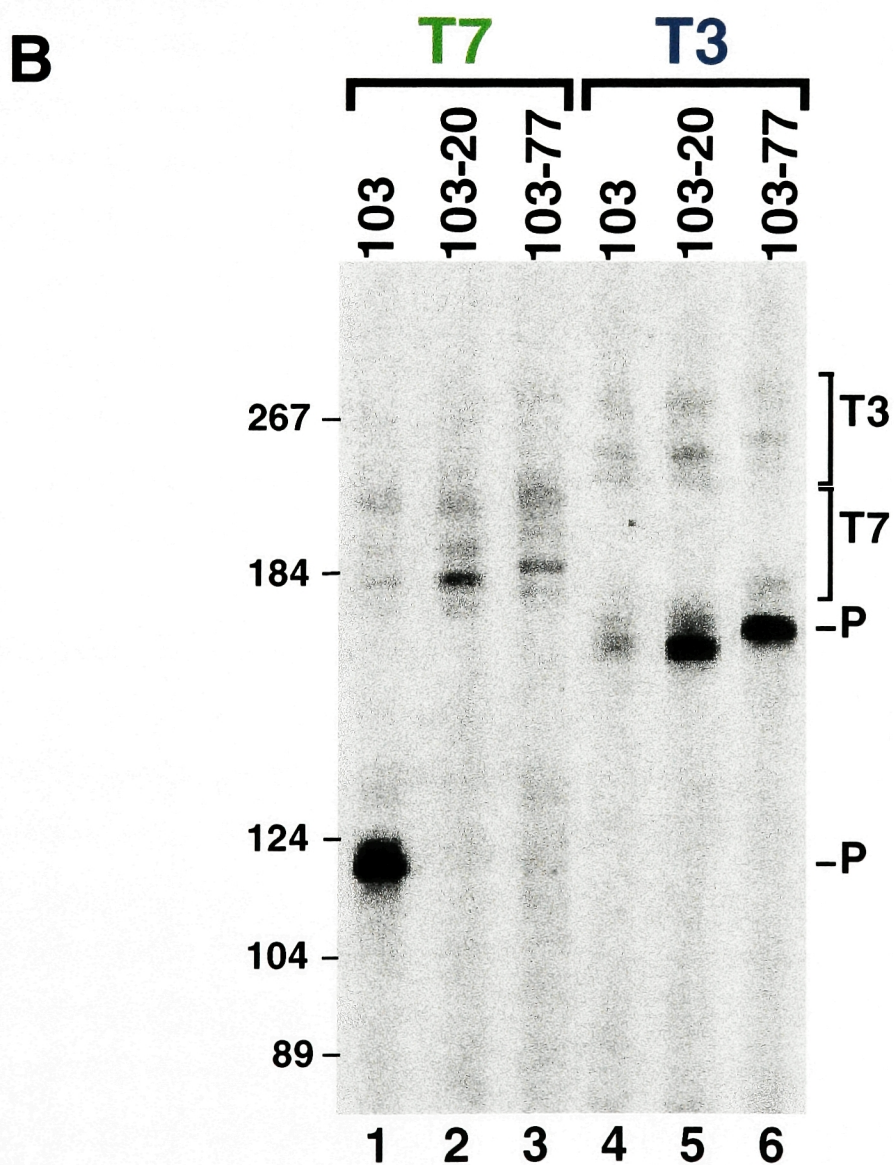
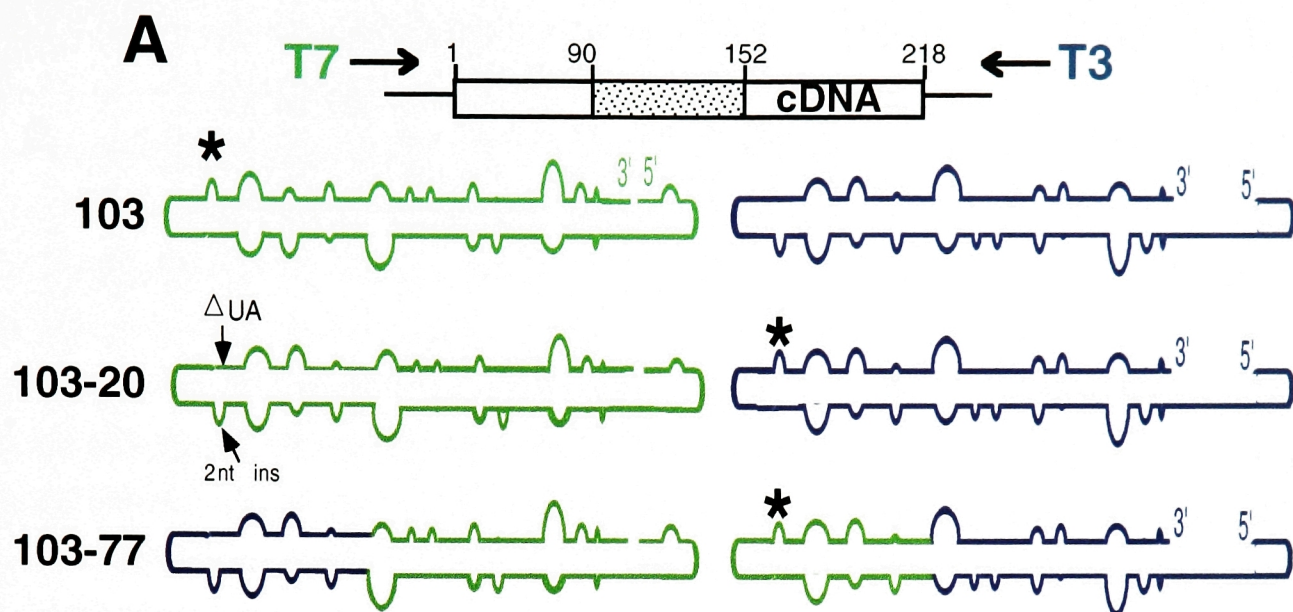
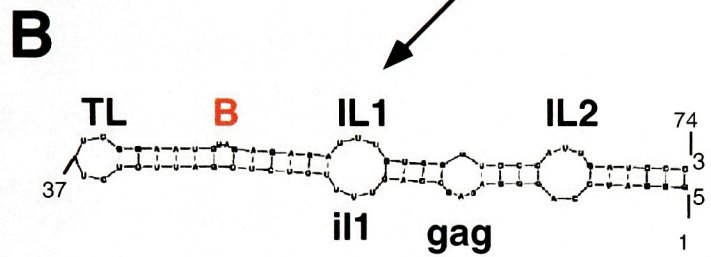
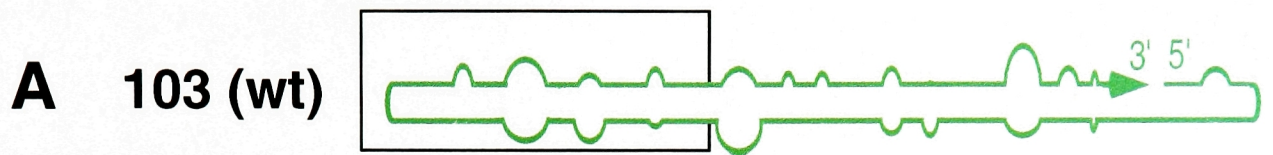
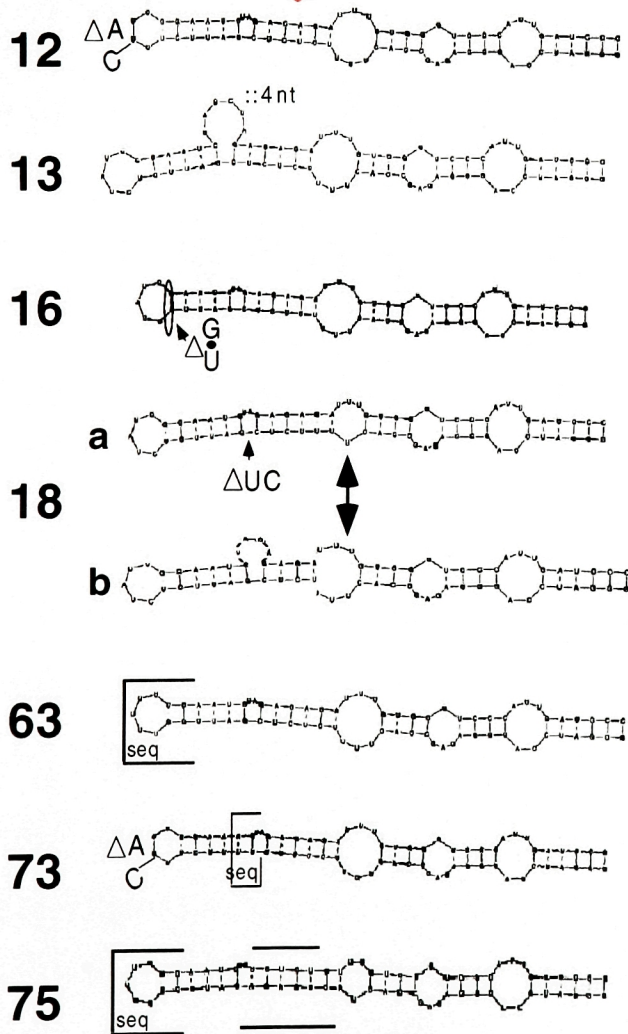


Figure 21. Compilation of secondary structures of AG HDV RNA templates. **A.** Schematic representation of the secondary structure of the wild type AG103 template as predicted by the Mufold program. **B.** Sequence in the terminal hairpin region (pos. 85-158 of the AG103 template, boxed region in A) of the wild type and mutant functional pol II templates. 18a represents the lowest energy secondary structure of the terminal hairpin of AG103-18 template predicted by the Mufold program. 18b represents its alternative conformation detected by the chemical mapping shown in Figure 24. **C.** Sequence of the terminal hairpin region of the nonfunctional pol II templates. TL-terminal loop; B-bulge 3' of the TL; b-bulge 5' of the TL; IL1-internal loop 1; IL2-internal loop 2. Deletions in the sequence are represented by Δ and, insertions by $::$. seq indicates that the region contains only sequence alterations without deletions or insertions.



Functional templates



C

Nonfunctional templates

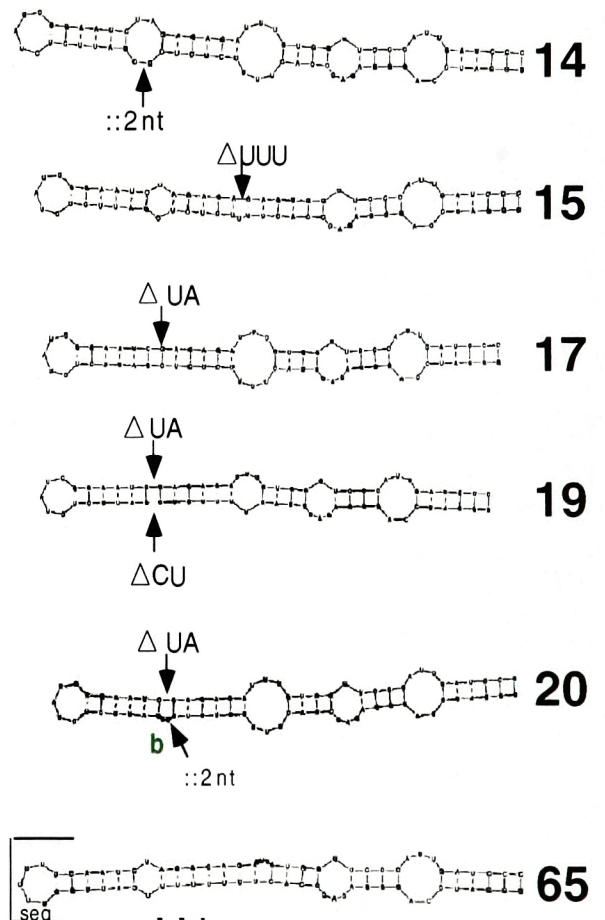
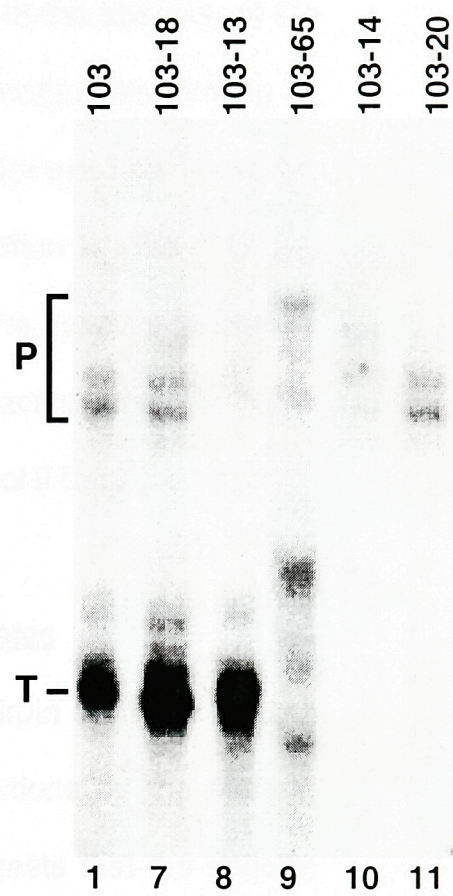
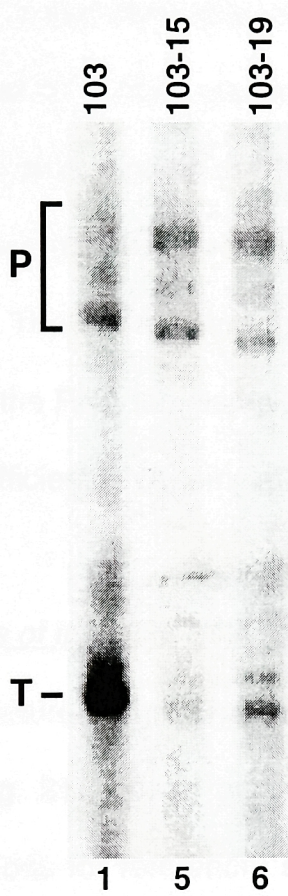
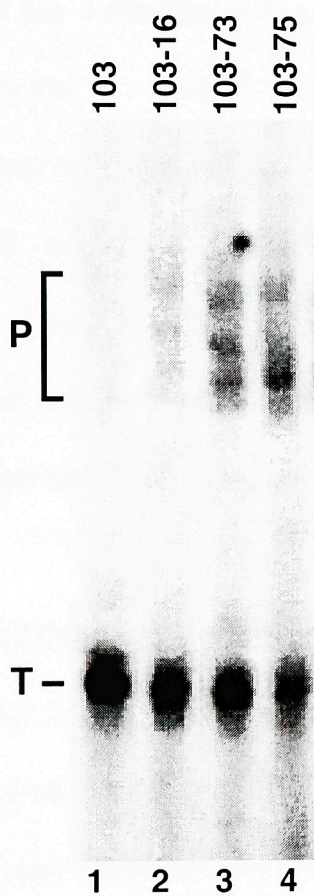


Figure 22. Transcription reactions in NE using structurally distinct RNA templates. The transcription reactions were performed under standard conditions in HeLa nuclear extract and templates whose terminal hairpin secondary structures are shown in Figure 21. Transcription products were resolved in a 5% denaturing/8M urea gel and detected by autoradiography. Positions of the labeled template (T) and, NE RNA products (P) are indicated. Transcription reaction using the wild type AG103 template is included for comparison (lane 1 in each gel).



The importance of the bulged nucleotides in HDV RNA templates at the transcription start site is further supported by the compilation analysis of the secondary structures of terminal hairpin in HDV RNAs that do not represent functional pol II templates (Fig. 21C). All these RNAs contain a variety of mutations (as indicated in Fig. 21C) that invariably result in the absence of the bulge 3' to the TL. Even RNAs that do contain unpaired nucleotides forming an internal loop or a mismatch at the corresponding position, for example RNAs 14, 15 and 65, are not recognized as templates for transcription in vitro (Fig. 22, lanes 10, 5 and 9, respectively). Thus, it appears that the specific secondary structure of the terminal hairpin of the RNA templates characterized by the bulge 3' of the terminal loop is required efficient RNA-templated pol II transcription.

Mapping of the secondary structure of the HDV RNA templates

Comparison of the secondary structure in the terminal hairpin of functional and nonfunctional RNA templates (Fig. 21), based on predictions by the MuFold program (see Materials and Methods for reference) suggests that the specific structure rather than the primary sequence of HDV RNA allows pol II to recognize and select the template for transcription. Mapping of the actual secondary structures of these RNA molecules shows an excellent agreement with the computer predictions (Figures 23 and 24), supporting the importance of the RNA secondary structure for efficient pol II transcription.

Figure 23. Enzymatic mapping of the secondary structure in the terminal hairpin of the HDV RNA templates. 5' end labeled RNAs encompassing the terminal hairpins of the wt and mutant templates represented in Figure 21 were subjected to partial digestion by RNase PhyM (lanes 2-6) or RNase T1 (lanes 7-18). The resulting fragments were resolved in a 10% polyacrylamide/8M urea gel and detected by autoradiography. A sequence ladder generated by alkaline hydrolysis is also included (lane 1). The numbering system is relative to the wild type hairpin segment (see Figure 21B). The positions of the fragments corresponding to the RNAs cleaved at the characteristic secondary structure elements (TL, B, b, IL1 and IL2) described in Figure 21 are indicated.

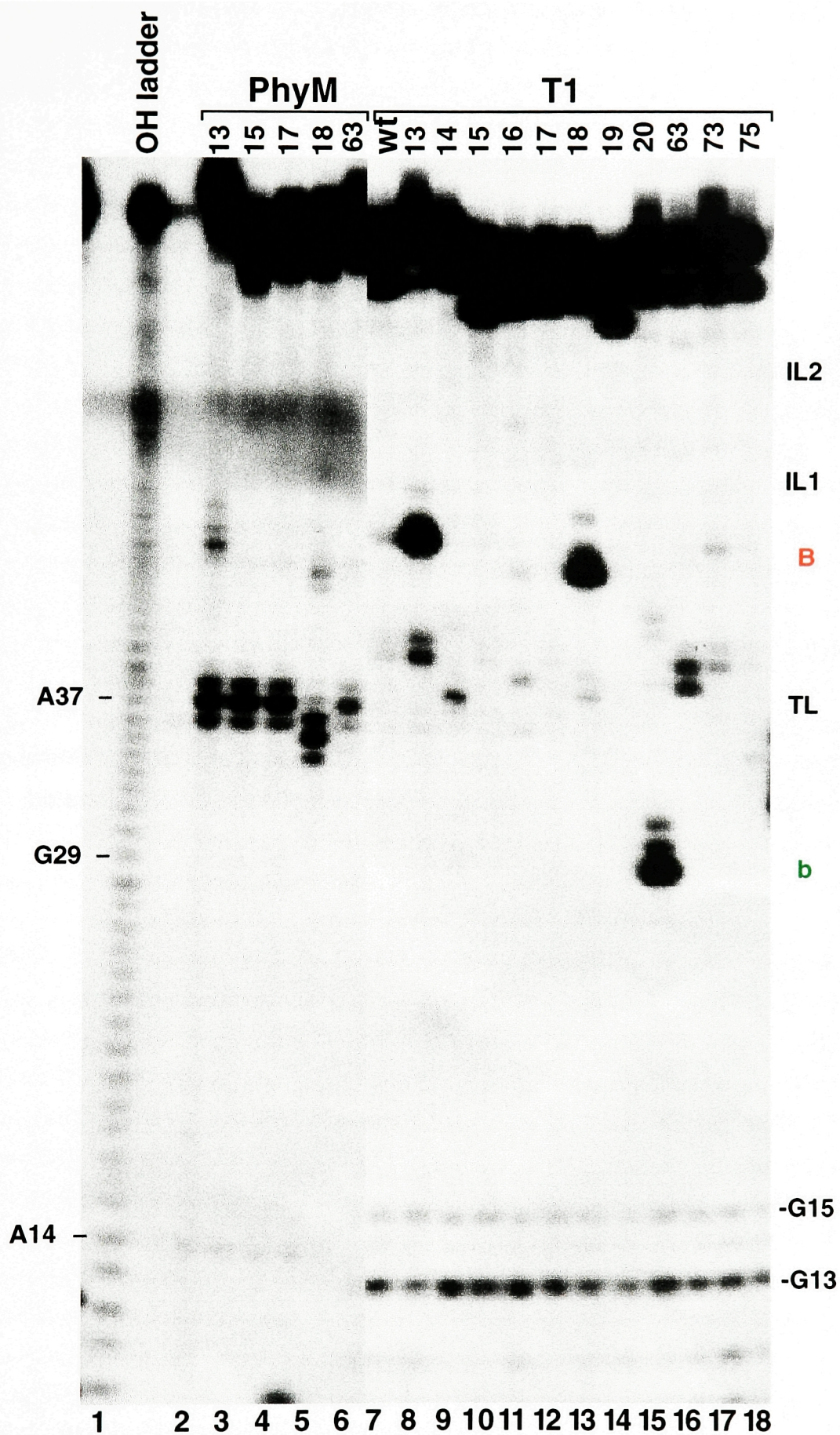
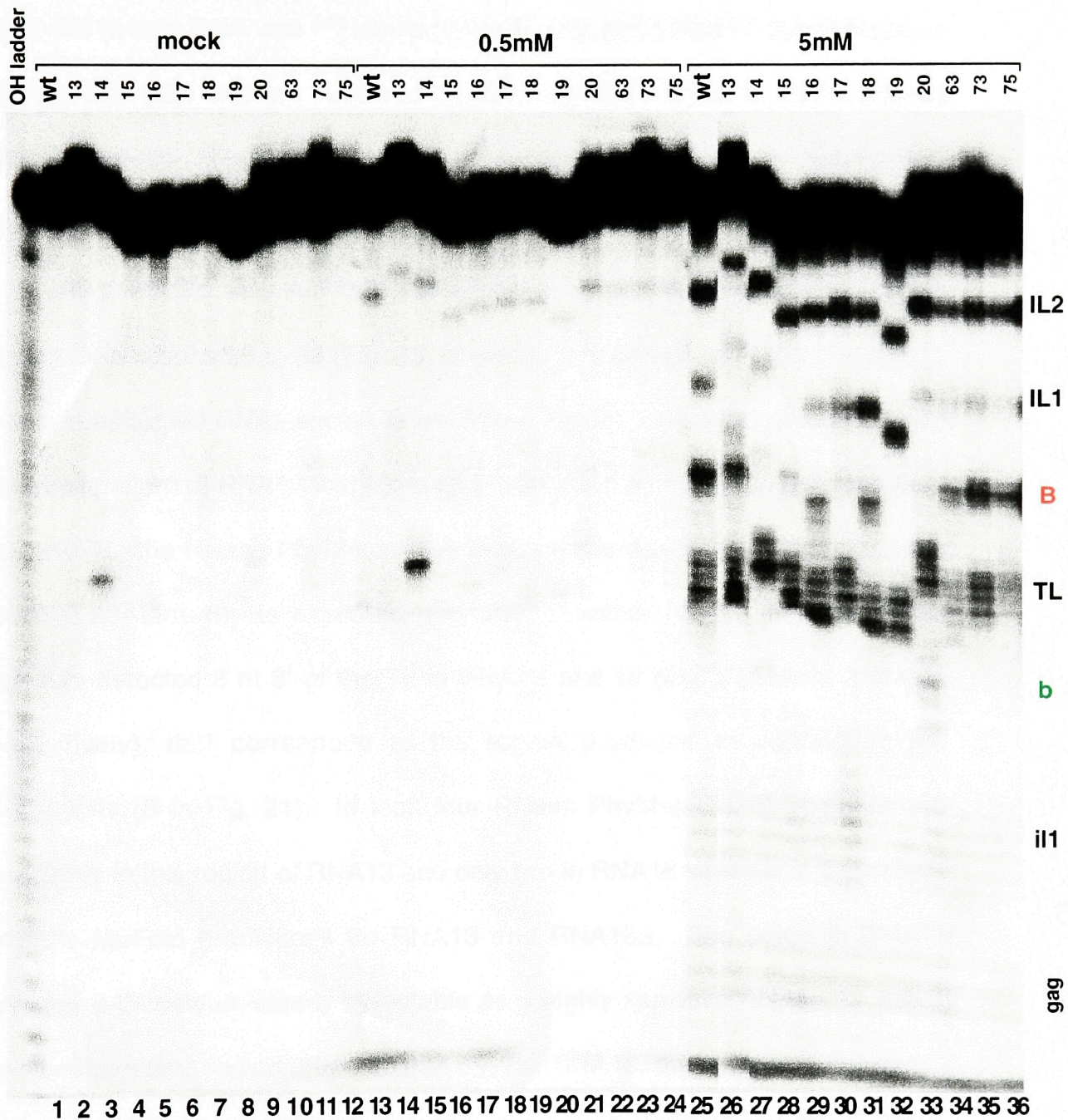


Figure 24. Lead cleavage mapping of the secondary structure in the terminal hairpin of the HDV RNA templates. 5' end labeled RNAs encompassing the terminal hairpins of the wt and mutant templates represented in Figure 21 (lanes as indicated) were subjected to mock (lanes 1-12), or Pb^{2+} catalyzed hydrolysis in the presence of 0.5mM (lanes 13-24) or 5mM lead acetate. The resulting fragments were resolved in a 10% polyacrylamide/8M urea gel and detected by autoradiography. A sequence ladder generated by alkaline hydrolysis is also included (lane M). The positions of the fragments corresponding to the RNAs cleaved at the characteristic secondary structure elements (TL, B, b, IL1 and IL2) described in Figure 21 are indicated.



For the mapping analysis, RNA molecules encompassing the terminal hairpin region (Fig. 21) were synthesized from the appropriate cDNA templates. Each ~74 nt long RNA was P^{32} labeled at its 5' end, subjected to partial digestion with RNases with different nucleotide specificities (Fig. 23) or to Pb^{2+} -catalyzed RNA hydrolysis (Fig. 24), and the resulting fragments were resolved in polyacrylamide gels.

As predicted, RNase PhyM (cleaves at A↓ and U↓) cleaves efficiently at nucleotide positions 36 to 38 (Fig. 23, lanes 2-4 and 6) that correspond to the TL region in all tested RNAs except RNA18 (see Fig. 21 for the sequence). In the digestion pattern of RNA18 that contains 2 nt deletion in the region between its 5' end and TL, the RNase PhyM-sensitive sequence is detected at positions 34 to 36 (Fig. 23, lane 5), as expected (Fig. 21). Another RNase PhyM-sensitive region is detected 8 nt 3' of the TL in RNA13 and 18 (Fig. 23, lanes 2 and 5, respectively), that correspond to the region predicted to contain bulged nucleotides (B in Fig. 21). In fact, four RNase PhyM-sensitive positions are detectable in this region of RNA13 and only two in RNA18 which is in agreement with the MuFold predictions for RNA13 and RNA18a. The bulge in RNA13 contains a G residue, clearly detectable as a highly sensitive RNase T1 site in the corresponding region (Fig. 23, lane 8). For RNA18, the program predicts a 2 nt UA bulge (18a in Fig. 21). However, there is a clear RNase T1 sensitive position in this region, similar to that observed in RNA13 (Fig. 23, lane 13 and 8, respectively), suggesting that at least a subset of RNA18 molecules can acquire

a secondary structure with the 4 nt bulge (designated 18b in Fig. 21B,). A prominent RNase T1 sensitive site is also detected in RNA20 at position 29 (Fig.23, lane 15), demonstrating the presence of a bulge 5' of the TL, as predicted (Fig. 21C). No similar RNase T1 sensitive sites are detected in any other of the tested RNAs (Fig. 23, lanes 7, 9-12, 14 and 16-18). This may be a consequence of the absence of any bulged nucleotides, as predicted for some of the RNAs (Fig. 21C). Alternatively, the sequence of the bulge may not be recognized by RNase T1 (Fig. 21B).

To detect single stranded regions independently of their primary sequence, Pb^{2+} cleavage mapping was performed (Fig. 24). At low Pb^{2+} concentrations (0.5 mM), single Pb^{2+} sensitive site is detected in all RNAs (Fig. 24, lanes 13-24) that corresponds to the region encompassing the internal loop 2, IL2 (Fig. 21). The high sensitivity of IL2 to Pb^{2+} catalyzed hydrolysis may indicate that this site represents a specific, high affinity metal binding site in the HDV RNA (Polacek & Barta, 1998; Moore 1999 and references therein). Higher Pb^{2+} concentrations (Fig. 24, lanes 25-36) cause increased levels of cleavage in the IL2, and reveal additional single stranded regions accessible to hydrolysis. In the region corresponding to TL there are 4-5 Pb^{2+} sensitive residues in all RNAs, as predicted. The electrophoretic mobility of the fragments generated by cleavage at the set of sensitive nucleotides in the TL region changes depending on the deletions or insertions in the corresponding RNA molecules. For example, in the pattern of RNAs 14 and 20 (Fig. 24, lanes 27 and 33, respectively), these

fragments migrate slower than the wt set of TL-cleaved fragments, consistent with a 2 nt insertion 5' of the TL (Fig. 21). In the pattern of RNAs 16, 18 and 19 (Fig. 24, lanes 29, 31 and 32, respectively), the set of fragments generated by the cleavage in the TL region migrates faster, consistent with the number of deleted nucleotides 5' of TL (1 nt in RNA16 and 2 nt in RNAs 18 and 19). These results clearly demonstrate that all HDV RNA templates fold into a hairpin structure with a 4 or 5 nt terminal loop, as predicted by the computer folding program (Fig. 21).

The lead cleavage mapping also demonstrates the presence of the predicted bulge, B, 3' of TL in RNAs wt, 13, 16, 18, 63, 73 and 75 (Fig. 24, lanes as indicated) and 5' of TL in RNA20 (region designated b, in Fig. 24, lane 33). Also in agreement with the predictions is the observed low sensitivity to Pb^{2+} hydrolysis in the corresponding region of RNAs 14, 15, 17 and 19 (Fig. 24, lanes as indicated). In particular, MuFold program does not predict any unpaired nucleotides for RNAs 17 and 19 (Fig. 21), and no lead sensitive sites are detected in these RNAs (Fig. 24, lanes 30 and 32, respectively). Interestingly, RNAs 14 and 15 show similarly low sensitivity to Pb^{2+} hydrolysis in this region (Fig. 24, lanes 27 and 28, respectively), although unpaired nucleotides are predicted for both these RNAs (Fig. 21). This indicates that in RNAs 14 and 15 the nucleotides predicted to be unpaired, may in fact form non Watson-Crick base pairs. Alternatively, and more likely in the case of RNA14, a symmetric loop is formed that is stabilized by cross-strand stacking interactions (Ciesiolka et

al., 1997; Moore, 1999). Similar interactions may be accounted for the observed weak Pb^{2+} cleavage in the IL1 region (Fig. 24, lanes 25-30 and 33-35) in all RNAs, for which the symmetric IL1 is predicted (Fig. 21). Consistent with this possibility is the observed increased sensitivity to lead cleavage in the IL1 region of RNAs 18 and 19 (Fig. 24, lanes 31 and 32, respectively), for which an asymmetric IL1 is predicted (Fig. 21). The cleavage pattern of RNA75 does not fully conform to this explanation, however, it clearly demonstrates the presence of all single stranded regions (Fig. 24, lane 36) as predicted (Fig. 21).

2.2 RNA SECONDARY STRUCTURE REQUIREMENTS FOR HDV REPLICATION IN VIVO

2.2.1 Description of a system to study RNA template requirements for HDV replication in vivo

The analysis of the template requirements for AG HDV RNA-templated pol II transcription in vitro demonstrates that the efficiency and specificity of the reaction depend on the secondary rather than the primary structure of the RNA template. A bulge structure (B), positioned 3' of the terminal loop (TL) is characteristic of all the RNAs that are recognized by pol II as templates for

Figure 25. Description of the strategy used to analyze the effects of RNA secondary structure alterations on HDV replication in vivo. pSVL expression vectors carrying a dimer (pSVL2) or 1.3X copy (pSVL1.3) HDV cDNA insert were transfected into Cos7 or HeLa cells. Positions of the cleavage sites of the AG (circle), G (triangle) ribozymes, the open reading frame for the delta antigen (Ag ORF) and some characteristic restriction sites in the HDV cDNA are indicated. The primary AG HDV RNA transcripts synthesized from the SV40 promoter contain two ribozyme domains that allow them to be processed into circular AG templates for G RNA synthesis. Total cellular RNA was isolated at different time points post transfection and hybridized to HDV specific, ³²P labeled RNA probe (red). The HDV-probe RNA hybrids were resolved in a 3.5% polyacrylamide/8M urea gel and visualized by autoradiography (hyb-shift assay).

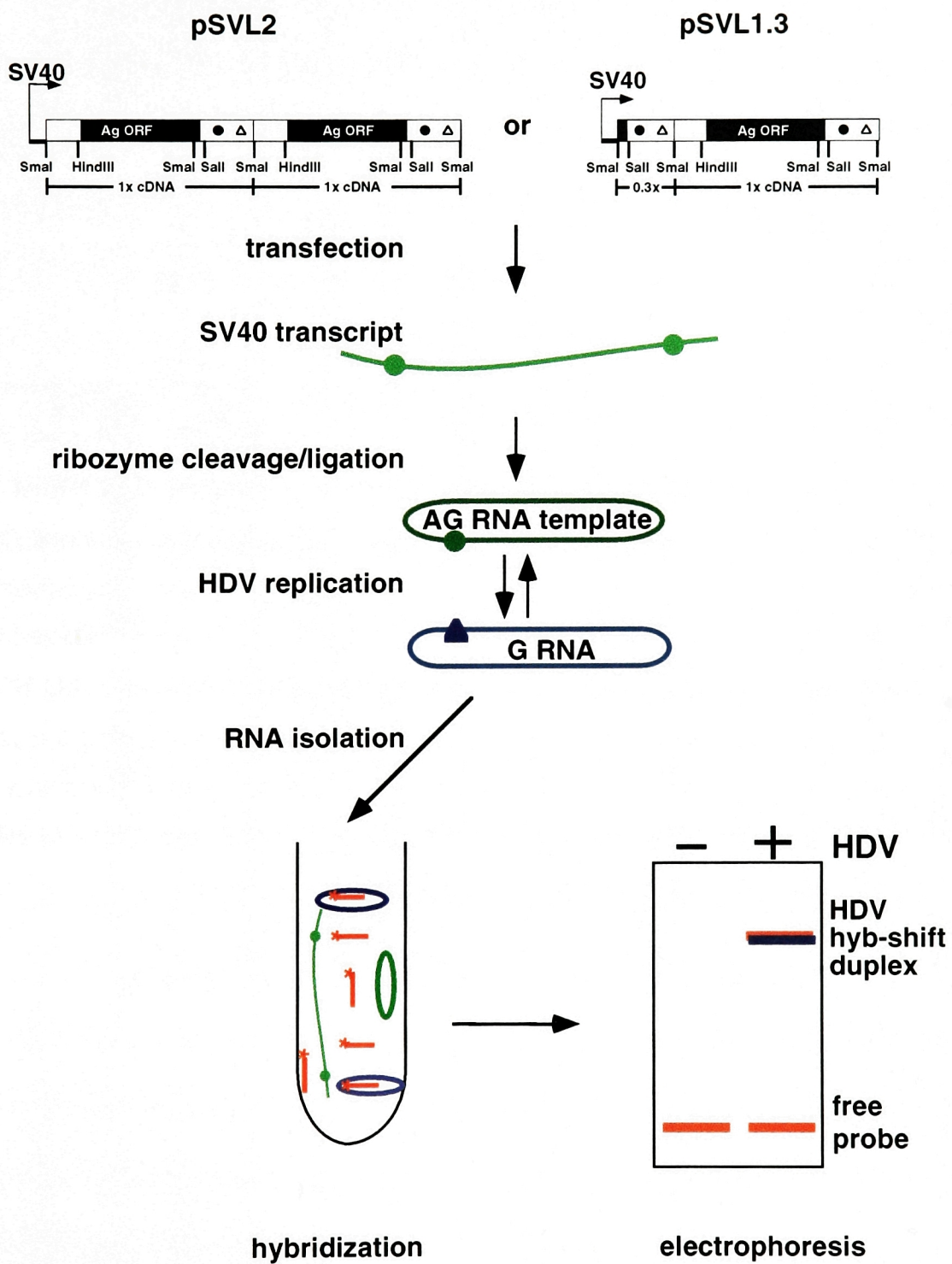


Figure 26. Specificity of the hyb-shift assay. Short ^{32}P labeled RNA probes specific for G (lanes 1-8) or AG (lanes 9-14) polarity HDV RNA were hybridized in solution to total RNA isolated from nontransfected (N) cells (lanes 1 and 9), pSVL2 transfected (T) cells (lanes 2 and 10), or synthetic genomic (lanes 3–7 and 14) or antigenomic (lanes 8 and 11-13) HDV RNAs (their length in kb is indicated). The resulting high molecular weight hybrids were resolved in a 3.5% polyacrylamide/8M urea gel and detected by autoradiography. AG HDV and G HDV indicate the positions of the hyb-shift complexes containing antigenomic or genomic polarity HDV RNA, respectively. C and L indicate the position of hyb-shift complexes of antigenomic probe with circular and linear G HDV RNA, respectively. Lane M contains λ xHindIII DNA size marker.

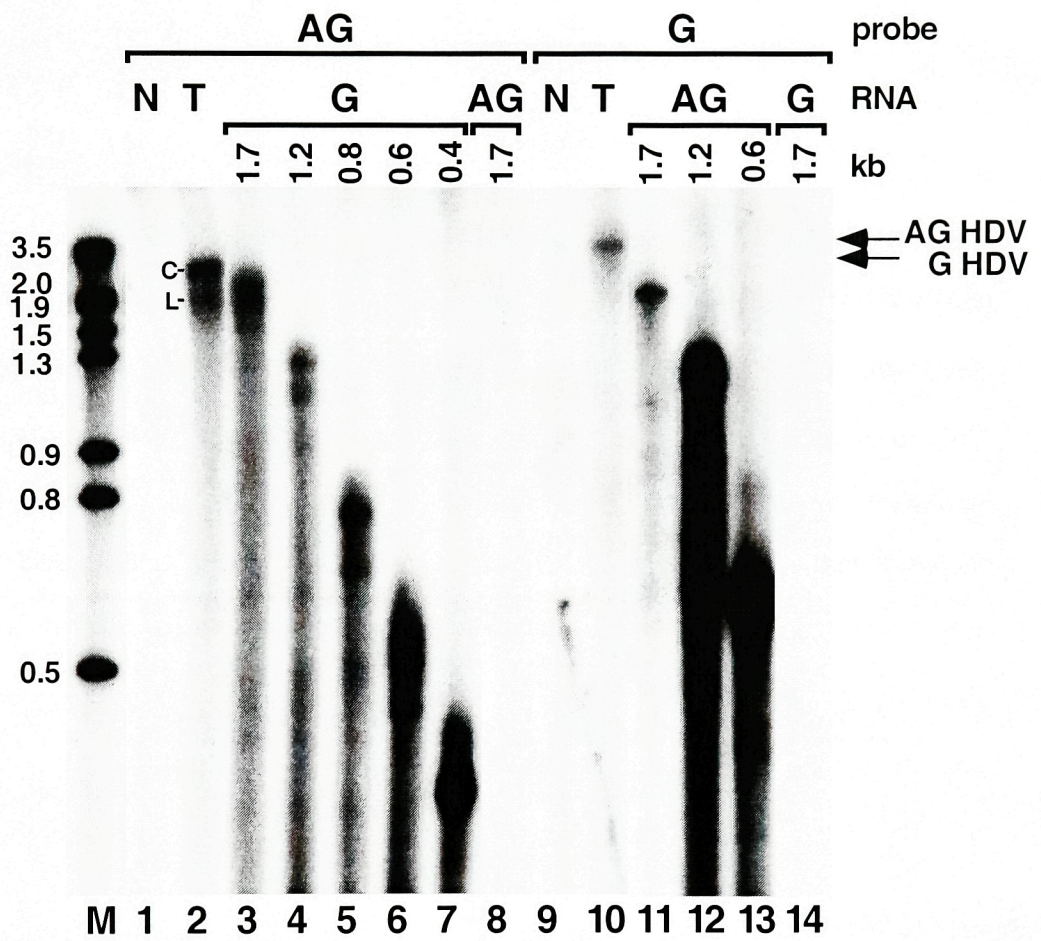
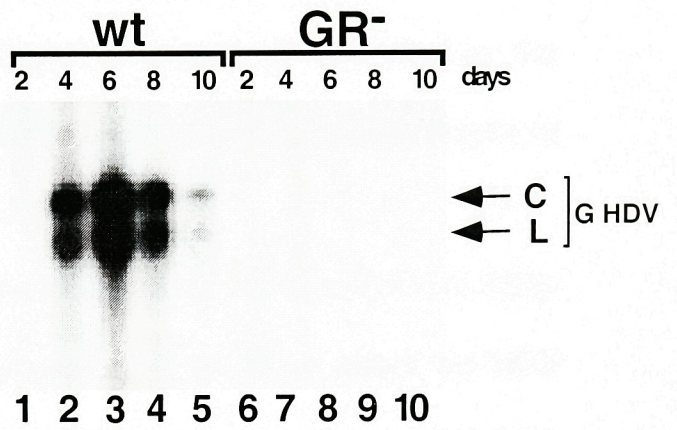
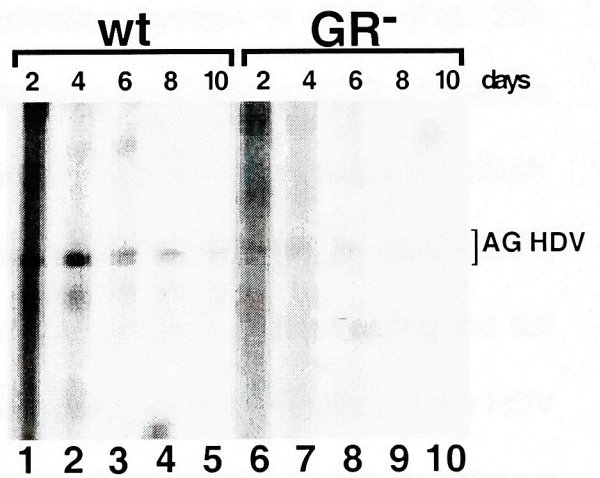


Figure 27. Replication of wild type (wt) and genomic ribozyme mutant (GR⁻) HDV in COS7 cells. Total cellular RNA was isolated 2, 4, 6, 8 and 10 days after transfection (lanes as indicated) with pSVL2-wt (lanes 1-5) or pSVL-GR⁻ (lanes 6-10) and genomic **(A)** or antigenomic **(B)** HDV RNA was detected by hyb-shift assay. C and L indicate the circular and linear genomic, GHDV RNA respectively. Mostly circular antigenomic, AG HDV RNA is detected.

A



B



transcription in vitro (Fig. 21B). Interestingly, the specific site of template cleavage/transcription initiation is positioned immediately adjacent to the bulge. The high specificity of the pol II mediated HDV RNA transcription in vitro indicates the possibility that HDV replication in vivo may be regulated by the similar secondary structure elements in the left-hand terminal hairpin region of AG HDV RNA.

To examine this possibility, mutations that diversely affect AG HDV RNA-templated pol II transcription in vitro were tested for their effect on HDV replication in vivo. As described in the section the Introduction, HDV RNA replication can be observed in various cell lines transfected with eukaryotic expression vectors carrying HDV cDNA. At different time points post transfection, the level of HDV RNA synthesized in the RNA dependent fashion is monitored to assess the efficiency of HDV replication. In this study, a modified version of the HDV cDNA based transfection system is used (Fig. 25). Specifically, COS7 or HeLa cells are transfected with pSVL2 or pSVL1.3 vectors (as indicated in the specific experiment) carrying wild type or mutant HDV cDNA. The pSVL2 and pSVL1.3 constructs contain dimer or 1.3X length HDV cDNA, respectively, placed under the control of SV40 promoter. Duplicating the full length or the third (containing both, the G and AG ribozyme domains) of the HDV cDNA in the pSVL2 and pSVL1.3, respectively, is necessary to assure efficient HDV replication in the HDV cDNA based system. Namely, following transfection in the cells, AG HDV RNA dimeric or 1.3X length, linear transcripts are

synthesized in a DNA-templated reaction from the SV40 promoter. These primary transcripts generated from the both types of expression vectors contain two AG ribozyme elements that allow for RNA cleavage and ligation into a circular monomers which represent templates for subsequent HDV replication through the RNA-templated RNA synthesis. The levels of G HDV RNA synthesized by copying of the AG RNA template in vivo is monitored by a Hyb-shift assay (Fig. 25 and 26) that was developed as an alternative to the Northern hybridization and RNase-protection assays. The Hyb-shift assay involves detection of probe-HDV RNA hybrid duplexes generated by hybridization under conditions typically used in the RNase-protection assay. These conditions allow for specific and efficient hybridization in solution, of the ^{32}P labeled RNA probe to the palindromic and highly structured HDV RNA, which was difficult to accomplish using the Northern assay. The probe-HDV RNA hybrids (hyb-shift duplexes) are separated from the excess of unhybridized probe by electrophoresis and visualized by autoradiography. The Hyb-shift assay allows for detection of full length HDV RNAs. Therefore, it is better suited to monitor the HDV RNA replication than the RNase-protection assay which detects any short G HDV RNA segment generated from endogenous promoters independently of the RNA replication (Macnaughton et al., 1993a, and data not shown).

The specificity of the Hyb-shift assay is demonstrated in Figure 26. Antigenomic (AG) and genomic (G) polarity RNA probes (see Materials and Methods for sequences) are hybridized to SP6 or T7 RNA polymerase-

synthesized HDV RNAs. When the AG probe is used, only genomic, G, and not AG HDV RNA is detected (Fig. 26, lane 3 and 8, respectively). In addition, electrophoretic mobility of the AG probe-G HDV RNA hybrid changes as expected when different length G HDV RNAs are tested (Fig. 26, lanes 3-7). Using the same AG probe and RNA isolated from HDV cDNA transfected cells, two distinct hyb-shift complexes are detected (Fig. 26, lane 2). One of them, L, comigrates with the control ~1.7 kb linear G HDV RNA and the other, C migrates slightly slower (compare lanes 2 and 3), consistent with the closed circular conformation of the replicating G HDV RNA.

Different length AG HDV RNAs are detected using G polarity probe (Fig. 26, lane 11-13) that does not generate hyb-shift complex with G HDV RNA (Fig. 26, lane 14). The same probe generates a hyb-shift complex with the circular AG HDV RNA from transfected cells (Fig. 26, lane 10), that similarly to the circular G HDV RNA migrates slower than the corresponding linear AG RNA control (compare lane 10 to 11).

2.2.2 In vivo replication of HDV mutants containing RNA sequence/structure alterations that modulate pol II transcription in vitro

Replication of wild type HDV RNA in the COS7 cells

A typical time course of wild type HDV replication in COS7 cells transfected with vectors carrying dimer HDV cDNA (pSVL2) is presented in Figure 27A, lanes 1-5. Two days after transfection, only AG HDV RNA synthesized from the SV40

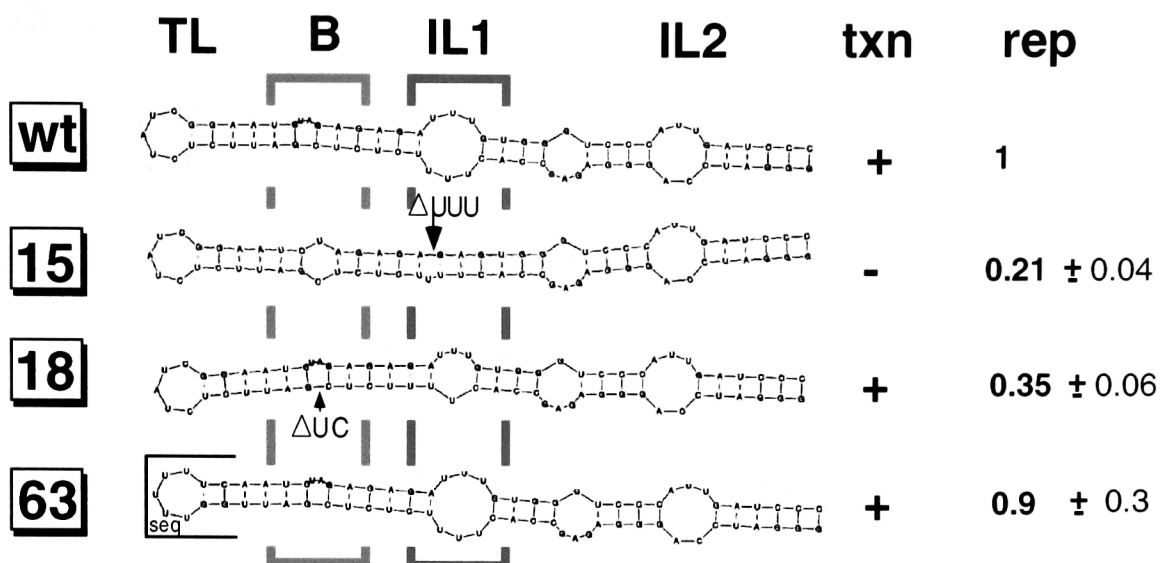
promoter (Fig. 27B, lane 1) but not G HDV RNA (Fig. 27A, lane 1) can be detected. Genomic HDV RNA synthesized from the AG RNA template accumulates at later time points (Fig. 27A, lanes 2-5), reaching a peak 6 days post transfection (lane 3). The decline of G HDV RNA at late time points (Fig. 27A, lanes 4 and 5) is caused by the inhibitory effect of the long form of delta antigen, L-HDAg, that is synthesized during the late phase of HDV replication (Glenn and White, 1991). In addition, lower G HDV RNA levels are expected at later time points post transfection because of the “dilution” effect of the growing population of untransfected COS7 cells. The HDV-GR⁻ mutant that has an impaired genomic ribozyme activity and thus can not replicate (Macnaughton et al., 1993), was also included in the analysis as a negative control (Figure 27A, lanes 6-10). At early time points, both, wt and GR⁻ antigenomic, AG, RNA is synthesized from the SV40 promoter (Fig. 27B, lanes 1-3 and 6-8, respectively, and data not shown). In contrast, G HDV RNA, synthesized by the RNA-templated transcription can be detected only in the cells transfected with the pSVL2-wt and not with the pSVL2-GR⁻ vector (Fig. 27A).

Replication of the HDV15, HDV18 and HDV63 mutants in COS7 cells

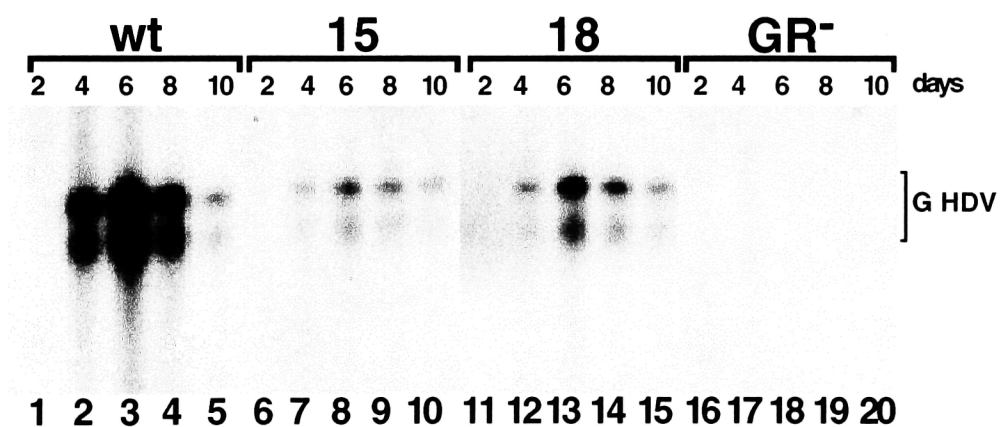
HDV15 mutant replicates ~5 fold less efficiently than the wt HDV in COS7 cells (Fig. 28B, compare lanes 6-10 to 1-5). The three nucleotide deletion in the terminal hairpin of the HDV15 RNA interferes with the formation of the

Figure 28. Replication of HDV15, HDV18, and HDV63 in COS7 cells. A. RNA sequences and the secondary structures in the terminal hairpins of wild type (wt) and mutant HDV RNAs analyzed in B and C. TL-terminal loop; B-bulge 3' of the TL; IL1-internal loop 1; IL2-internal loop 2. Functional and nonfunctional pol II templates in vitro are designated + and – respectively. Deletions in the sequence are represented by Δ and, insertions by $::$. seq indicates that the region contains only sequence alterations without deletions or insertions. The efficiency of replication of the mutant HDV RNAs is represented relative to the replication of wild type HDV RNA that is designated 1 (see Materials and Methods for description of the quantitative analysis). **B. and C.** Total cellular RNA was isolated 2, 4, 6, 8 and 10 days after transfection (lanes as indicated) and genomic, G HDV RNA was detected by hyb-shift assay. **B.** RNA isolated from cells transfected with pSVL2-wt (lanes 1-5), pSVL2-15 (lanes 6-10), pSVL2-18 (lanes 11-15) or pSVL-GR⁻ (lanes 16-20). **C.** RNA isolated from cells transfected with pSVL2-wt (lanes 1-4), with pSVL2-63 (lanes 5-8).

A



B



C

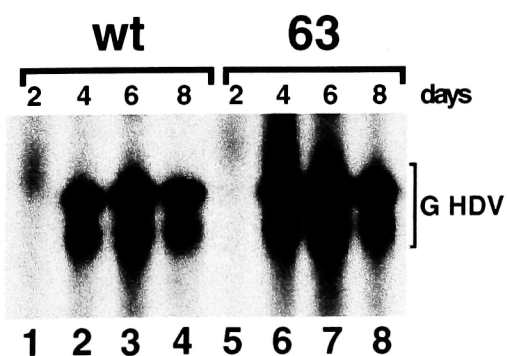
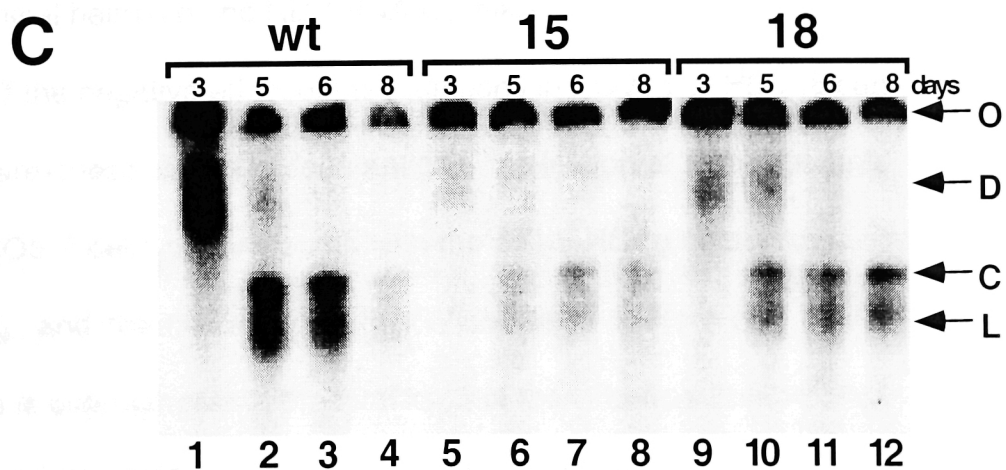
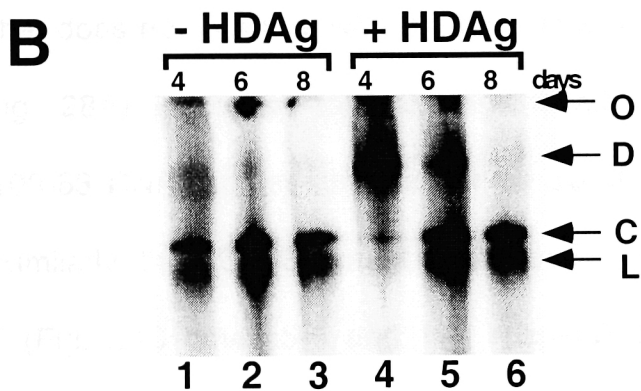
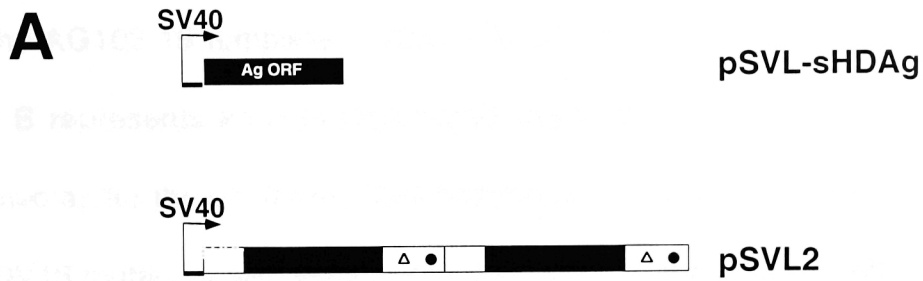


Figure 29. Replication of the HDV15 and HDV18 in COS7 cells when delta antigen is overexpressed. **A.** Schematic representations of the expression vectors used. The pSVL-sHDAg vector contains the open reading frame encoding for the delta antigen under the control of the SV40 promoter. The pSVL2-based vectors are the same as those used in the experiment represented in Figure 28. Total cellular RNA was isolated after transfection (at days as indicated) and genomic, G HDV RNA was detected by hyb-shift assay. **B.** RNA isolated from cells transfected with pSVL-wt only (lanes 1-3) or pSVL-wt cotransfected with the pSVL-sHDAg (lanes 4-6). **C.** RNA is from cells cotransfected with pSVL-sHDAg and pSVL-wt (lanes 1-4) or pSVL2-15 (lanes 5-8) or pSVL2-18 (lanes 9-12). O-gel origin; D-dimer HDV RNA; C-circular monomer HDV RNA; L-linear monomer HDV RNA.



characteristic bulge, B (Fig. 28A) and completely abolishes pol II-mediated transcription from the AG103-15 template in vitro (Fig. 22, lane 5). This result suggests the bulge B represents an important regulatory element for the HDV RNA replication in vivo as for the pol II-mediated transcription in vitro. However, replication of the HDV18 mutant is also significantly less efficient than that of wild type (Fig. 28B, compare lanes 11-15 to 1-5), although the deletion does not interfere with either bulge, B formation (Fig. 28A) or with pol II transcription in vitro (Fig. 22, lane as indicated). To examine the effect of a different type of sequence change that does not interfere with the formation of the bulge B, the HDV63 mutant (Fig. 28A) was also tested for replication in vivo. The corresponding AG103-63 RNA represents a functional pol II template in vitro (Fig. 11A, lane1). Similarly, the HDV63 mutant replicates in vivo comparably to the wild type HDV (Fig. 28C lanes 5-8 and 1-4, respectively). This result demonstrates that similarly to the transcription in vitro, replication in vivo is not affected by extensive sequence changes that do not alter the secondary structure in the terminal hairpin of the HDV RNA template.

To examine if the negative effect of the mutations in HDV15 or HDV18 can be alleviated by overexpression of the delta antigen, replication of these mutants was examined in COS 7 cells cotransfected with the pSVL-HDAg vector (carries the ORF for HDAg) and the corresponding pSVL2-based vectors (Fig. 29A). When delta antigen is overexpressed, the total level of the wild type G HDV RNA is increased ~1.4 fold (Fig. 29B, compare lanes 4-6 to 1-3) most likely because

delta antigen increases the stability of HDV RNA (Lazinski and Taylor, 1994). Under these conditions also, the HDV15 and HDV18 mutants replicate less efficiently than the wt HDV (Fig. 29C, lanes 5-8 and 9-12, respectively). Thus, HDAg does not alleviate the effects of these RNA secondary structure alterations on replication in vivo.

The correlation between the effects of the secondary structure alterations in HDV15 and HDV63 RNA on pol II-mediated transcription in vitro, in HeLa cell NE and on HDV RNA replication in vivo, in COS7 cells, strongly suggests that similar structure elements are involved in the regulation of both processes. However, the opposite effect of the 2 nt deletion in HDV18 on replication in the two systems, rises a possibility that availability of cell specific protein factors may contribute to the observed difference. To address this possibility replication of these mutants was also examined in HeLa cells.

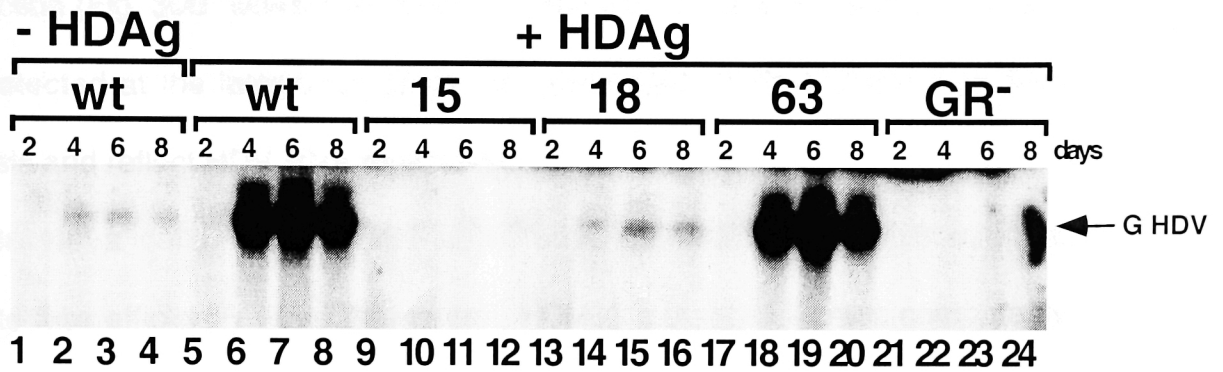
Replication of wild type, and HDV15, HDV18, and HDV63 mutants in HeLa cells

HDV replication in HeLa cells transfected with the pSVL-based vectors is very inefficient (Fig. 30A, lanes 1-4). This is most likely due to the lower levels of AG RNA transcript generated from the SV40 promoter in this cell type as compared to COS7 cells (note the absence of AG HDV early post transfection in Figure 30B, lanes 1 and 5). Consistent with this, HDV RNA replication is stimulated (Fig. 30A, lanes 5-8) when delta antigen that increases the stability of HDV RNA

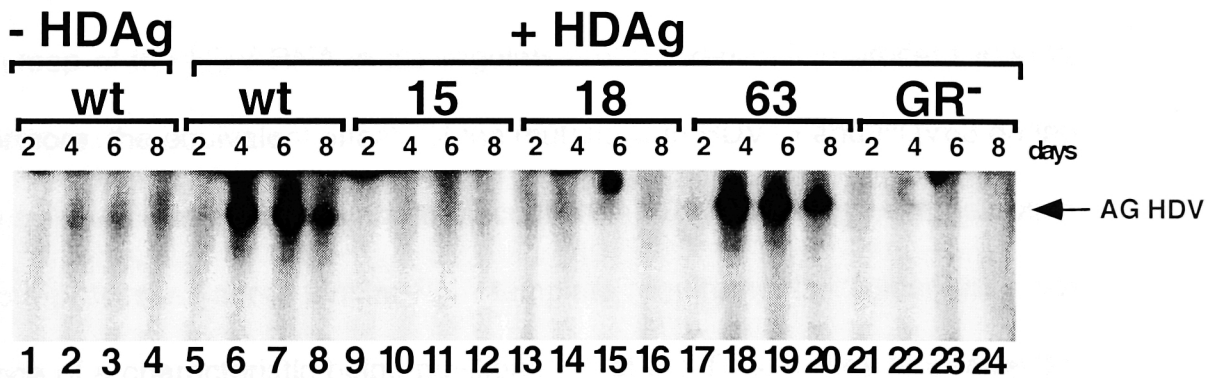
Figure 30. Replication of the HDV15, HDV18, and HDV63 in HeLa cells.

Total cellular RNA was isolated 2, 4, 6 and, 8 days after transfection (lanes as indicated) and genomic, G HDV RNA (**A**) or antigenomic, AG HDV RNA (**B**) was detected by hyb-shift assay. RNA isolated from cells transfected with pSVL2-wt (lanes 1-4) or from cells cotransfected with the pSVL-sHDAg (lanes 5-24) and pSVL2-wt (lanes 5-8) or pSVL2-15 (lanes 9-12) or pSVL2-18 (lanes 13-16) or pSVL2-63 (lanes 17-20) or pSVL-GR⁻ (lanes 21-24).

A



B



(Lazinski and Taylor, 1994) and stimulates replication in systems where its synthesis is otherwise limited (Kuo et al., 1989), is provided in trans, from the pSVL-HDAg vector. Under these conditions the primary transcripts synthesized from the SV40 promoter in HeLa cells, can not be detected early post transfection (Fig. 30B, lanes 1, 5, 9, 13, 17, and 21). Thus both, G and AG HDV RNA detected at the later time points, are generated by RNA-dependent RNA synthesis and reflect HDV RNA replication.

In HeLa cells, as in COS7 cells, both, HDV15 and HDV18 mutants replicate less efficiently while the neutral, HDV63 mutant replicates comparably to the wild type HDV (Fig. 30A and B, compare lanes 9-12, 13-16, and 17-20 respectively, to lanes 5-8). The equivalent effect of the various sequence/structure alterations on HDV replication in the two different cell types further supports the importance of the specific secondary structure in the terminal hairpin loop of the HDV RNA in the regulation of the replication process in vivo. Furthermore, the equivalent effects of the mutations in HDV15 and HDV63 on the efficiency of replication in vivo and transcription in vitro further support the notion that both processes share similar RNA template requirements. Specifically, the presence of a characteristic bulge positioned 3' of the TL specifies functional pol II RNA templates in vitro and is also required for efficient HDV replication in vivo. However, the inefficient replication of HDV18 that has a similarly positioned bulge, B and represents a functional RNA template for pol II in vitro suggests that the regulation of the replication process in vivo is more complex and involves

additional RNA secondary structure elements and/or other factors. Significantly, the mutations in both, HDV18 and HDV15 interfere not only with the formation of the bulge, B, but also with the formation of the symmetric internal loop 1, IL1, characteristic for the wild type and HDV63 RNA. Taken together, these results suggest that the internal loop 1, IL1, in addition to the bulge, B may play a role in the regulation of HDV replication.

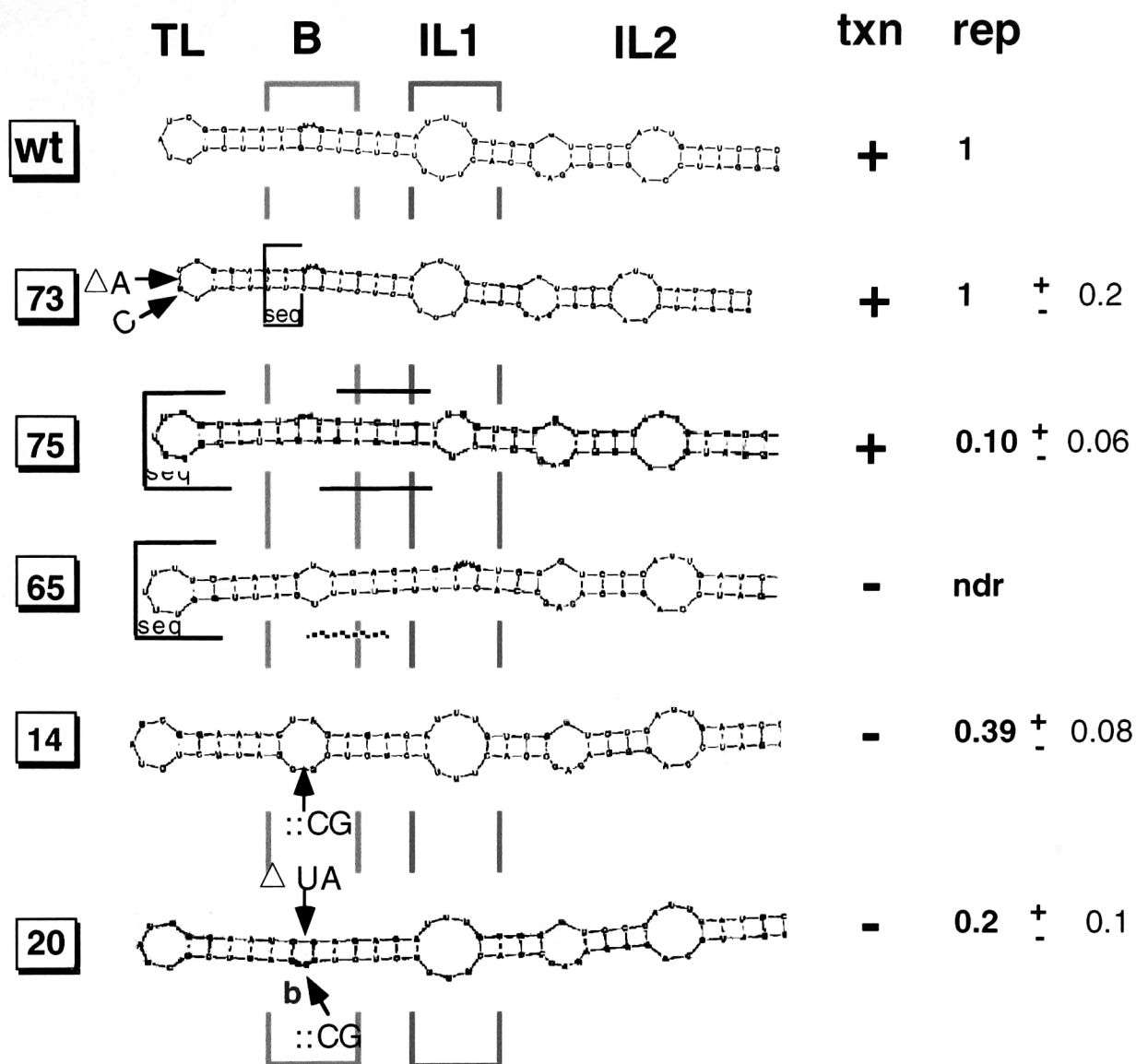
Replication of HDV73, HDV75, HDV65, HDV14, and HDV20 mutants in COS7 cells

To further examine the possible involvement of IL1, and more definitively establish the importance of bulge B for HDV replication in vivo, five additional mutants were tested. The mutations and the secondary structure in the left-hand terminal hairpin of these HDV RNAs are represented in Figure 31A. The effects of the structure alterations on pol II-mediated transcription in vitro are demonstrated in Figure 22 and summarized in Figure 31A (designated 'txn'). Effects of the corresponding mutations on HDV replication in vivo was examined using pSVL1.3 based vectors (see Fig. 25A) in both COS7 and HeLa cells.

Replication of these HDV mutants in COS7 cells is demonstrated in Figure 31B. Similarly to the "neutral" HDV63 mutant, HDV73 replicates as efficiently as the wild type (Fig. 31, compare lanes 4-6 to 1-3). As demonstrated above, the sequence changes in HDV73, as those in HDV63, do not interfere with formation of the bulge B or with pol II transcription in vitro (Fig. 22, lane 3).

Figure 31. Replication of HDV73, HDV75, HDV65, HDV15, and HDV20 in COS7 cells. **A.** RNA sequences and the secondary structures in the terminal hairpins of wild type (wt) and mutant HDV RNAs analyzed in B. TL-terminal loop; B-bulge 3' of the TL; IL1-internal loop 1; IL2-internal loop 2. Functional and nonfunctional pol II templates in vitro are designated + and – respectively. Deletions in the sequence are represented by Δ and, insertions by ::. seq indicates that the region contains only sequence alterations without deletions or insertions. The efficiency of replication of the mutant HDV RNAs is represented relative to the replication of wild type HDV RNA that is designated 1 (see Materials and Methods for description of the quantitative analysis). ndr- designates that no G HDV RNA could be detected **B.** Total cellular RNA was isolated 4, 6 and, 8 days after transfection (lanes as indicated) and genomic, G HDV RNA was detected by hyb-shift assay. RNA isolated from cells transfected with pSVL2-wt (lanes 1-3), pSVL2-73 (lanes 4-6), pSVL2-75 (lanes 7-9), pSVL2-65 (lanes 10-12), pSVL2-14 (lanes 13-15) or pSVL-20 (lanes 16-18).

A



B

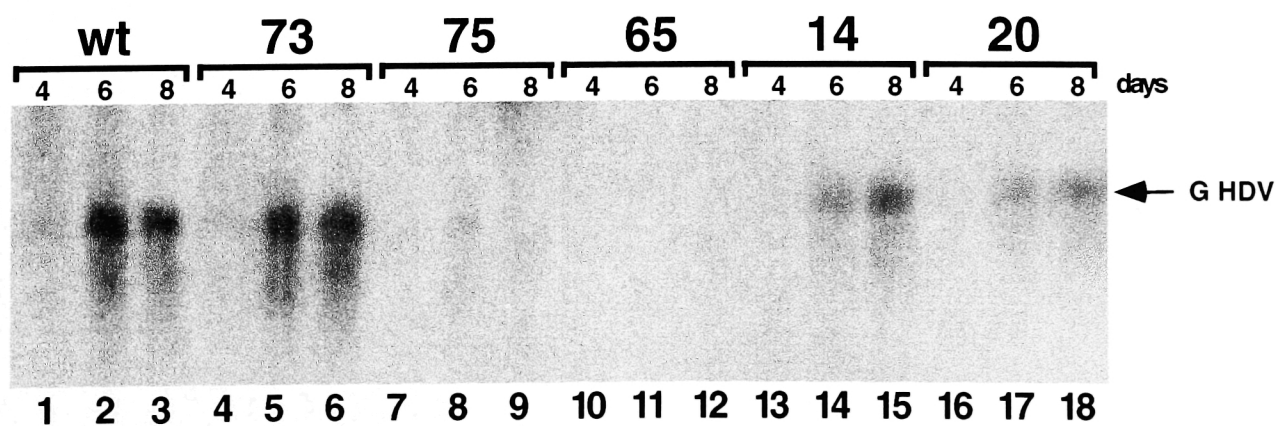
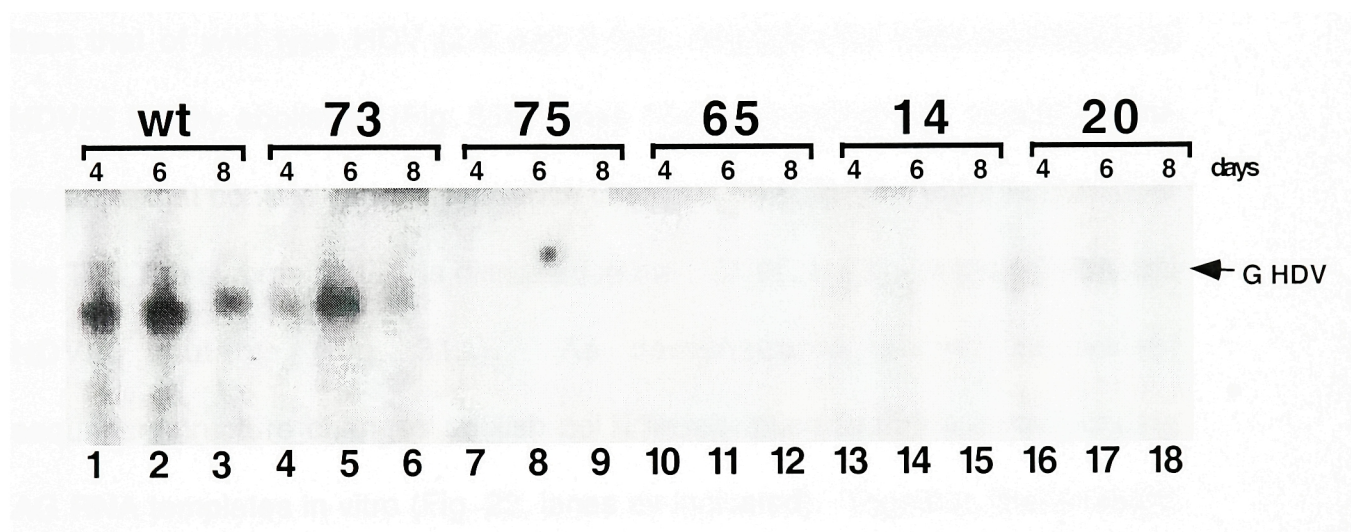


Figure 32. Replication of HDV73, HDV75, HDV65, HDV15, and HDV20 in HeLa cells. Total cellular RNA was isolated 4, 6 and, 8 days after transfection (lanes as indicated) and genomic, G HDV RNA was detected by hyb-shift assay. RNA isolated from cells cotransfected with the pSVL-sHDAg and pSVL2-wt (lanes 1-3) or pSVL2-73 (lanes 4-6) or pSVL2-75 (lanes 7-9) or pSVL2-65 (lanes 10-12) or pSVL2-14 (lanes 13-15) or pSVL-20 (lanes 16-18).



The symmetric IL1 can also be formed in HDV73 RNA, generating a terminal hairpin that fully resembles the TL/B/sym.IL1 structure characteristic for the wild type (Fig. 31A) and the “neutral” HDV63 mutant (Fig.28A). This result again demonstrates that sequence changes that do not interfere with formation of the characteristic secondary structure in the terminal hairpin of HDV RNA do not interfere with the replication in vivo.

Replication of HDV14 and HDV20 mutants is significantly less efficient than that of wild type HDV (2.5 and 5 fold, respectively) while replication of HDV65 is fully abolished (Fig. 31B, lanes 10-18, as indicated). In each of the mutants, that contain various sequence changes, no bulge B can be formed 3' of the TL. The symmetric IL1, is disrupted in the HDV65, but not in the HDV14 and HDV20 mutants (Fig. 31A). As demonstrated above, equivalent sequence/structure changes abolish pol II transcription from the corresponding AG RNA templates in vitro (Fig. 22, lanes as indicated). Together, these results further demonstrate that the bulge B is an important RNA secondary structure element for the regulation of both HDV replication in vivo and pol II transcription in vitro. However, similarly to the HDV18 mutant (Fig. 28), the HDV75 does not replicate in COS7 cells (Fig. 31B, lanes 7-9), although the sequence changes in this mutant do not interfere with the formation of the bulge, B, or with pol II transcription in vitro (Figures 31A and 22, respectively). Notably, different sequence changes in the HDV75 and HDV18 mutants generate a similar secondary structure alteration in the corresponding RNAs. Specifically, the

symmetric IL1 can not be formed in the terminal hairpin of the both mutants. While the absence of symmetric IL1 does not interfere with the efficiency of transcription in vitro it does negatively affect replication of the two mutants in vivo.

Replication of HDV73, HDV75, HDV65, HDV14, and HDV20 mutants in HeLa cells

In HeLa, as in COS7 cells, the HDV73 mutant that contains the characteristic TL/B/sym.IL1 hairpin structure, replicates comparably to the wt HDV (Fig. 32, lanes 4-6). Furthermore, the various sequence alterations that invariably interfere with formation of bulge B, in the terminal hairpin of RNAs 65, 14, and 20, and abolish pol II transcription in vitro, inhibit HDV replication in HeLa cells (Fig. 32, lanes 1-18, as indicated). Also similarly to the situation in COS7 cells, the HDV75 mutant does not replicate in HeLa cells (Fig. 32, lanes 7-9). As mentioned above, HDV75 and HDV18 mutants have significantly different sequence but very similar secondary structures in their terminal hairpins. Specifically, they both have a characteristic asymmetric instead of the symmetric IL1, characteristic for wild type HDV RNA. Thus, it appears that the symmetric IL1, in addition to bulge B, represents an important regulatory element for the HDV replication in vivo.

Taken together, these results demonstrate that the secondary structure of the left-hand terminal hairpin of the HDV RNA, characterized by the bulge B

positioned 3' of the TL, determines efficient pol II-mediated HDV RNA transcription in vitro and HDV RNA replication in vivo. This correlation supports the pol II involvement in HDV replication. Furthermore, it suggests that HDV RNA replication in vivo may initiate by a mechanism that involves RNA template cleavage at a specific site, as observed in vitro. The different effects of the secondary structure changes in HDV18 and HDV75 RNAs observed in the two experimental systems indicate that in vivo initiation and/or other steps of the RNA synthesis may be regulated by additional RNA secondary structure elements including the symmetric IL1. A model for HDV RNA replication that accounts for the similarities and differences observed in vitro and in vivo is presented in section 3.5.

Chapter 3

DISCUSSION

3.1 Does HDV RNA represent a specific template for pol II-mediated transcription?

A number of studies have indicated that pol II may be involved in replication of a group of RNA pathogens whose life cycle does not include a DNA intermediate. These RNA pathogens are represented by the group B plant viroids and the Hepatitis Delta Virus (HDV). Plant pol II has been found to associate with viroid RNA in transcriptionally active nuclear homogenates (Warrilow and Symons, 1999), and viroid RNA associated to pol II has been observed by electron microscopy (Goodman et al., 1984). Functional studies that implicate pol II as an RNA-dependent polymerase include mostly indirect evidence, based on the inhibition of viroid and HDV RNA synthesis in vivo and in vitro in the presence of the toxin α -amanitin (Mühlbach and Sängner, 1979; Rackwitz et al., 1981; MacNaughton et al., 1991; Fu and Taylor, 1993). Thus, direct evidence concerning the ability of pol II to specifically transcribe viroid or HDV RNA templates is still lacking.

The HDV RNA-based transcription system used in this study demonstrates more directly that pol II can specifically recognize HDV RNA as a template for precise transcription. In this system, a defined short segment (~200 nt) encompassing the left-hand terminal tip of the AG HDV rod-like RNA represents a template for α -amanitin sensitive RNA synthesis in HeLa cell nuclear extracts (NE) (Fig. 2 and 4). Furthermore, the reaction is partially resistant to the toxin in NE from PMG cells that contain an allele of the α -

amanitin resistant subunit of pol II (Fig. 2). These results strongly support the involvement of pol II in the observed HDV RNA-templated transcription reaction. Furthermore, purified RNA polymerase II supplemented with the transcriptional coactivator PC4 can recognize the same segment of AG HDV RNA and use it as a template for transcription (Gottardo 1998,). Together, these results establish the involvement of RNA polymerase II in the HDV RNA-templated RNA synthesis and provide a simplified system for studying the RNA-dependent pol II function.

Considering the rod-like structure of some viroids and HDV, it has been proposed that pol II recognizes these RNAs as templates for transcription because their double stranded segments resemble the classical DNA template (Rackwitz et al., 1981). This notion implies that many highly structured RNAs, and certainly many regions of the viroid and HDV double-stranded segments would represent templates for pol II transcription. However, no pol II-mediated transcription can be observed using a variety of highly structured RNA molecules as snRNAs, tRNAs and 7SL RNA, or a number of HDV RNA segments (A. Peterhans and M. M. Konarska, unpublished results). The fact that only one specific region of the AG HDV RNA was shown to serve as template for pol II transcription in vitro (Fig. 2 to 4) indicates that the enzyme may recognize a specific region in the HDV RNA to initiate complementary RNA synthesis and replication in vivo. The high specificity and precision of the HDV RNA-templated pol II transcription in vitro and the similar RNA structure requirements for HDV

RNA synthesis in vivo (discussed in the following sections) further support this possibility.

The inability to detect G HDV RNA-directed RNA synthesis in the in vitro system described in this study may reflect a lesser efficiency of transcription from the genomic template. In this context it is interesting to note that estimates of the amounts of the different HDV RNA species in infected cells account for 10 to 1 ratio of the genomic and antigenomic HDV RNA (Chen et al., 1986). The difference in the infected cells may be due to a differential stability of the G and AG HDV RNAs, but also due to a different efficiency of their synthesis regulated by distinct sequence/structure determinants in the AG and G RNA templates, as suggested above. In fact, others have observed HDV RNA-templated RNA synthesis in vitro using HeLa cell NE and a segment of genomic polarity RNA encompassing the right-hand tip of the HDV rod (opposite of the one analyzed in this study) (Beard et al., 1996). Detection of the AG RNA product in that system implied the presence of a G HDV RNA promoter at the right-hand terminal hairpin of the HDV RNA. However, no α -amanitin sensitivity of that reaction was demonstrated and thus pol II involvement and other details concerning the specificity and mode of initiation remain unclear. Assuming that the reaction in the system of Beard and colleagues is carried out by pol II, it is possible that the right and left-hand tip of the HDV RNA rod can initiate synthesis of AG and G HDV RNA, respectively. In such case, further increasing the sensitivity of our transcription system in vitro by testing alternative transcription, template

preparation, and/or product detection conditions, may result in detection of the less efficient AG HDV RNA-directed transcription. Alternatively, G RNA-templated transcription can not be detected in our in vitro system, because distinct RNA polymerases may direct the synthesis of genomic and antigenomic HDV RNA. In fact, two distinct RNA polymerases, α -amanitin resistant and α -amanitin sensitive polymerase were implicated in the synthesis of full length, AG HDV RNA and delta antigen mRNA, respectively, by recent studies of inhibition of HDV RNA synthesis by this toxin in vivo (Modahl and Lai, 2000). If indeed, α -amanitin insensitive polymerase is involved in the synthesis of the full length AG HDV RNA, as suggested by the authors, the G HDV promoter directing its synthesis could not be detected in our transcription system, which depends on pol II function.

3.2 Initiation of pol II-mediated transcription on the HDV RNA-template in vitro

In the 'typical' DNA templated transcription, selection of the pol II start site is highly regulated and involves interaction of specific accessory protein factors (GTFs) with the promoter sequences positioned internally on the DNA template. Pol II transcription from such internal promoter sequences initiates de novo i.e. by extending the new transcript on the 3'OH of the initiating NTP, generating a transcript with a free 5' end accessible for modifications by the RNA capping enzymes (Bentley, 1999). As demonstrated by the deletion analysis presented in

Figure 3, the HDV RNA-templated transcription in vitro also initiates at the position internal to the RNA template. This result indicated that pol II recognizes a specific region near the terminal loop of the HDV RNA template and may initiate transcription de novo, by a process similar to that observed on the DNA templates. However, the detailed analysis of the NE product (Fig. 7 and 8-13) revealed that it represents a chimeric molecule, in which the G RNA transcript is covalently linked to the 5' half of the AG RNA template, suggesting that pol II initiation in this system involves cleavage of the RNA template. The new end of the AG RNA template generated by the cleavage base pairs to the template strand and thus it is appropriately positioned to be used as a primer for transcription (see Fig. 10 and 17). The site of cleavage and initiation of transcription is unique in all of the functional RNA templates (e.g. Fig. 14), suggesting that the RNA-templated reaction, like the 'typical' DNA-templated pol II initiation, represents a precisely regulated process. Furthermore, transcription proceeds by accurate copying of the sequence of the HDV RNA template (Fig. 5-6 and 10-15), as expected for a reaction that reflects a stage in the RNA replication process.

The RNA template cleavage in the nuclear extract is most likely accomplished by the polymerase itself since purified pol II supplemented only with PC4 that has not been shown to possess a detectable endonuclease activity, generates a chimeric transcription product with an identical template/transcript junction (Gottardo, 1998). Furthermore, a number of earlier

studies concerning pol II activity, point out that this enzyme possesses an intrinsic endoribonuclease activity that is under certain conditions, activated by the elongation factor TFIIS (Izban and Luse, 1992; Reines, 1992; Izban and Luse, 1993; Gu and Reines, 1995). In fact, endonucleolytic cleavage of the nascent transcripts has been demonstrated in vitro in paused or arrested ternary complexes with all three eukaryotic RNA polymerases, as well as with *E. coli* and vaccinia RNA polymerases (see Reines, 1994; Uptain et al., 1997; and references therein). In the arrested complexes the nascent RNA chain is repositioned relative to the polymerase subunits and an internal phosphodiester bond of the nascent transcript is positioned at the active site. In the case of pol II arrested complexes, the RNA cleavage is stimulated by TFIIS that interacts with the largest subunit of pol II (Archambault et al., 1992; Wu et al., 1996) and has been found to crosslink to the 3'OH end of the cleaved transcript (Powell et al., 1996). The observation that pyrophosphate can functionally substitute TFIIS suggests that the endonucleolytic cleavage represents a reverse polymerization reaction accomplished by the pol II catalytic center (Rudd et al., 1994).

In light of this model for the regulation of pol II elongation, it is plausible that the endonucleolytic activity of pol II can be activated upon its binding to the HDV RNA template, generating the new 3'OH that is used as a primer for transcription. In such case the initiation of the HDV RNA-templated transcription observed in our in vitro system may mechanistically reflect the elongation phase of the pol II function on the 'typical' DNA template. Consistent with this possibility

is the observation that purified pol II supplemented with TFIIIS only, can generate transcription product similar to that generated in the HDV RNA transcription system containing only pol II and PC4 (Gottardo, 1998). Interestingly, TFIIIS-dependent RNA cleavage-extension has been also observed in binary complexes of yeast pol II and *E. coli* RNA polymerase with CPG79 RNA encompassing a prokaryotic pause site (Altmann et al., 1994; Johnson and Chamberlin, 1994). The RNA cleavage in these reactions was very slow (detectable after ~8-20 hr), and inefficient followed by addition of only a few nucleotides. In contrast, HDV RNA cleavage-extension described here is very fast (detectable at 10 min) and proceeds efficiently by precise transcription of the RNA template for up to ~40 nt. Despite the kinetic differences, the CPG79 and the AG HDV RNA-templated cleavage/extension reaction appear to be mechanistically related, since both the cleavage and extension of CPG79 and HDV RNA are sensitive to α -amanitin. Furthermore, close examination of the CPG79 RNA sequence suggests that it can be folded into a hairpin structure similar to that of the terminal hairpin of HDV RNA template used in our HDV RNA-based system.

Although the physiological role of transcript cleavage by pol II or other RNA polymerases has not been precisely established, one commonly accepted hypothesis suggests that in vivo it also stimulates pol II activity by alleviating transcriptional arrest (Reines, 1994; Reines et al., 1999). Another proposed role for the transcript cleavage is to increase fidelity of transcription by facilitating

removal of misincorporated nucleotides, as demonstrated in vitro (Erie et al., 1993; Thomas et al., 1998). Pol II mediated HDV RNA cleavage/transcription may be involved in regulation of some step during the HDV RNA replication in vivo. However, although highly precise and specific, such a reaction, that includes cleavage of the HDV RNA template cannot be easily reconciled with the rolling circle model for HDV replication, especially if envisioned as a mode of initiation of replication. On the other hand, the observed correlation between the RNA template requirements for the HDV RNA synthesis in vitro and in vivo (discussed in detail in the following sections) supports such a hypothesis and allows for refinement of the HDV replication model by incorporating a specific cleavage/transcription step.

3.3 Fidelity and processivity of the pol II-mediated HDV RNA transcription in vitro and the effect of delta antigen

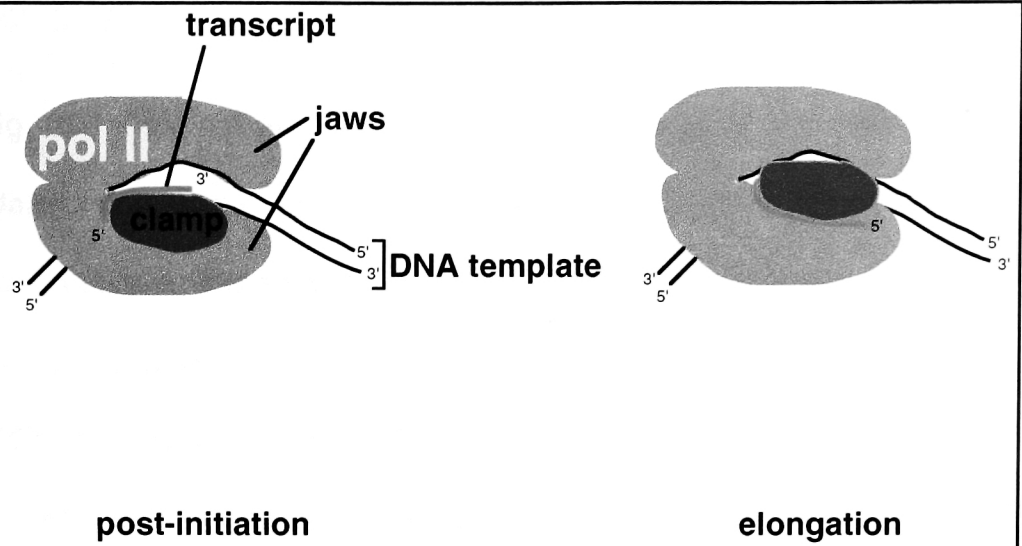
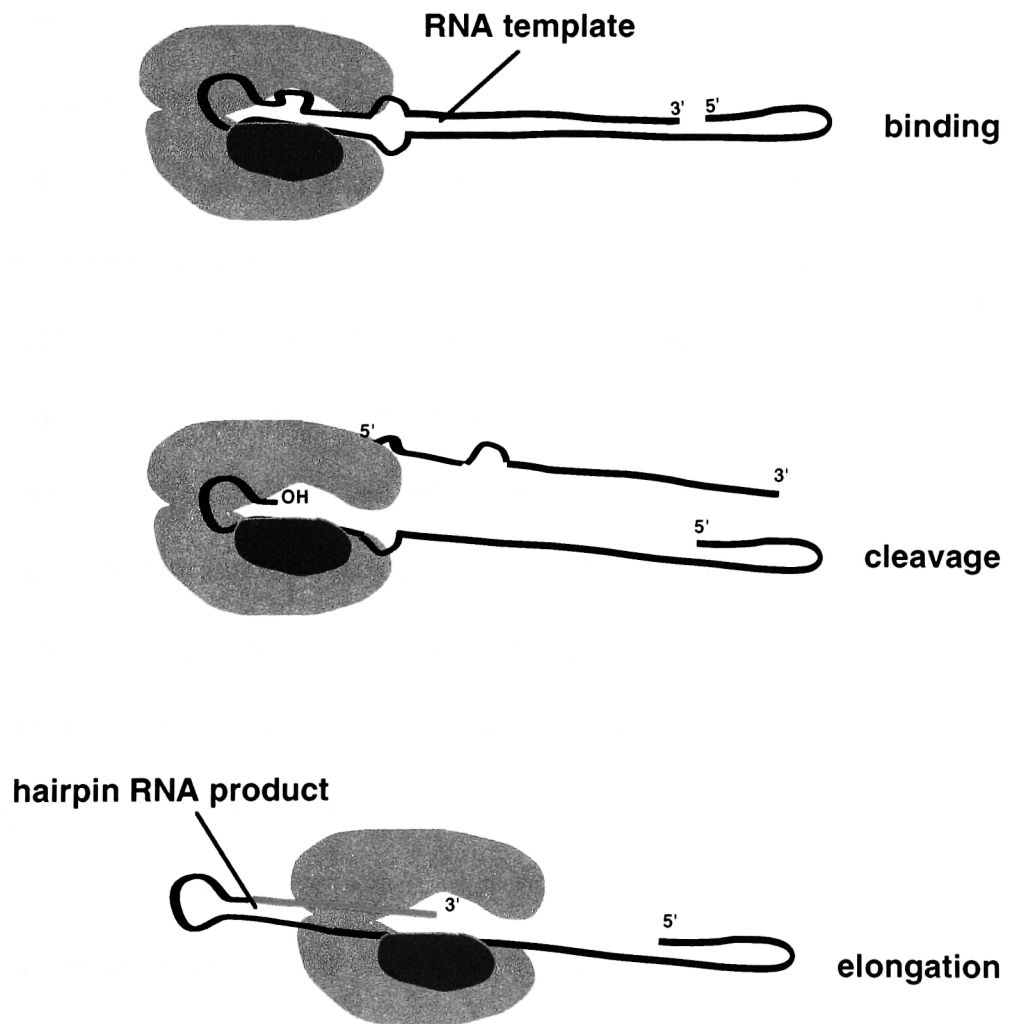
Once the HDV RNA template is recognized and pol II is able to initiate transcription, polymerization is highly precise, as mutations in the template sequence are faithfully represented in the transcript (Figures 5-7 and 10-16). However, the RNA-templated pol II transcription in vitro does not proceed to the 5' end of the template but terminates ~40 nt following initiation. What prevents pol II from accomplishing the full-length run-off transcription on the RNA template is not clear at this point and at least a couple of scenarios can be considered. First, the structure of the transcription product that involves a stable hairpin loop

(see Fig. 17) may interfere with pol II processivity, either directly or through an interaction with protein factors from the nuclear extract. It is also possible that some stimulatory protein activities, possibly characteristic to the RNA-templated pol II function, are limiting in our nuclear extracts that efficiently support DNA-templated pol II run-off transcription. Also a differential sensitivity of the RNA- and DNA-templated pol II elongation complexes to the balance of negative and positive elongation factors in the NE may affect the strong pausing of the RNA-transcribing polymerase. Finally, any combination of the proposed scenarios might also be possible.

It is tempting to further speculate on some of these possibilities in light of the recent high resolution structure for the yeast pol II and the emerging models for its function (Cramer et al., 2000). Specifically, the model by Cramer and colleagues, proposed to explain the high stability of the DNA-templated pol II elongation complexes (Linn and Luse, 1991; Uptain et al., 1997; Nudler, 1999) may also help to partly explain the inability of pol II to generate long run-off transcripts on the HDV RNA template. According to this model, early post-initiation complexes (Fig. 33A, post-initiation) are unstable and produce short abortive transcripts because the DNA template binding channel has an “open” conformation in which the DNA is held in place only by a pair of movable jaws. The late elongation complexes (Fig. 33A, elongation) are stabilized by the locking of the DNA template holding clamp that occurs when the 20 nt or longer RNA transcript occupies the product binding site that winds around the base of the

clamp. In the case of the HDV RNA-templated pol II transcription, the largely double stranded RNA template is most likely positioned at the DNA template

Figure 33. Schematic representation of DNA- and HDV RNA-templated pol II transcription. **A.** Schematic representation (modified from Cramer et al., 2000) of the predicted positioning of the DNA template and the primary transcript during the DNA-templated pol II transcription (details in the text). **B.** Schematic representation of the steps of the HDV RNA-templated pol II transcription. The positioning of the RNA template and the transcription product is purely speculative.

A**B**

binding channel (Fig. 33B, binding) and not at putative RNA binding channels that can accommodate only a single chain of nucleic acid (Cramer et al., 2000). This would also allow proper positioning of the RNA template strand relative to the pol II catalytic site and allow the specific cleavage and initiation of transcription (Fig. 33B, cleavage). However, the stable hairpin structure of the HDV RNA product that unlike the de novo initiated DNA-templated transcript includes the template sequence (Fig. 33B elongation) may hinder the entry of the transcript in the RNA channel. In such scenario, the clamp can not close the template binding channel, causing instability of the transcription complex and premature disengagement of the polymerase.

Even if the HDV RNA hairpin product could be extruded through the RNA pore, it may in fact induce a different conformational change that results in displacement of the 3' end of the transcript from the catalytic site and consequently pol II pausing. Such product-induced pol II pausing would in a way resemble the interaction of the prokaryotic RNA polymerase with the intrinsic pause and ρ -independent termination RNA hairpins (Uptain et al., 1997; Wang et al., 1997a; Artsimovitch and Landick, 1998; Platt, 1998; Nudler, 1999). In the case of pol II, no specific RNA secondary structures have been documented to influence pausing, except in the case of HIV-1 promoter proximal pausing (Palangat et al., 1998). However, there is a growing evidence that promoter proximal sequences are involved in the regulation of pol II elongation and that pausing is associated with a conformational change in the enzyme (O'Brien and

Lis, 1991; Gu et al., 1993; Krumm et al., 1995; Reeder and Hawley, 1996). The HDV hairpin product-pol II interactions may represent another example and may provide a simple and convenient model for studying the regulation of pol II elongation and understanding the nature of the eukaryotic pause signals.

As suggested above, pol II in NE may pause after only ~40 transcribed nucleotides because the stimulatory protein factors specific to the RNA-dependent pol II function are present in limiting concentrations and need to be recruited or activated by a mechanism that involves the viral protein delta antigen (HDAg). Alternatively, cellular factors may inhibit the RNA-dependent pol II function, and again the delta antigen may interfere with their function. The involvement of delta antigen in regulating pol II processivity on the RNA HDV template was examined experimentally. The results demonstrate that HDAg stimulates pol II to transcribe beyond the initial ~40 nucleotides (Fig. 18 and 19). The mechanism by which HDAg stimulates pol II transcription is not fully clear at this point. However, emerging data from the field of regulation of DNA-templated pol II elongation may provide some interesting insights. Namely, sequence analysis of the components of the pol II negative elongation factor NELF revealed a homology of one of its subunits to the HDAg (Yamaguchi and Handa, unpublished results). This finding has initiated a series of experiments which demonstrated that similarly to the HDV RNA-templated, the DNA-templated pol II transcription is stimulated by HDAg (Yamaguchi and Handa, unpublished results). It has been suggested that the stimulation of the DNA-templated pol II

elongation occurs because HDAg disrupts the NELF multisubunit complex by competitive binding and in that way interferes with its inhibitory effect on elongation (Yamaguchi and Handa, unpublished results). As described in the Chapter 1, NELF is required for the repression of pol II elongation induced by the DRB that inhibits pol II CTD phosphorylation by the p-TEFb kinase (Yamaguchi et al., 1999). Interestingly, the HDV RNA-templated transcription in NE is also sensitive to DRB and consistent with the model, HDAg alleviates this inhibition (Yamaguchi and Handa, unpublished results).

Under physiological conditions, in the absence of DRB, NELF inhibits elongation by binding to pol II that may be progressively dephosphorylated by the pol II specific phosphatase (Archambault et al., 1998). This interpretation was based on the finding the in vitro NELF binds to the hypophosphorylated but not the hyperphosphorylated form of pol II (Yamaguchi et al., 1999). Similarly, in the HDV RNA-templated transcription system, after transcribing ~40 nt, the elongating pol II could become dephosphorylated and thus accessible to DSEF/NELF binding. Addition of HDAg to the transcription reaction could disrupt functional NELF complexes, prevent binding of the negative elongation factor to the polymerase resulting in synthesis of longer transcripts, as observed (Fig. 19).

Finally, a mechanism that involves interactions between cellular positive elongation factors and the secondary structure of the RNA template and product may also be considered to explain the observed pol II pausing on the HDV RNA template in vitro. In fact, the positive elongation factor p-TEFb (the DRB target)

is likely to be involved in the RNA-templated reaction since it is sensitive to DRB (Yamaguchi and Handa, unpublished results). It is unclear how the p-TEFb kinase complex is recruited to the pol II elongating complex, even for most of the DNA-based transcription systems. In the case of HIV-1 LTR promoter, the regulatory CycT1 subunit of p-TEFb binds cooperatively with the viral Tat protein to a specific hairpin structure (TAR) generated by the 5' segment of the nascent transcript. This results in CTD phosphorylation and stimulation of the pol II elongation (Parada and Roeder, 1996; Garber and Jones, 1999). Interestingly, a strong pol II pausing occurs just before the TAR formation (Palangat et al., 1998). This pol II pausing coincides with the formation of a distinct secondary structure in the nascent transcript, a pause hairpin (Palangat et al., 1998). The mutational and functional analyses suggested that pol II pausing in this context could help fine-tune the timing of TAR formation, Tat/p-TEFb binding, and the appropriate positioning of the polymerase relative to this complex that allows for efficient CTD phosphorylation and elongation (Palangat et al., 1998). It is possible that pol II pausing and elongation on the RNA template can be similarly regulated and/or fine-tuned by complex interactions involving the specific TL/B/IL structure of the AG HDV RNA template, the product hairpin, and regulatory protein factors including also the HDAg. Such fine-tuning that favors pausing and arrest of the polymerase may be significant in the context of the HDV replication cycle (discussed in the following sections).

Addition of HDAg to the RNA-templated reaction stimulated transcription ~3-4 fold. As mentioned in the Results section, this effect can not be accounted for only by the stimulation of elongation and the incorporation of additional ^{32}P labeled G residues in the product. The transcript segment in the extended NE+HDAg product contains 9 G residues in addition to the 11 present in the ~40 nt region also transcribed in the NE product (Fig. 19). Furthermore, the δP product whose intensity was quantitatively compared to that of the NE product, contains also a small fraction of products with shorter (~40 nt) transcripts (Fig. 19B, compare lane 2 to 1 and 3). Thus, HDAg appears to also increase the number of transcripts by stimulating pol II initiation on the RNA template. The mechanism for this stimulation may include facilitating pol II recognition of the RNA template by the viral protein. As described in Chapter 1, HDAg binds HDV RNA. If it also interacts with the polymerase, the stimulation of transcription by recruitment or stabilization of pol II binding to the characteristic secondary structure of the RNA template may contribute to the process. Interaction of pol II with the HDAg, that would support this scenario, has not been demonstrated as yet and remains to be tested. The alternative possibility that HDAg stimulates the pol II-induced template cleavage in a reaction similar to the TFIIIS-induced transcript cleavage by the pol II complexes can not be excluded. Therefore, a simple in vitro system in which the cleavage and extension reactions can be uncoupled would be useful. Understanding the mechanism of HDAg function

during RNA-templated transcription in vitro may also shed light on the mechanism of HDAg simulation of HDV replication.

3.4 Specific secondary structure of the HDV RNA template determines the specificity and efficiency of the HDV RNA-templated pol II transcription in vitro

RNA secondary structure represents a structural and functional component in the regulation of a variety of biological processes, transcription by the DNA-dependent polymerases included (Simons and Grunberg-Manago, 1998). As mentioned above, in prokaryotes there are a number of well characterized examples of regulation of transcription that involve different hairpin loop structures that alone, or in concert with protein factors they recruit, regulate specific steps of the process (Platt, 1998). Studies of transcription in eukaryotes have mostly focused on the DNA promoter sequences and the complex polymerase machinery. The elongation phase of the transcription process, including the structure of the RNA transcript, has come into focus in the recent years. A number of protein factors involved in the regulation of elongation have been described and continue to be characterized. (Uptain et al., 1997; Reines et al., 1999), but the contribution of specific sequence/structure motifs in the nascent transcript is still not clear. In the case of pol II transcription, promoter proximal sequences have been shown to regulate pausing or arrest in a number of genes (Uptain et al., 1997). However, no specific secondary structure elements have

been implicated with the exception of HIV-1 pause and TAR elements (described above).

The HDV RNA-dependent function of pol II is clearly regulated by the secondary structure of the RNA template. The initial deletion analysis (Fig. 3) suggested that hairpin structure of the AG RNA template significantly affects the reaction in vitro. A single deletion in either, 5' or 3' half of the template hairpin, positioned close to the transcription start site, decreases the efficiency of the reaction (132 and 138 in Figure 3), while combining two such deletions restores the hairpin structure of the RNA and efficient transcription from these templates (129 and 103 in Figure 3). Compilation of all the mutant templates derived from the wild type AG103 template (represented in Figure 21), demonstrates that the secondary structure, rather than the primary sequence of the HDV RNA template, specifies the functional pol II templates. The large number of mutants available for the analysis pointed out that a particular secondary structure element, bulge (B), positioned adjacent to the cleavage/initiation site, specifies the functional pol II templates and affects the efficiency of the transcription reaction. The sequence and size of the other structural elements including the internal loop 1 (IL1), do not appear to be crucial for the transcription reaction in vitro, since it varies in some of the functional templates (Fig. 21B).

Enzymatic and chemical mapping of the mutant HDV RNA hairpins (Fig. 23 and 24) demonstrates that their secondary structure reflects the Mfold program predictions. In addition, enzymatic cleavage mapping pointed out to some

features and interactions in the basic TL/B/IL1/IL2 structure that can not be modeled by the program and may be important for the transcription reaction in vitro and for the replication process in vivo as suggested by the experimental data. For example, the enzymatic mapping demonstrates that a substantial proportion of the RNA18 molecules acquire a conformation with a 6 nt bulge (Fig. 23, lanes 8 and 13), similar to that of RNA13 and not the predicted 2 nt wild type-like bulge (Fig. 21). Consistent with this, both RNA18 and RNA13 are functionally comparable, since pol II transcription from these templates is reproducibly more efficient than that from the wild type template (Fig. 22).

The lead cleavage mapping (Fig. 24) confirms that all of the program-predicted secondary structure elements, i.e. the terminal loop (TL), the bulge (B, 3' of TL or b, 5' of TL), and the internal loops (IL1 and IL2), are present in the RNA templates. In addition, this secondary structure mapping demonstrates that nucleotides in the symmetric IL1 may be involved in intra- or inter-strand stacking interactions (Ciesiolka et al., 1998; Moore, 1999), since they are relatively resistant to lead catalyzed cleavage (compare IL1 to IL2 region in Figure 24). The lead cleavage is increased in the IL1 loop of RNAs 18 and 19 (Fig. 24, lanes 31 and 32). This result is consistent with presence of the predicted asymmetric IL1 loop in these RNAs (Fig. 21), and with the observations that similar asymmetric loops are thermodynamically less stable than symmetric loops of the same base content (Peritz et al., 1991; Schroeder and Turner, 2000). An exception here is the RNA75, for which asymmetric IL1 is also predicted, but no increased cleavage

is observed in the IL1 region of this RNA (Fig. 24, lane 36). Since all other structural elements are clearly detectable in this RNA it is possible that the inefficient cleavage in its asymmetric IL1 is a consequence of the different sequence context, i.e. the closing base pairs of the loop, also shown to influence the stability of the internal loops (Schroeder and Turner, 2000). In any case, alterations of the IL1 are likely to contribute to the overall geometry of the terminal hairpin of the HDV RNA. While the effects of such changes may be minute in the context of standard transcription reactions in vitro where template cleavage at the bulge B is crucial for efficient product formation, they may become apparent under certain conditions in vitro (e.g. presence of HDAg) or during the HDV replication in vivo, as suggested by the experiments discussed in the following section.

How does the secondary structure of the HDV RNA template affect the HDV RNA-templated pol II transcription is certainly interesting from the point of view concerned with the HDV replication but also with pol II function in general. It is possible that the specific secondary structure of the AG HDV terminal hairpin determines template recognition and its positioning within the polymerase. This in turn allows for the efficient template cleavage at the unique position and the correct placement of the priming 3'OH in the pol II catalytic center. The delta antigen, HDAg may facilitate the process, and other cellular protein factors are also likely to participate as the specificity of the RNA template selection is reduced in the minimal pol II/PC4 system (Gottardo, 1998).

The high specificity of the RNA-templated transcription reaction in vitro suggests that HDV replication in vivo may be regulated by similar structural elements in the viral RNA. Thus, understanding the template requirements for the pol II-mediated reaction in vitro allows for systematic examination of the structural features of HDV RNA important for its replication in vivo.

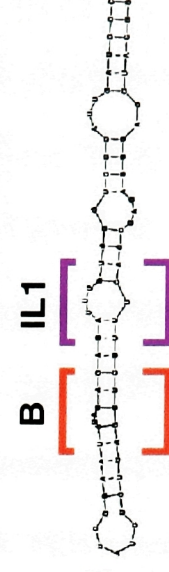
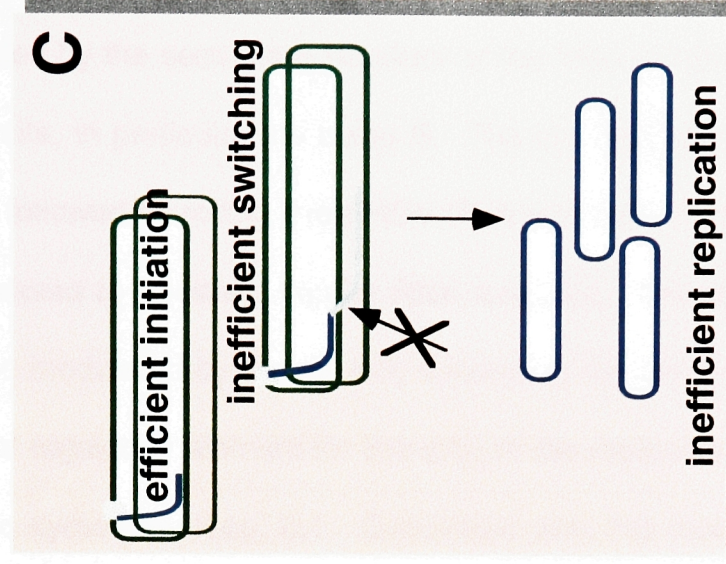
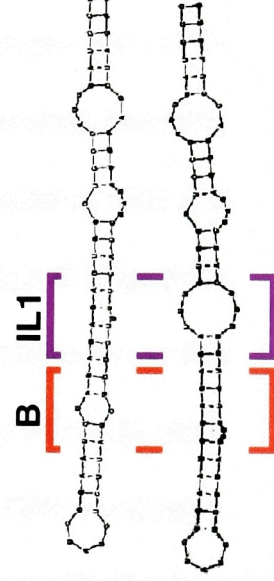
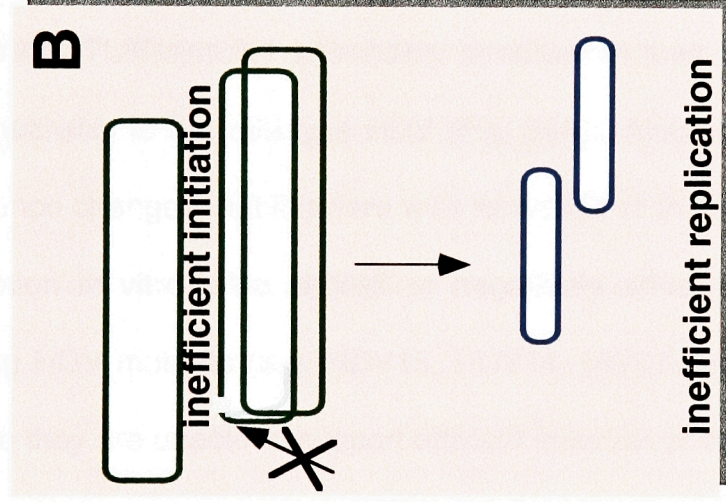
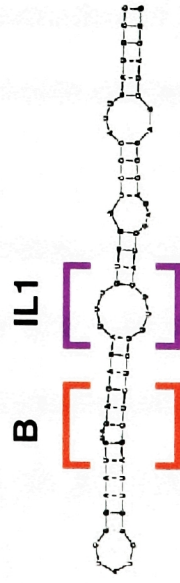
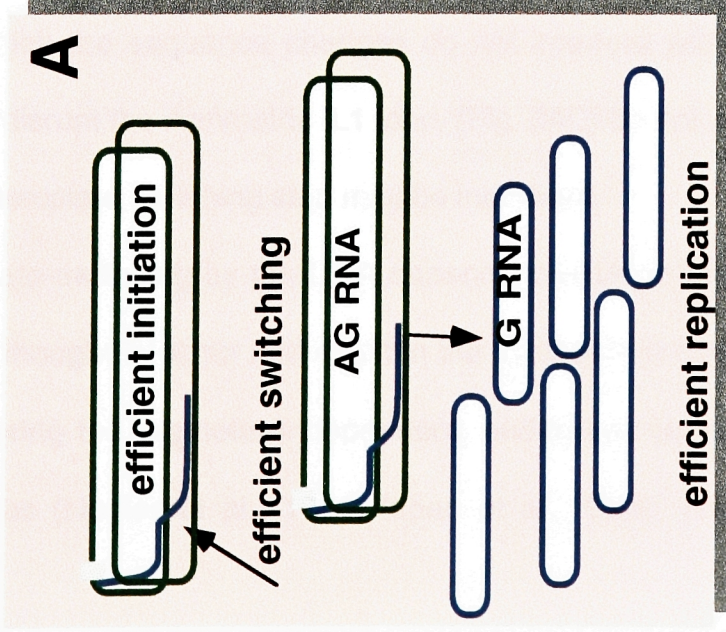
3.5 Effects of alterations of the RNA secondary structure on HDV replication in vivo- a template switching model for HDV replication

The mutational analysis of HDV replication in vivo demonstrates that similar secondary structure elements in the terminal hairpin of the AG HDV RNA regulate pol II transcription in vitro and HDV replication in vivo. More precisely, all the mutations that eliminate the bulge B positioned at the start site of pol II transcription in vitro, down-regulate or abolish HDV replication in vivo, in two cell types (RNA 15, 65, 14, and 20 see Figures 28, 29, 30, 31, and 32). Furthermore, sequence changes that do not affect the characteristic wild type structure of the terminal hairpin, replicate comparably to the wild type (HDV 63, and HDV73 in Figures 28, 29, 31, and 32). These results strongly suggest pol II involvement in the HDV replication process. Moreover, they suggest that HDV RNA synthesis in vivo initiates by a similar mechanism to the pol II-mediated AG HDV RNA-templated transcription in vitro, i.e. by cleavage of the circular template and using the newly created 3'OH as a primer for RNA synthesis.

Interestingly, mutations 18 and 75 (2 nt deletion and sequence alterations, respectively) that do not disrupt the characteristic bulge B, and do not interfere with pol II activity in vitro, have a negative effect on HDV replication in vivo, in both, COS7 and HeLa cells (Fig. 28, 29, 30, 31 and 32). These results reveal differences between the experimental systems in vitro and in vivo, and point out to secondary structure elements that in addition to the characteristic bulge B are involved in regulation of HDV replication in vivo. In fact, a closer examination of the secondary structure in the hairpin region of these RNAs shows that they contain an asymmetric, instead of the wild type-like, symmetric internal loop IL1. It should be recalled here that the symmetry of the IL1 had no effect on the pol II transcription in vitro where only appropriate positioning of the bulge B relative to the TL, is sufficient for the template cleavage and formation of the transcription product. Thus, the negative effect of the IL1 alteration on HDV replication suggests that in vivo, initiation and/or additional steps of HDV RNA synthesis may also be regulated by specific secondary structure elements within the terminal hairpin region.

A modified rolling circle model for HDV replication that helps explain the correlation between the secondary structure requirements for the HDV RNA synthesis in vitro and in vivo, as well as the differences observed in the case of HDV18 and HDV75 mutants, is illustrated in Figure 34. This model considers that, similarly to the HDV RNA-templated pol II transcription in vitro, initiation of HDV RNA synthesis in vivo, may involve cleavage of the circular RNA template

Figure 34. Template switching model for HDV RNA replication. A . Complementary RNA synthesis initiates by cleavage of the AG HDV RNA template adjacent to a specific bulge structure, as in vitro. After a segment of the template is transcribed pol II 'jumps' to a new circular template and proceeds by rolling circle copying of the HDV RNA. Similarly to the initiation step, the template switching step is regulated by the specific secondary structure of the terminal hairpin of the RNA templates that contain a symmetric internal loop IL1. Thus, HDV mutants that can fold into the wild type-like TL/B/IL1/IL2 structure at the left-hand terminal tip of the AG RNA replicate as efficiently as wt HDV. **B.** HDV mutants in which the characteristic bulge B can not be formed fail to replicate efficiently because of inefficient initiation of transcription. **C.** HDV mutants in which the characteristic symmetric internal loop IL1 can not be formed fail to replicate efficiently because of inefficient pol II template switching.



that is regulated by the secondary structure of the RNA template encompassing the initiation site, in particular the bulge B. The cleavage-mediated initiation *in vivo*, may be followed by rolling circle HDV RNA synthesis if pol II switches from the initial, cleaved to another, circular RNA template. The template switching would be also modulated by the specific secondary structure of the HDV RNA, but the crucial regulatory element for this step of the replication process is likely to include the symmetric loop, IL1. Consistent with this model, HDV mutants (e.g. HDV63 and HDV73) that contain extensive sequence changes but retain the characteristic TL/B/sym.IL1 secondary structure in their terminal hairpins, replicate comparably to the wild type HDV (Fig. 34A). Also consistent with this model, sequence changes that interfere with formation of the bulge B (Fig. 34B) and transcription *in vitro*, also abolish or negatively affect replication of the corresponding HDV mutants (e.g. HDV15, HDV14, HDV20, and HDV65) most likely because they are unable to support efficient initiation of the RNA synthesis. Finally, and again consistent with the proposed model, mutants HDV18 and HDV75 in which the sequence changes do not interfere with formation of the bulge, B, but disrupt the symmetric IL1 loop, (Fig. 34C) do not replicate efficiently because the template switching step may be inefficient.

Template-switching by the DNA dependent RNA polymerases, including pol II, is also thought to occur and result in the synthesis of longer than template transcripts during the promoter-independent, end-to-end transcription on linear DNA templates (Nudler et al., 1996; Izban et al., 1998; Rong et al., 1998).

Furthermore, sequence comparison of a number of viroid variants suggested that RNA template switching may frequently occur during viroid replication, and the terminal hairpin structures have been specifically implicated in this process (Hammond et al., 1989; Koltunow and Rezaian, 1989; Rezaian, 1990; Hernandez and Flores, 1992; Semancik et al., 1994; Diener, 1995; Kofalvi et al., 1997). It has also been proposed that RNA template switching may have occurred earlier in evolution during the pol II-mediated RNA replication in the cell. Namely, identification of the Hepatitis Delta Interacting Protein A (DIPA), a human homolog of HDAg, led to the hypothesis that HDV evolved from a viroid-like RNA that acquired HDAg ORF as a result of pol II switching templates from a viroid ancestor to the cellular mRNA (Brazas and Ganem, 1996). In fact, such an event would have to occur during the genomic RNA synthesis at position 3' from the end of the HDAg ORF i.e. in the left-hand terminal tip of the AG RNA template (see Figure 2A). This is the same region, in which we observe pol II pausing after the initiation of transcription on the AG RNA template in vitro and where the symmetric IL1, proposed to regulate the switching step during the HDV replication, is located. Thus, it seems feasible that pol II can switch templates during the replication in the left-hand terminal hairpin of the AG HDV RNA. A post-initiation pausing in this region may in fact promote a conformational change of the polymerase and facilitate the acquisition of the new template. Distinct, but overlapping secondary structure elements of the AG HDV RNA templates and the product hairpin, together with host cell factors and the delta antigen, may

regulate each step of the process and fine-tune their coupling in a functional replication process. In fact, the RNA binding activity of HDAg, together with its ability to dimerize, could promote pol II template switching by bringing two template molecules in a close proximity to each other. The observed stimulatory effect of HDAg on pol II elongation (Fig. 18 and 19) may also contribute to this process through an interaction with some auxiliary factors or the polymerase itself, as discussed above.

The template-switching model that includes AG RNA cleavage and initiation of transcription in the cells, predicts the existence of chimeric AG/G RNA molecules, at least early in the replication cycle. Attempts to identify such chimeric RNA species early post-transfection failed, but this result can be easily explained by inability of the otherwise sensitive RT-PCR assay to detect these highly structured RNA molecules. Testing the prediction of the template-switching model, i.e. recombination between the replicating HDV RNAs in the cell, by a more functional assay would also be complicated and a number of points should be considered. First, two different mutant HDV RNAs would have to be delivered to the same cell. This can be accomplished by the transient transfection of a single expression vector containing two mutant HDV cDNAs under two different promoters (e.g. pBudCE4, Stratagene) or by the transient transfection of a pBudCE4 vector carrying one HDV mutant and a marker (e.g. GFP) cDNA into a cell line that steadily expresses HDV RNA for the second mutant. The marker protein would allow for an identification and FACS

(Fluorescence Activated Cell Sorting) sorting of the transfected cells so that an equal number of transiently transfected control or cells steadily expressing mutant HDV RNA can be analyzed. Although more time consuming, the second strategy may be more useful, because it allows for testing different ratios of the two mutants. It may also help avoid the serious problem of the DNA recombination between the HDV cDNA sequences, at least during the subcloning steps. To avoid the DNA recombination problem, an RNA cotransfection assay can be considered, however, such an assay may be harder to control, especially in terms of the quality of RNA preparations and the efficiency of transfection. Even if an efficient and controllable method of introducing two different HDV mutants into a single cell can be established, it will be difficult to detect a clear difference between the co-replication of two mutant HDV RNAs and replication of each mutant alone. In fact, there is no combination of mutants defective in one of the two considered steps of replication (initiation and switching) that when combined, could produce an efficiently replicating progeny. For example, combination of the mutants described in Figure 34B (containing the sym. IL1) and C may be considered including the possibility that inefficient template switching could occur from the initiation competent mutant C to an acceptor mutant B. However, subsequent initiation from the progeny, synthesized by rolling circle copying of the HDV RNA B could not initiate new replication cycles and thus no positive effect on the co-replication can be expected in vivo. Similar considerations are valid for a number of other combinations of mutants.

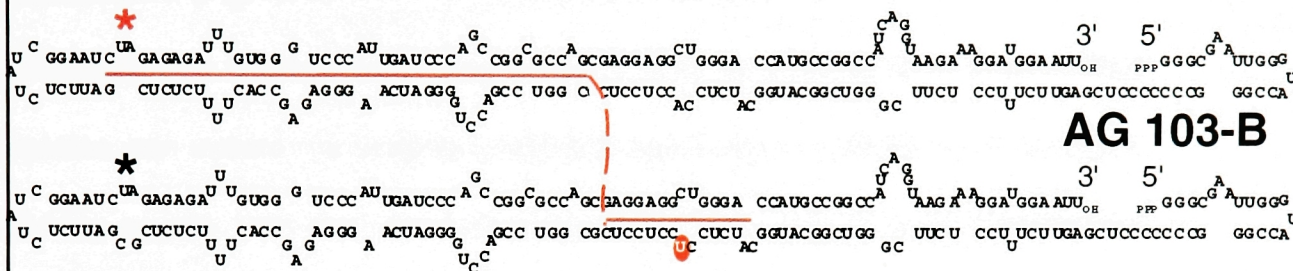
At this point, a modified in vitro transcription system seems to offer a more appropriate alternative for testing the pol II template-switching model. One prediction of the model is that pol II switches templates in the vicinity of IL1, most likely in the region at which it pauses during transcription in standard nuclear extracts. In addition, since HDAg stimulates pol II elongation in this region, and has the ability to bind RNA and dimerize, it may release the pause and promote RNA template switching, if an appropriate recipient mutant template is provided in the reaction. A possible experiment is outlined in Figure 35A. Standard transcription reaction can be initiated using the wild type AG103 template or a neutral mutant containing the wild type secondary structure with sequence alterations that could be informative in the final sequence analysis (AG103-A). At different time points an initiation-incompetent mutant (AG103-B), containing a symmetric IL1 loop (predicted to be crucial for the switching step) and some marker mutation (alters the RNase digestion pattern of wt (A) and mutant (B) or switch (S) complementary RNA) would be added together with or without HDAg. If the direct sequence analysis of the transcription products of the reactions containing both templates, demonstrates that the marker sequence is copied, this would strongly suggest that pol II switched transcription from the functional to the initiation-incompetent mutant. The comparison of the product analysis from the reactions with and without HDAg would demonstrate if the viral protein plays a role and modulates the pol II RNA template switching.

Figure 35. Description of two different approaches designed to examine if pol II RNA template switching can occur in vitro. Sequences of the AG103-based RNA templates that could be used in a standard **(A)** or pulse-chase **(B)** transcription the experiment. The segments transcribed in the presence of α -³²P-NTP or 'cold' NTP are underlined by the red or black bars respectively. The red and black asterisks designate the site corresponding to the template cleavage and initiation of pol II transcription in functional and nonfunctional templates, respectively. The arrows indicate the predominant 3' end of the pol II transcript generated in NE in the absence of HDAg. P-A, P-A1, and P-A2 designate transcription products (full length **(A)**, or only the pol II transcribed sequence **(B)**) generated using the templates AG103-A, AG103-A1, and AG103-A2, respectively. P-S designates the full length **(A)**, or only the pol II transcribed sequence **(B)** of a hypothetical product that generated by pol II RNA template switching in vitro. α -³²P-NTP labeled segments are represent in red. The RNase A sites in each of the transcription products are designated by the vertical bars and the length (in nt) of the resulting oligonucleotides is indicated. The RNase A digestion patterns of the products expected in the experiment outlined in panel B is schematically represented in the lower right corner (digestion of products generated in standard (lanes 1-4) or pulse-chase (lane 5) txn. reactions). Detailed description of the proposed experiment is presented in the text.

A

B IL1

AG 103-A



P-A



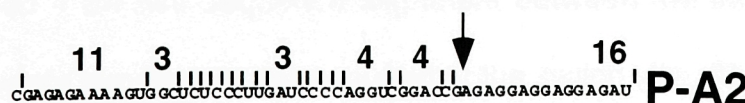
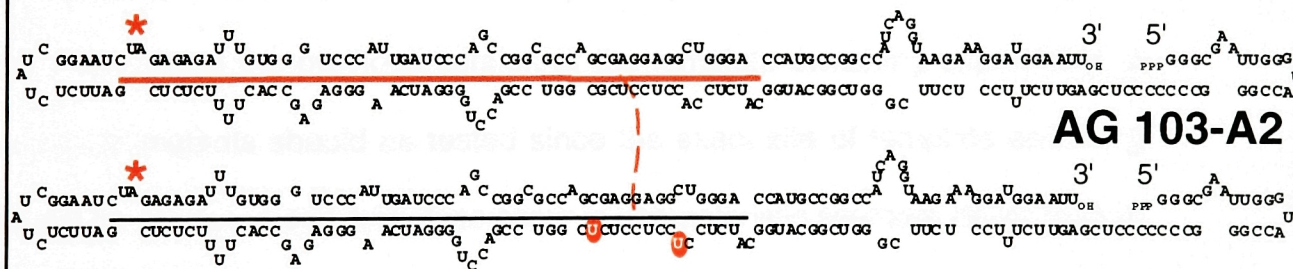
P-S



B

B IL1

AG 103-A1



A	B	A	B	S	HDAg
-	-	+	+	+	
					16
					11
					8
					6
					4
					3
1	2	3	4	5	

An alternative modification of the in vitro switching experiment could also be considered (Fig. 35B) that includes two functional templates (AG103-A1 and AG103-A2) that contain both, the bulge B and symmetric IL1. These RNA templates can initiate transcription, and are predicted to support pol II template switching as well. Instead standard of the standard a pulse-chase transcription would be more appropriate with those templates since both can initiate transcription. NE transcription reaction should be initiated in the presence of labeled NTP with or without AG103-A1 template. At different time points excess of a cold NTP, and the template AG103-A2, should be added in the presence or absence of HDAg. This would allow pol II to synthesize the A1 labeled transcript segment and if it switches, to link it to the cold transcript of the marker region of the AG103-A2 template. If the digestion pattern of the product of the pulse-chase transcription reaction with the two templates contains labeled fragments that correspond to both, A1 and A2 templates (lane S in the hypothetical gel in Fig. 35B). As in the previous version of the template switching experiment, a number of mutants should be tested since the exact site of template switching can not be predicted and is likely to occur in characteristic segment rather than at a specific single nucleotide. Therefore, these in vitro tests of pol II RNA template switching may pose different limitations. For example, although the RNA structure requirements for pol II switching may be fulfilled, the reaction may be hampered if precise sequence alignment between the incoming transcript and the acceptor template is also required at the switch site. Thus, further analysis of

these in vitro and in vivo HDV RNA-based systems is needed to better understand their specific requirements and limitations, and develop new experimental approaches. In the process, combining the limited information from the available experimental systems, can point out new directions and in long run, allow for a better view and understanding of the RNA-templated pol II function, the putative template switching reaction and HDV replication.

Finally, the results with the AG HDV RNA-templated pol II transcription presented here, and similar studies with the 'atypical' pol II templates, combined with the increasingly available structural and functional data from the DNA-based systems, may help generate new insights into pol II function and the greater puzzle of gene expression.

Chapter 4

CONCLUSIONS

A segment of the HDV RNA represents a specific template for pol II transcription in vitro. This RNA template (~218) encompasses the left-hand terminal hairpin region the rod-like AG HDV RNA (1.7 kb) (Figures 2-4).

Specificity of the HDV RNA-templated pol II transcription in vitro is determined by the secondary structure, rather than the primary sequence, within the 62 nt terminal hairpin region of the RNA template (Figure 20). The terminal hairpin of all functional pol II templates contains the terminal loop (TL) and a bulge B, positioned 3' of the TL, followed by internal loops IL1 and IL2 (Figure 21). Extensive sequence changes in this region that do not change its characteristic secondary structure do not affect pol II transcription of the mutant templates in vitro. On the other hand, even small sequence changes that specifically interfere with the formation of the bulge B, fully abolish transcription (Figure 21).

Efficient replication of HDV RNA in vivo, also depends on the presence and the positioning of the characteristic bulge B (Figures 28-32). This correlation between the template requirements for efficient HDV RNA transcription in vitro and in vivo strongly supports pol II involvement in HDV replication. The mutational analysis in vivo, also reveals that another secondary structure element, the symmetric IL1, in addition to the bulge B, is crucial for HDV RNA replication (Figures 28-32).

The product of HDV RNA-templated transcription reaction in vitro is a chimeric molecule (Figure 17). It contains the 5' segment of the AG template and

genomic transcript segment covalently linked to it. Thus, initiation of HDV RNA-templated pol II transcription must involve cleavage of the RNA template.

The template/transcript junction is located at a specific and unique position relative to the secondary structure of all the functional templates. In fact, RNA template cleavage occurs just adjacent to the characteristic bulge B (Figure 17). This generates a new 3'OH end that is used as a primer for precise pol II transcription. The pol II-mediated HDV RNA cleavage/transcription reaction mechanistically resembles the endonucleolytic cleavage of the nascent RNA that occurs in transcription complexes arrested during the DNA-templated pol II elongation as well as in binary complexes of yeast pol II with an RNA molecule encompassing a prokaryotic pause site (see Uptain et al., 1997). Thus, initiation of the HDV RNA-templated transcription reaction in vitro most likely involves the elongation mode of pol II function.

The common RNA structure requirements for the processes in vitro and in vivo, suggest that initiation of HDV RNA replication also involves cleavage of the circular AG RNA template. However, coupling of a cleavage-induced initiation to a rolling circle RNA synthesis that generates longer than unit length HDV RNA molecules detected in the cells, must involve pol II switching from the initial cleaved, to a new circular RNA template. Such template switching or RNA recombination step is most likely, as in the case of the other RNA-dependent RNA polymerases (Nagy et al., 1999; Nagy and Simon, 1997; Nagy and Simon, 1998) regulated by specific secondary structure elements in the donor and

acceptor RNA templates. Since the symmetric internal loop IL1, positioned close to the initiation bulge B is clearly required for the replication process in vivo but not for transcription in vitro, it is most likely involved in the regulation of a post-initiation step that involves pol II template switching.

In HeLa nuclear extracts, pol II elongation on the RNA template stops after only ~40 transcribed nucleotides. Further elongation is stimulated by addition of the viral protein, delta antigen (Figures 18 and 19), but still, transcription does not efficiently proceed to the 5' end of the template. Pol II elongation on DNA templates is also stimulated by the delta antigen (Y. Yamaguchi, personal communication). These results suggest that the HDV RNA- and the DNA-templated pol II transcription are not only mechanistically similar but they also share similar regulatory mechanisms.

The reasons for the inability of pol II to transcribe the full length of the RNA template remain to be further investigated. One interesting possibility is that the stop of pol II elongation on the HDV RNA template observed in vitro may in fact, reflect a regulatory point of HDV RNA replication in vivo responsible for efficient coupling of the cleavage-induced initiation to the template switching step.

Thus, integrative interpretation of the results presented in this study of HDV RNA-templated pol II transcription in vitro and in vivo, allows for refinement of the rolling circle model for HDV replication. The new model is represented in Figure 34: Initiation of G HDV RNA synthesis occurs at specific site in the

terminal hairpin of the AG HDV RNA and involves cleavage of the RNA template that is followed by limited elongation and subsequent RNA template switching by the polymerase. After template switching, rolling circle RNA synthesis proceeds on the newly acquired circular template. The initiation, as well as the template switching step, is regulated by secondary structure elements in the terminal hairpin loop of the AG HDV RNA. Since both, the bulge B and the symmetric loop IL1 are required for efficient HDV replication *in vivo*, it is possible that the former is crucial for the regulation of the initiation, and the latter for the regulation of the template switching step. The perfect base paired hairpin structure of the post-initiation AG/G chimeric product may also be involved in the regulation of the process by directly interacting with the polymerase or other proteins involved in the regulation of pol II processivity and/or pausing on the RNA template (see Discussion sections 3.3 and 3.5 for more details).

This study also reveals mechanistic and regulatory similarities between the RNA- and the DNA- templated elongation process and provides a new and simple system to study this important regulatory step of pol II transcription and gene expression in general.

Chapter 5

MATERIALS AND METHODS

5.1 Recombinant DNA technology

All recombinant DNA manipulations and procedures were carried out essentially as described in (Ausubel et al., 1994; Sambrook et al., 1989).

5.2 Plasmids and construction of mutants for in vitro analysis

HDV cDNA of the Italian isolate (a gift from N. Houghton) was used in the construction of the HDV-derived plasmids. The pSG200 plasmid used for synthesis of AG200 RNA template contains Sma I fragment of HDV cDNA between pos. 481 and 1119 (Wang et al. 1986), cloned into pSKII⁺ vector. NcoI cut pSG200 was used for T3 transcription to generate AG200 RNA containing HDV sequences from pos. 576 to 1109. As compared to pSG200, pSG128 contains a 183 nt deletion between HDV pos. 481 and 664 and a 198 nt deletion between HDV pos. 912 and 1119. All plasmids were linearized with EcoR I for synthesis of AG RNA templates. As compared to pSG128, pSG130 contains a 28 nt deletion between HDV pos. 863 and 890, pSG131 contains a 26 nt deletion between pos. 693 to 718, pSG132 contains a 33 nt deletion between HDV pos. 718 and 768 and pSG138 contains a 31 nt deletion between HDV pos. 822 and 858. pSG129 contains both deletions described for pSG131 and pSG130, while pHS103 contains both deletions described for pSG132 and pSG138.

The mutated versions of the pHS103 used for RNase mapping experiments were constructed using a PCR-based approach, except for pHS103-77 mutant, in which Taq I fragment of HDV cDNA, encompassing the terminal

loop of the template was introduced in an opposite orientation by Taq I digestion and religation.

The DNA templates for synthesis of SP6-G103 and SP6-G103-63 were generated from the corresponding pHS103-based constructs by PCR using the primers **HDV71**: CGCGATTTAGGTGACACTATAGAAGAATCGAGAGAA

or **HDV82**: CGCGATTTAGGTGACACTATAGAAGAAATCTCTC

respectively, (contain SP6 promoter sequence)

and **HDV 72**: GGGCGAATTGGGTACCGGGC .

The DNA templates for synthesis of the terminal hairpin RNAs used for secondary structure mapping were generated from the corresponding mutant pHS-103-based constructs by PCR using the following HDV primers:

HDV52: CCGAATTCTGGGATCAATGGG

HDV53: GTAATACGACTCACTATAGGGGATCCAGGGAG (contains the T7 promoter sequence)

All sequences were confirmed by sequencing.

5.3 Plasmids and construction of mutants for in vivo analysis

The starting plasmid for construction of HDV expression vectors was pGEHX4HDV that contains a monomer HDV cDNA (Sma I 480/481 fragment) at the Sma I site of the polylinker. The vector is pGEM4-based from which all the restriction sites between the Hind III and Xba I (they included) were eliminated. The mutations described in the text were introduced into the HDV cDNA sequence by PCR based approach.

Plasmids containing wild type or mutant HDV cDNA dimer sequence (pGEHX4HDV2) were generated by subcloning a monomer HDV cDNA (pos. 480-481) excised from the pGEHX4HDV construct into a Sma I linearized corresponding pGEHX4HDV vector. To make the plasmids containing 1.3X copy of wild type or mutant HDV cDNA sequence (pGEHXS4HDV1.3) the Sma I at the pos. 480 of the HDV cDNA in the pGEHX4HDV vector was modified to Srf I site by PCR to generate the pGEHXS4HDV. A wild type or mutant Sma I (pos. 481-1119) HDV cDNA fragment was then subcloned into the corresponding Srf I linearized pGEHXS4HDV vector to generate pGEHXS4HDV1.3.

The dimer or 1.3X copy HDV cDNA was excised from the corresponding vectors by Bam H I and Ecl I digestion, the 5' protruding ends were filled in by Klenow enzyme and subcloned into Ecl I linearized pSVL (Pharmacia) expression vector to generate pSVL2 or pSVL1.3 respectively.

The pSVL-HDAg expression vector was generated by subcloning the Sca I Sal I blunted HDV cDNA fragment (pos. 1624-991) into a Xba I/BamH I digested and Klenow filled pSVL expression vector.

All sequences were confirmed by sequencing and the orientation of the inserts was examined by appropriate restriction analysis.

5.4 Preparation of RNA templates for HeLa NE transcription

RNA templates were prepared from linearized plasmids using T3 or T7 RNA polymerase in the presence of 40mM Tris-HCl, pH8.0, 8mM MgCl₂, 2mM

spermidine, 50mM NaCl, 10mM DTT and 1mM of each NTP. Typically, transcription was carried out at 37°C for 2.5 hrs, after which the RNA was phenol extracted and ethanol precipitated. The RNA was then incubated in the folding buffer (0.5M NaCl, 50mM CAPSO, pH8.6 at 60°C, 2mM MgCl₂, and 0.1mM EDTA, pH8.0) for 2 hrs at 60°C, ethanol precipitated, resuspended in 25mM Tris-glycine/25% glycerol and resolved in a 5% non-denaturing polyacrylamide gel in 50mM Tris-glycine pH 8.8. The RNA visualized by UV shadowing was eluted in 0.5M NH₄OAc, 10mM MgCl₂, 0.1mM EDTA, 0.1% SDS buffer and ethanol precipitated. RNA was dissolved in H₂O and stored at -20°C.

5.5 Preparation of control RNAs for the RNase digestion analyses

Body labeled RNA controls for the digestion analyses were prepared by transcription from linearized plasmids or DNA templates generated by PCR (described above) using T3, T7 or SP6 RNA polymerase in the presence of 40mM Tris-HCl, pH8.0, 8mM MgCl₂, 2mM spermidine, 50mM NaCl, 10mM DTT, 50μM GTP, 1mM of ATP, CTP and UTP each and 2.5 μM α-³²P-GTP (800 Ci/mmol, New England Nuclear). Control capped RNA was synthesized by SP6 RNA transcription as above except a GpppG dinucleotide (40mM) was also added to prime transcription (Konarska et al., 1984). Typically, transcription was carried out at 37°C for 2.5 hrs, after which the RNA was phenol extracted, ethanol precipitated, dissolved in H₂O and stored at -20°C.

5.6 Preparation of 5' end labeled RNAs for secondary structure mapping

The RNAs used for secondary structure mapping were synthesized as above except no α - ^{32}P -GTP was added. After synthesis, RNA was phenol extracted, ethanol precipitated and subjected to Calf Intestinal Phosphatase (CIP) treatment for 1hr and again phenol extracted and ethanol precipitated. The CIP-treated RNA was then 5' end labeled by T4 polynucleotide kinase (New England Biolabs) with γ - ^{32}P -ATP (3000 Ci/mmol, New England Nuclear) in the presence of 100mM Tris-HCl pH7.5, 30mM β -mercaptoethanol, 20mM MgCl_2 , 3.5mM spermidine for 1hr at 37°C. 5' end ^{32}P labeled RNA was gel purified, eluted in the buffer described in the procedure for preparation RNA template for transcription in NE, ethanol precipitated, dissolved in H_2O and stored at -20°C .

5.7 Transcription reactions in HeLa NE

Nuclear extracts from HeLa or PMG cells were prepared as described by (Dignam et al., 1983). Optimal conditions for NE transcription with regard to temperature, NE, NTP, Mg^{2+} , and salt concentration were determined for both AG200 and AG103 templates and shown to be in the range typical of RNA pol II reactions. Transcription reaction (15 μl) typically contained 5 to 7 μl of NE (protein concentration $\sim 10 \mu\text{g}/\mu\text{l}$) 12mM HEPES, pH7.9, 12% glycerol, 60mM KCl, 0.12mM EDTA, 0.5mM DTT, 8mM MgCl_2 , 6mM of three NTPs and 0.4 μM (800 Ci/mmol, New England Nuclear) [^{32}P]- α -NTP. The ^{32}P labeled nucleotide was GTP, unless otherwise specified. His-tagged delta antigen, purified from

BL21 cells (kindly provided by J. Doudna at Yale University) was added only where indicated. Approximately 600 fmol of the RNA template was added (although in many cases titration of the RNA template was used to determine the optimal concentration). In control, DNA-templated reactions, 125ng (60pmol) Sma I-digested pSmaF DNA (Weil et al., 1979) was used to generate Ad2 MLP RNA. Reactions were incubated for 1hr at 30°C, RNA products were phenol extracted, ethanol precipitated and resolved in 5 or 7% polyacrylamide/8M urea gels.

5.8 Phosphatase and β -elimination treatment of the RNA

Where indicated eluted RNA was subjected to Calf Intestinal Phosphatase (CIP) treatment (1 hour at 60°C with 1U AP (Boehringer)) or β -elimination treatment prior to RNase digestion. The β -elimination treatment was essentially as described in Muthukrishnan et al. (1975). Briefly, the RNA was first incubated in 0.1M sodium acetate pH5.6 and 0.1M NaIO₄ for 1.5 hrs at room temperature in dark after which glycerol was added to final concentration of 7% and the reaction was incubated for 15 more minutes. The finally aniline was added to 0.5M final concentration and the reaction was incubated for 2 more hrs in a dark place, after which the RNA was phenol extracted, ethanol precipitated and used for RNase digestion.

5.9 RNase T1 and RNase A digestion

RNA products of T3, SP6 or NE transcription reactions carried out in the presence of α -³²P-GTP unless indicated otherwise, were gel-purified and eluted in buffer containing 50mM Tris-HCl pH8.3, 2mM EDTA, 0.1% SDS and 0.2M NaCl. The eluted RNA was then precipitated with ethanol in the presence of 40 μ g of glycogene and dissolved in 10 μ l H₂O. Prior to digestion RNA was denatured for 10 min at 95°C, and immediately placed on ice. The digestion buffer was then added to a final concentration of 50mM Tris-HCl pH7.7, 0.2 μ g/ μ l yeast tRNA, and 2U/ μ l RNase T1 (Calbiochem) or 2.5 ng/ μ l RNase A (Boehringer) and the reaction was incubated for 2 hrs at 50°C or 30 min at 37°C respectively. Digestion products were dried, dissolved in FA dye without bromophenol blue (90% deionized formamide, 1mM EDTA, 0.05% xylene cyanol) and resolved in a 25% polyacrylamide/8M urea gel.

5.10 RNase H analysis

For RNase H mapping the RNA to be analyzed was gel purified and precipitated as described for RNase T1 and RNase A digestion. The annealing was performed by denaturing for 5 min at 95°C followed by slow cooling to 37°C of 6.6 μ l reaction containing the RNA to be analyzed, 50mM Tris-HCl pH8.3, 10mM DTT, 60mM NaCl and 1 μ g of a specific DNA oligonucleotide. The sample was then placed on ice and 0.3U RNase H (Boehringer) and 4mM final MnCl₂ was added. The reaction was incubated for 30 min at 37°C. The products were

phenol extracted, ethanol precipitated and resolved in a 7% polyacrylamide/8M urea gel. Sequences of the used DNA oligonucleotides are as follows:

a-5' CTCGAGGGGGGGCCCGGT3'

b-5' TGGAGATGCCATGCCGA3'

c-5' CGCGGATCCAGAGCCACTTTTCTCTC3'

d-5' CGCGGATCCGACCTGGGGATCAAGGG3'

5.11 RNA secondary structure predictions

The predicted RNA secondary structures were generated using the Mfold program by Zuker and Turner at <http://mfold2.wustl.edu/~mfold/rna/form1-2.3.cgi>.

The lowest energy structure for each mutant RNA template at 30 °C is shown.

5.12 Enzymatic mapping of RNA secondary structure

The 5' end labeled RNA (10K cpm for each mutant) was digested for 40 min at room temperature with 5U RNase PhyM or for 7.5 min with 0.01U RNase T1 in the presence of 50mM Tris-HCl pH7.7, 60mM KCl, 8mM MgCl₂. The digestion reactions were stopped by 200 fold dilution with a lithium extraction buffer (5mM Tricine pH7.5, 0.25M LiCl, 0.25 mM EDTA and 0.05% SDS), RNA was immediately phenol extracted, ethanol precipitated, dissolved in 5µl of FA dye.

5.13 Lead cleavage mapping of RNA secondary structure

The 5' end labeled RNA (10K cpm for each mutant) was incubated for 7 min at room temperature in 20µl volume containing 6µg of carrier tRNA, 10 mM Tris-HCl pH7.2, 60mM KCl, 8mM MgCl₂ and 0.5 or 5mM lead acetate (as indicated in the specific experiment). The cleavage reaction was stopped by adding 2.5µl 0.5M EDTA and 200 fold dilution with the lithium extraction buffer (5mM Tricine pH7.5, 0.25M LiCl, 0.25 mM EDTA and 0.05% SDS). RNA was then ethanol precipitated, dissolved in 5µl of FA dye and resolved in a 10% polyacrylamide/8M urea gel.

The RNA ladder marker was generated by incubating 10K cpm of wild type RNA in 50mM sodium carbonate pH 9.0 for 6 min at 97°C. The reaction was stopped by adding sodium acetate pH 4.6 to 0.3M final concentration. RNA was then ethanol precipitated, dissolved in 5µl FA dye and resolved in a 10% polyacrylamide/8M urea gel.

5.14 Cells and transfections

Cos7 and Hela cells were cultivated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf serum, 100U/ml penicillin and 100µg/ml streptomycin, 1X MEM non-essential amino acids and 2mM L-glutamine. All solutions except the calf serum from GibcoBRL. On the day of transfection the cells were 80% confluent, they were trypsinized and washed twice and then resuspended in phosphate buffered saline (PBS) pH7.4 to 10⁷ cells/ml (COS7) or

5×10^6 cells/ml (HeLa). 500 μ l of the cell suspension was transfected by electroporation by 800V and 25 μ F (Cos7) or 900V and 25 μ F (HeLa) with 5 μ g Green Lantern-1 Vector (Life Technologies) carrying the ORF for a mutated version of the Green Fluorescent Protein (GFP) under the control of CMV promoter. The amounts of the HDV expression vectors that generated maximal wild type replication signal were as follows in Cos7 cells 8.5 μ g of the pSVL2-based or 11.1 μ g of the pSVL1.3-based vector. Where indicated 5.7 μ g of the pSVL-HDAg vector was cotransfected. In HeLa cell the DNA amounts were as follows: 8.5 μ g of the pSVL2-based or 7.4 μ g of the pSVL1.3-based vector and 5.7 μ g of the pSVL-HDAg. Twenty four hours post transfection the percentage of transfected cells was determined by Fluorescence Activated Cell Sorting (FACS) analysis. Only the experiments in which the efficiency of transfection between the samples varied less than 3% were included in the quantitative analysis of HDV replication.

5.15 RNA isolation from transfected cells

At indicated days post transfection total cellular RNA was isolated by TRIzol (GibcoBRL) as recommended by the manufacturer.

5.16 Detection of replicating HDV RNA by hyb-shift assay

High specific activity HDV RNA probes for hyb-shift assay were synthesized from linearized plasmids using T3 or T7 RNA polymerase in the presence of 40mM

Tris-HCl pH8, 8mM MgCl₂, 2mM spermidine, 50mM NaCl, 10mM DTT, 1mM of each, ATP, CTP and GTP and 6.25 μ M α -³²P- UTP (800 Ci/mmol, New England Nuclear). The genomic- and antigenomic-specific probes encompass the HDV sequences between pos. 780-960 and pos. 650-780, respectively. Total cellular RNA, 5 μ g from the Cos7 or 15 μ g from the HeLa cells was incubated with 50% formamide, 40mM PIPES pH6.7, 0.4M NaCl, 1mM EDTA and 50K cpms of the HDV specific RNA probe for 18 hrs at 65°C. The RNA was then ethanol precipitated, dissolved in 5 μ l FA dye and resolved in 3.5% polyacrylamide/8M urea gel. Dried gels were scanned by PhosphorImager (Molecular Dynamics) and the signals corresponding to the HDV RNA hybrids were quantified by ImageQuant software.

5.17 Quantification of the relative replication efficiency of HDV mutants

HDV RNA replication was calculated for each, the wild type and mutant HDV RNA as an area below the corresponding kinetic curves defined by intensity of the hyb-shift signal at days 4, 6 and 8. The relative replication of each mutant was calculated as the ratio of the mutant to the wild type HDV RNA replication. The final value represents a mean of three independent experiments.

Chapter 6

BIBLIOGRAPHY

- Altmann, C. R., Solow-Cordero, D. E., and Chamberlin, M. J. (1994). RNA cleavage and chain elongation by *Escherichia coli* DNA-dependent RNA polymerase in a binary enzyme/RNA complex. *Proc Natl Acad Sci USA* *91*, 3784-3788.
- Archambault, J., Lacroute, F., Ruet, A., and Friesen, J. D. (1992). Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Mol Cell Biol* *12*, 4142-4152.
- Archambault, J., Pan, G., Dahmus, G. K., Cartier, M., Marshall, N., Zhang, S., Dahmus, M. E., and Greenblatt, J. (1998). FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* *283*, 28593-601.
- Artsimovitch, I., and Landick, R. (1998). Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev* *12*, 3110-22.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994). *Current protocols in molecular biology*. Wiley Interscience Press).
- Bartolomei, M. S. , and Corden, J. C. (1997). Localization of an α -amanitin resistance mutation in the gene encoding the largest subunit of mouse RNA polymerase II. *MCB* *7*, 586-594
- Beard, M. R., MacNaughton, T. B., and Gowans, E. J. (1996). Identification and characterization of a hepatitis delta virus RNA transcription promoter. *J Virol* *70*, 4986-4995.
- Beaudry, D., Busiere, F., Lareau, F., Lessard, C., and Perreault, J. P. (1995). The RNA of both polarities of the peach latent mosaic viroid self- cleaves in vitro solely by single hammerhead structures. *Nucleic Acids Res* *23*, 745-52.
- Bengal, E., and Aloni, Y. (1998). A block of transcription elongation by RNA polymerase II at synthetic sites in vitro. *J Biol Chem* *264*, 9791-8.
- Bentley, D. (1999). Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr Opin Cell Biol* *11*, 347-51.
- Bergmann, K. F., and Gerin, J. L. (1986). Antigens of hepatitis delta virus in the liver and serum of humans and animals. *J Infect Dis* *154*, 702-6.

- Bjorklund, S., and Kim, Y.-J. (1996). Mediator of transcriptional regulation. *TIBS* 21, 335-337.
- Boege, F., Rohde, W., and Sanger, H. L. (1982). In vitro transcription of viroid RNA into full-length copies by RNA-dependent RNA polymerase from healthy tomato leaf tissue. *Biosci Rep* 2, 185-194.
- Bonino, F., Heermann, K. H., Rizzetto, M., and Gerlich, W. H. (1986). Hepatitis delta virus: protein composition of delta antigen and its hepatitis B virus-derived envelope. *J Virol* 58, 945-950.
- Branch, A. D., and Robertson, H. D. (1984). A replication cycle for viroids and other small infectious RNAs. *Science* 223, 450-5.
- Branch, A. D., and Robertson, H. D. (1991). Efficient trans cleavage and a common structural motif for the ribozymes of the human hepatitis δ agent. *Proc Natl Acad Sci USA* 88, 10163-10167.
- Brazas, R., and Ganem, D. (1996). A cellular homolog of hepatitis delta antigen: implications for viral replication and evolution. *Science* 284, 90-94.
- Buratowski, S., Hanh, S., Guarente, L., and Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56, 549-561.
- Burley, S. K., and Roeder, R. G. (1996). Biochemistry and structural biology of transcription factor IID. *Annu Rev Biochem* 65, 769-799.
- Bussiere, F., Ouellet, J., Cote, F., Levesque, D., and 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, 2647-5.
- Casey, J. L., and Gerin, J. L. (1995). Hepatitis D virus RNA editing: specific modification of adenosine in the antigenomic RNA. *J Virol* 69, 7593-600.
- Chamberlin, M. J. (1974). The selectivity of transcription. *Annu Rev Biochem* 43, 721-775.
- Chang, F. L., Chen, P. J., Tu, S. J., Wang, C. J., and Chen, D. S. (1991). The large form of hepatitis delta antigen is crucial for assembly of hepatitis delta virus. *Proc Natl Acad Sci U S A* 88, 8490-4.

- Chang, M., Sun, C., Chen, C., and Chang, S. C. (1993). Functional Motifs of Delta Antigen Essential for RNA Binding and Replication of Hepatitis Delta Virus. *Journal of Virology* 67, 2529-2536.
- Chang, M. F., Baker, S. C., Soe, L. H., Kamahora, T., Keck, J. G., Makino, S., Govindarajan, S., and Lai, M. M. (1988). Human hepatitis delta antigen is a nuclear phosphoprotein with RNA-binding activity. *J Virol* 62, 2403-10.
- Chao, M., Hsieh, S. Y., and Taylor, J. (1991). The antigen of hepatitis delta virus: examination of in vitro RNA-binding specificity. *J Virol* 65, 4057-62.
- Chen, P. J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., and Taylor, J. (1986). Structure and replication of the genome of the hepatitis delta virus. *Proc Natl Acad Sci USA* 83, 8774-8.
- Chou, H. C., Hsieh, T. Y., Sheu, G. T., and Lai, M. M. (1998). Hepatitis delta antigen mediates the nuclear import of hepatitis delta virus RNA. *J Virol* 72, 3684-90.
- Ciesiolka, J., Michalowski, D., Wrzesinski, J., Krajewski, J., and Krzyzosiak, W. J. (1998). Patterns of cleavages induced by lead ions in defined RNA secondary structure motifs. *J Mol Biol* 285, 211-20.
- Conaway, J. W., Shilatifard, A., Dvir, A., and Conaway, R. C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem Sci* 25, 375-380.
- Cramer, P., Bushnell, D. E., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. P., Edwards, A. M., David, P. R., and Kornberg, R. D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640-9.
- Dahmus, M. E. (1995). Phosphorylation of the C-terminal domain of RNA polymerase II. *Bioch Bioph Acta* 1261, 171-182.
- Daros, J. A., Marcos, J. F., Hernandez, C., and Flores, R. (1994). Replication of avocado sunblotch viroid: evidence for a symmetric pathway with two rolling circles and hammerhead ribozyme processing. *Proc Natl Acad Sci U S A* 91, 12813-7.
- Diener, T. O. (1979). Viroid discovery. *Science* 206, 886.
- Diener, T. O. (1987). Autonomous and helper-dependent small pathogenic RNAs of plants: viroids and satellites. *Prog Clin Biol Res* 234, 3-18.

- Diener, T. O. (1993). The viroid: big punch in a small package. *Trends Microbiol* **1**, 289-94.
- Diener, T. O. (1995). Origin and evolution of viroids and viroid-like satellite RNAs. *Virus Genes* **11**, 119-131.
- Dignam, J. D., Martin, P. L., Shastry, B. S., and Roeder, R. G. (1983). Eukaryotic gene transcription with purified components. *Meth Enzymol* **101**, 582-598.
- Erie, D. A., Hajiseyedjavadi, O., Young, M. C., and Hippel, P. H. v. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* **262**, 867-873.
- Feaver, W. J., Gileadi, O., Li, Y., and Kornberg, R. D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell* **67**, 1223-1230.
- Flores, R., and Semancik, J. S. (1982). Properties of a cell-free system for synthesis of citrus exocortis viroid. *Proc Natl Acad Sci USA* **79**, 6285-6288.
- Fu, T.-B., and Taylor, J. (1993). The RNAs of hepatitis delta virus are copied by RNA polymerase II in nuclear homogenates. *J Virol* **67**, 6965-6972.
- Garber, M. E., and Jones, K. A. (1999). HIV-1 Tat: coping with negative elongation factors. *Curr Opin Immunol* **11**, 460-5.
- Glenn, J. S., Taylor, J. M., and White, J. M. (1990). In vitro-synthesized hepatitis delta virus RNA initiates genome replication in cultured cells. *J Virol* **64**, 3104-3107.
- Glenn, J. S., and White, J. M. (1991). trans-dominant inhibition of human hepatitis delta virus genom replication. *J Virol* **65**, 2357-61.
- Goodman, T. C., Nagel, L., Rappold, W., Klotz, G., and Riesner, D. (1984). Viroid replication: equilibrium association constant and comparative activity measurements for the viroid-polymerase interaction. *Nucleic Acids Res* **12**, 6231-46.
- Gottardo, M. (1998) Molecular characterization of an RNA-dependent transcription reaction catalyzed by RNA polymerase II. The Rockefeller University, New York.
- Gross, H. J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H., and Sanger, H. L. (1978). Nucleotide sequence and secondary structure of potato spindle tuber viroid. *Nature* **283**, 203-8.

Gross, H. J., Krupp, G., Domdey, H., Raba, M., Jank, P., Lossow, C., Alberty, H., Ramm, K., and Sanger, H. L. (1982). Nucleotide sequence and secondary structure of citrus exocortis and chrysanthemum stunt viroid. *Eur J Biochem* 121, 249-57.

Gu, W., Powell, W., Mote, J. J., and Reines, D. (1993). Nascent RNA cleavage by arrested RNA polymerase II does not require upstream translocation of the elongation complex on DNA. *J Biol Chem* 268, 25604-25616.

Gu, W., and Reines, D. (1995). Variation in the size of nascent RNA cleavage products as a function of transcript length and elongation competence. *J Biol Chem* 280, 30441-30447.

Gudima, S., Dingle, K., Wu, T. T., Moraleda, G., and Taylor, J. (1999). Characterization of the 5' ends for polyadenylated RNAs synthesized during the replication of hepatitis delta virus. *J Virol* 73, 6533-9.

Gudima, S., Wu, S. Y., Chiang, C. M., Moraleda, G., and Taylor, J. (2000). Origin of hepatitis delta virus mRNA. *J Virol* 74, 7204-10.

Hammond, R., Smith, D. R., and Diener, T. O. (1989). Nucleotide sequence and proposed secondary structure of Columnea latent viroid: a natural mosaic of viroid sequences. *Nucleic Acids Research* 17, 10083-10094.

Hawley, D. K., and Roeder, R. G. (1985). Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. *J Biol Chem* 260, 8163-8172.

Hernandez, C., and Flores, R. (1992). Plus and minus RNAs of peach latent mosaic viroid self-cleave in vitro via hammerhead structures. *Proc Natl Acad Sci USA* 89, 3711-5.

Horikoshi, M., Bertuccioli, C., Takada, R., Wang, J., Yamamoto, T., and Roeder, R. G. (1992). Transcription Factor TFIID induces DNA bending upon binding to the TATA element. *Proc Natl Acad Sci USA* 89, 1060-1064.

Horikoshi, M., Wang, C. K., Fujii, H., Cromlish, J. A., Weil, P. A., and Roeder, R. G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* 341, 299-303.

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

Hutchins, C. J., Rathjen, P. D., Forster, A. C., and Symons, R. H. (1986). Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res* 14, 3628-40.

Isambert, H., and Siggia, E. D. (2000). Modeling RNA folding paths with pseudoknots: application to hepatitis delta virus ribozyme. *Proc Natl Acad Sci USA* 97, 6515-20.

Izban, M. G., and Luse, D. S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'-to-5' direction in the presence of elongation factor SII. *Genes Dev* 6, 1342-1356.

Izban, M. G., and Luse, D. S. (1993). The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. *J Biol Chem* 268, 12873-12885.

Izban, M. G., Parsons, M. A., and Sinden, R. R. (1998). Template end-to-end transposition by RNA polymerase II. *J Biol Chem* 283, 28009-16.

Johnson, T. L., and Chamberlin, M. J. (1994). Complexes of yeast RNA polymerase II and RNA are substrates for TFIIIS-induced RNA cleavage. *Cell* 77, 217-224.

Kao, S. Y., Calman, A. F., Luciw, P. A., and Peterlin, B. M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330, 489-93.

Kim, J. L., and Burley, S. K. (1994). 1.9-Angstrom resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Struct Biol* 1, 638-653.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993a). Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365, 520-528.

Kim, Y., Geiger, J. H., Hahn, S., and Sigler, P. B. (1993b). Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365, 512-520.

Kim, Y.-J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase. *Cell* 77, 599-608.

Kofalvi, S. A., Marcos, J. F., Canizares, M. C., Pallas, V., and Candresse, T. (1997). Hop stunt viroid (HSVd) sequence variants from *Prunus* species:

- evidence for recombination between HSVd isolates. *Journal of General Virology* 78, 3177-3186.
- Koleske, A. J., and Young, R. A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* 368, 466-369.
- Koltunow, A. M., and Rezaian, M. A. (1989). Grapevine viroid 1B, a new member of the apple scar skin viroid group contains the left terminal region of tomato planta macho viroid. *Virology* 170, 575-578.
- Konarska, M. M., Padgett, R. A., and Sharp, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* 38, 731-6.
- Kos, A., Dijkema, R., Arnberg, A. C., van der Maide P. H., and Schellekens, H. (1986). The hepatitis delta virus possesses a circular RNA. *Nature* 323, 558-60.
- Kretzschmar, M., Kaiser, K., Lottspeich, F., and Meisterernst, M. (1994). A novel mediator of class II gene transcription with homology to viral immediate-early transcriptional regulators. *Cell* 78, 525-534.
- Krumm, A., Hickey, L. B., and Groudine, M. (1995). Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation. *Genes Dev* 9, 559-72.
- Kuo, M. Y.-P., Chao, M., and 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.
- Lafontaine DA, Deschenes P, Bussiere F, Poisson V, Perreault JP (1999). The viroid and viroid-like RNA database. *Nucleic Acids Res* 27, 186-7
- Lai, M. M. C. (1995). The molecular biology of hepatitis delta virus. *Annu Rev Biochem* 64, 259-286.
- Landick, R., Jr., C. L. T., and Yanofsky, C. (1996). Transcription attenuation. In *Escherichia coli* and *Salmonella*: cellular and molecular biology. F. V. Neidhardt, ed. (Washington, D.C., ASM Press), pp. 1263-1286.
- Lazinski, D., Grzadzińska, E., and Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* 59, 207-218.
- Lazinski, D. W., and Taylor, J. M. (1993). Relating Structure to Function in the Hepatitis Delta Virus Antigen. *Journal of Virology* 67, 2672-2680.

Lazinski, D. W., and Taylor, J. M. (1994). Expression of Hepatitis Delta Virus RNA Deletions: cis and trans Requirements for Self-Cleavage, Ligation, and RNA Packaging. *Journal of Virology* 68, 2879-2888.

Lee, C.-Z., Lin, J.-H., Chao, M., McKnight, K., and Lai, M. M. C. (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, 2221-2228.

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

Linn, S. C., and Luse, D. S. (1991). RNA polymerase II elongation complexes paused after the synthesis of 15- or 35-base transcripts have different structures. *Mol Cell Biol* 11, 1508-22.

Lis, J., and Wu, C. (1993). Protein traffic on the heat-shock promoter: parking, stalling, and trucking along. *Cell* 74, 1-4.

Liu, Y. H., and Symons, R. H. (1998). Specific RNA self-cleavage in coconut cadang cadang viroid: potential for a role in rolling circle replication. *RNA* 4, 418-29.

Lu, H., Zawel, L., Fisher, L., Egly, J.-M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* 358, 641-645.

Luo, G., Chao, M., Hsieh, S. Y., Sureau, C., Nishikura, K., and Taylor, J. (1990). A Specific Base Transition Occurs on Replicating Hepatitis Delta Virus RNA. *J Virol* 64, 1021-1028.

MacNaughton, T. B., Wang, Y.-J., and Lai, M. M. C. (1993). Replication of hepatitis delta virus RNA: effects of mutations of the autocatalytic cleavage sites. *J Virol* 67, 2228-2234.

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

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

Makino, S., Chang, M. F., Shieh, C. K., Kamahora, T., Vannier, D. M., Govindarajan, S., and Lai, M. M. (1987). Molecular cloning and sequencing of a human hepatitis delta virus RNA. *Nature* 329, 343-6.

Maldonado, E., Ahiekhattar, R., Sheldon, M., Cho, H., Drapkni, R., Rickert, P., Lees, E., Anderson, C. W., Linn, S., and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA-repair proteins. *Nature* 381, 86-89.

Marshall, N. F., Peng, J., Xie, Z., and Price, D. H. (1996). Control of RNA polymerase II elongation potential by a novel carboxy-terminal domain kinase. *J Biol Chem* 281, 28176-28183.

Mason, S. W., Li, J., and Greenblatt, J. (1992). Host factor requirements for processive antitermination of transcription and suppression of pausing by the N protein of bacteriophage λ . *J Biol Chem* 267, 19418-19426.

Matsui, T., Segall, J., Weil, P. A., and Roeder, R. G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 255, 11992-11996.

Matsuzaki, H., Kassavetis, G. A., and Geiduschek, E. P. (1994). Analysis of RNA chain elongation and termination by *Saccharomyces cerevisiae* RNA polymerase III. *J Mol Biol* 235, 1173-92.

Matysiak, M., Wrzesinski, J., and Ciesiolka, J. (1999). Sequential folding of the genomic ribozyme of the hepatitis delta virus: structural analysis of RNA transcription intermediates. *J Mol Biol* 291, 283-94.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* 11, 306-18.

McCracken, S., Rosonina, E., Fong, N., Sikes, M., Beyer, A., O'Hare, K., Shuman, S., and Bentley, D. (1998). Role of RNA polymerase II carboxy-terminal domain in coordinating transcription with RNA processing. *Cold Spring Harb Symp Quant Biol* 63, 301-9.

Meininghaus, M., Chapman, R. D., Horndasch, M., and Eick, D. (2000). Conditional Expression of RNA Polymerase II in Mammalian Cells. DELETION OF THE CARBOXYL-TERMINAL DOMAIN OF THE LARGE SUBUNIT

AFFECTS EARLY STEPS IN TRANSCRIPTION. *J Biol Chem* 285, 24375-24382.

Misteli, T., and Spector, D. L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol Cell* 3, 697-705.

Modahl, L. E., and Lai, M. M. (2000). The large delta antigen of hepatitis delta virus potently inhibits genomic but not antigenomic RNA synthesis: a mechanism enabling initiation of viral replication. *J Virol* 74, 7375-80.

Modahl, L. E., Macnaughton, T. B., Zhu, N., Johnson, D. L., and 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.

Moore, P. B. (1999). Structural motifs in RNA. *Annu Rev Biochem* 68, 287-300.

Muesing, M. A., Smith, D. H., and Capon, D. J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus trans- activator protein. *Cell* 48, 691-701.

Mühlbach, H.-P., and Sängner, H. (1979). Viroid replication is inhibited by alpha-amanitin. *Nature* 288, 185-188.

Muthukrishnan, S., Both, G. W., Furuichi, Y., and Shatkin, A. J. (1975). 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* 255, 33-37.

Nagy, P. D., Pogany, J., and Simon, A. E. (1999). RNA elements required for RNA recombination function as replication enhancers in vitro and in vivo in a plus-strand RNA virus. *EMBO* 18, 5653-65.

Nagy, P. D., and Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* 235, 1-9.

Nagy, P. D., and Simon, A. E. (1998). In vitro characterization of late steps of RNA recombination in turnip crinkle virus. The role of the priming stem and flanking sequences. *Virology* 249, 393-405.

Nakajima, N., Horikoshi, M., and Roeder, R. G. (1988). Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. *Mol Cell Biol* 8, 4028-4040.

- Nikolov, D. B., and Burley, S. K. (1997). RNA polymerase II transcription initiation: A structural view. *Proc Natl Acad Sci USA* *94*, 15-22.
- Nikolov, D. B., Chen, H., Halay, E. D., Hoffmann, A., Roeder, R. G., and Burley, S. K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci USA* *93*, 4956-4961.
- Nudler, E. (1999). Transcription elongation: structural basis and mechanisms. *J Mol Biol* *288*, 1-12.
- Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription Processivity: Protein-DNA Interactions Holding Together the Elongation Complex. *Science* *283*, 211-217.
- Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A. (1995). Coupling between transcription termination and RNA polymerase inchworming. *Cell* *81*, 351-357.
- O'Brien, T., and Lis, H. T. (1991). RNA polymerase II pauses at the 5' end of the transcriptionally induced *Drosophila hsp70* gene. *Mol Cell Biol* *11*, 5285-5290.
- Orphanides, G., Lagrange, T., and Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes Dev* *10*, 2657-2683.
- Ossipow, V., Tassan, J.-P., Nigg, E. A., and Schibler, U. (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. *Cell* *83*, 137-146.
- Owens, R. A., and Diener, T. O. (1977). Synthesis of RNA complementary to potato spindle tuber viroid using Q β replicase. *Virology* *79*, 109-120.
- Palangat, M., Meier, T. I., Keene, R. G., and Landick, R. (1998). Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure. *Mol Cell* *1*, 1033-42.
- Parada, C. A., and Roeder, R. G. (1996). Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain. *Nature* *384*, 375-378.
- Parker, C. S., and Topol, J. (1984). A *Drosophila* RNA polymerase II transcription factor contains promoter-region-specific DNA-binding activity. *Cell* *36*, 357-369.
- Pelchat, M., Levesque, D., Ouellet, J., Laurendeau, S., Levesque, S., Lehoux, J., Thompson, D. A., Eastwell, K. C., Skrzeczowski, L. J., and Perreault, J. P. (2000). Sequencing of peach latent mosaic viroid variants from nine North

- American peach cultivars shows that this RNA folds into a complex secondary structure. *Virology* 281, 37-45.
- Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991). Thermodynamic study of internal loops in oligoribonucleotides: symmetric loops are more stable than asymmetric loops. *Biochemistry* 30, 6428-36.
- Perrotta, A. T., and Been, M. D. (1990). The self-cleaving domain from the genomic RNA of hepatitis delta virus: sequence requirements and the effects of denaturant. *Nuc Acids Res* 18, 6821-6828.
- Platt, T. (1998). RNA Structure in Transcription Elongation, Termination, and Antitermination. In *RNA structure and function*, R. W. Simons, and M. Grunberg-Manago, eds. (Plienvew, New York, Cold Spring Harbor Laboratory Press), pp. 541-574.
- Polson, A. G., Bass, B. L., and 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., Bass, B. L., and 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, 1919-26.
- Powell, W., Bartholomew, B., and Reines, D. (1996). Elongation factor SII contacts the 3'-end of RNA in the RNA polymerase II elongation complex. *J Biol Chem* 281, 22301-22304.
- Rackwitz, H.-R., Rohde, W., and Sanger, H. L. (1981). DNA-dependent RNA polymerase II of plant origin transcribes viroid RNA into full-length copies. *Nature* 291, 297-301.
- Reeder, T. C., and Hawley, D. K. (1996). Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. *Cell* 87, 767-777.
- Reid, C. E., and Lazinski, D. W. (2000). A host-specific function is required for ligation of a wide variety of ribozyme-processed RNAs. *Proc Natl Acad Sci USA* 97, 424-9.
- Reinberg, D., Horikoshi, M., and Roeder, R. G. (1987a). Factors involved in specific transcription in mammalian RNA polymerase: functional analysis of initiation factors IIA and IID and identification of a new factor operating at sequences downstream of the initiation site. *J Biol Chem* 262, 3322-3330.

- Reinberg, D., and Roeder, R. G. (1987b). Factors involved in specific transcription by mammalian RNA polymerase II: transcription factor IIS stimulates elongation of RNA chains. *J Biol Chem* 262, 3331-3337.
- Reines, D. (1992). Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J Biol Chem* 267, 3795-3800.
- Reines, D. (1994). Nascent RNA cleavage by transcription elongation complexes. In *Transcription: mechanisms and regulation*. R. C. Conaway, and J. W. Conaway, eds. (New York, Raven Press, Ltd), pp. 263-288.
- Reines, D., Conaway, R. C., and Conaway, J. W. (1999). Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Curr Opin Cell Biol* 11, 342-6.
- Reines, D., Wells, D., Chamberlin, M. J., and Kane, C. M. (1987). Identification of intrinsic termination sites in vitro for RNA polymerase II within eukaryotic gene sequences. *J Mol Biol* 196, 299-312.
- Rezaian, M. A. (1990). Australian grapevine viroid-evidence for extensive recombination between viroids. *Nucleic Acids Research* 18, 1813-1818.
- Richardson, J. P., and Greenblatt, J. (1996). Control of RNA chain elongation and termination. In *Escherichia coli and Salmonella: cellular and molecular biology*. F. V. Neidhardt, ed. (Washington, D.C., ASM Press), pp. 822-848.
- Riesner, D., and Gross, H. J. (1985). Viroids. *Annu Rev Biochem* 54, 531-64.
- Rivera-Bustamante, R. F., and Semancik, J. S. (1989). Properties of a viroid-replicating complex solubilized from nuclei. *J Gen Virol* 70, 2807-2816.
- Rizzetto, M., Canese, MG., 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.
- Rizzetto, M., Hoyer, B., Canese, M. G., Shih, J. W., Purcell, R. H., and Gerin, J. L. (1980). Delta Agent: association of delta antigen with hepatitis B surface antigen and RNA in serum of delta-infected chimpanzees. *Proc Natl Acad Sci U S A* 77, 6124-8.
- Roeder, R. G. (1976). Eukaryotic nuclear RNA polymerases. In *RNA polymerase*. R. Losick, and M. J. Chamberlin, eds. (Cold Spring Harbor, New York, Cold Spring Harbor Laboratory), pp. 285-330.

- Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* 21, 328-335.
- Roeder, R. G., and Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224, 234-237.
- Rohde, W., Rackwitz, H. R., Boege, F., and Sanger, H. L. (1982). Viroid RNA is accepted as a template for in vitro transcription by DNA-dependent DNA polymerase I and RNA polymerase from *Escherichia coli*. *Biosci Rep* 2, 929-39.
- Rong, M., Durbin, R. K., and McAllister, W. T. (1998). Template Strand Switching by T7 RNA Polymerase. *J Biol Chem* 283, 10253-10260.
- Rosenstein, S. P., and Been, M. B. (1990). Self-cleavage of hepatitis delta virus genomic strand RNA is enhanced under partially denaturing conditions., *Biochemistry* 29, 8011-6.
- Rudd, M. D., Izban, M. G., and Luse, D. S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc Natl Acad Sci USA* 91, 8057-8061.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press).
- Schiebel, W., Pelissier, T., Riedel, L., Thalmeir, S., Schiebel, R., Kempe, D., Lottspeich, F., Sanger, H. L., and Wassenegger, M. (1998). Isolation of an RNA-directed RNA polymerase-specific cDNA clone from tomato. *Plant Cell* 10, 2087-101.
- Schnapp, G., Graveley, B. R., and Grummt, I. (1996). TFIIS binds to mouse RNA polymerase I and stimulates transcript elongation and hydrolytic cleavage of nascent rRNA. *Mol Gen Genet* 252, 412-419.
- Schroeder, S. J., and Turner, D. H. (2000). Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA. *Biochemistry* 39, 9257-74.
- Semancik, J. S., and 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 USA* 81, 4429-4433.
- Semancik, J. S., Szychowski, J. A., Rakowski, A. G., and Symons, R. H. (1994). A stable 463 nucleotide variant of citrus exocortis viroid produced by terminal repeats. *Journal of General Virology* 75, 728-732.

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

Sharp, P. A., and Marciniak, R. A. (1989). HIV TAR: an RNA enhancer? *Cell* 59, 229-230.

Shiekhattar, R., Mermelstein, F., Fisher, R. P., Drapkin, R., Dynlacht, B., Wessling, H. C., Morgan, D. O., and Reinberg, D. (1995). Cdk-activating kinase complex is a component of human transcription factor TFIID. *Nature* 374, 283-287.

Simons, R. W., and Grunberg-Manago, M., eds. (1998). RNA structure and function. (Plenum, New York, Cold Spring Harbor Laboratory Press).

Spiesmacher, E., Mühlbach, H.-P., Tabler, M., and Sanger, H. L. (1985). Synthesis of (+) and (-) RNA molecules of potato spindle tuber viroid (PSTV) in isolated nuclei and its impairment by transcription inhibitors, *Biosci Rep* 5, 251-265.

Struhl, K. (1996). Chromatin structure and RNA polymerase II connection: implications for transcription. *Cell* 84, 179-182.

Symons, R. H. (1997). Plant pathogenic RNAs and RNA catalysis. *Nucleic Acids Res* 25, 2683-9.

Taylor, J. M. (1992). The structure and replication of hepatitis delta virus. *Annu Rev Microb* 46, 253-76.

Thomas, M. J., Platas, A. A., and Hawley, D. K. (1998). Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93, 628-637.

Uptain, S. M., Kane, C. M., and Chamberlin, M. J. (1997). Basic mechanisms of transcript elongation and its regulation. *Ann Rev Biochem* 66, 117-172.

Wada, T., Orphanides, G., Hasegawa, J., Kim, D. K., Shima, D., Yamaguchi, Y., Fukuda, A., Hisatake, K., Oh, S., Reinberg, D., and Handa, H. (2000). FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIID. *Mol Cell* 5, 1067-72.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F., Buratowski, S., and Handa, H. (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II

- processivity, is composed of human Spt4 and Spt5 homologs. *Genes & Development* **12**, 343-356.
- Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. *EMBO J* **17**, 7395–7403.
- Wang, D., Severinov, K., and Landick, R. (1997a). Preferential interaction of the *his* pause RNA hairpin with RNA polymerase β subunit residues 904-950 correlates with strong transcriptional pausing. *Proc Natl Acad Sci USA* **94**, 8433-8438.
- Wang, H. W., Wu, H. L., Chen, D. S., and Chen, P. J. (1997b). 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., and Houghton, M. (1986). Structure, sequence and expression of the hepatitis delta (delta) viral genome [published erratum appears in *Nature* 1987 Jul 30-Aug 5;328(6129):456], *Nature* **323**, 508-14.
- Warrilow, D., and 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**, 2367-75.
- Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979). Selective and accurate initiation of transcription at the Ad2 Major Late Promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* **18**, 469-484.
- Wind, M., and Reines, D. (2000). Transcription elongation factor SII. *BioEssays* **22**, 328-336.
- Wu, H.-N., and Lai, M. M. (1989). Reversible cleavage and ligation of hepatitis delta virus RNA. *Science* **243**, 652-654.
- Wu, H.-N., Lin, Y.-J., Lin, F.-P., Makino, S., Chang, M.-F., and Lai, M. M. (1989). Human hepatitis δ virus RNA subfragments contain an autocleavage activity. *Proc Natl Acad Sci USA* **86**, 1831-1835.
- Wu, J., Awrey, D. E., Edwards, A. M., Archambault, J., and Friesen, J. D. (1996). In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS. *Proc Natl Acad Sci USA* **93**, 11552-11557.

- Wu, N. H., and Huang, Z. S. (1992). Mutagenesis analysis of the self-cleavage domain of hepatitis delta virus antigenomic RNA. *Nucleic Acids Res* 20, 5937-41.
- Wu, T. T., Netter, H. J., Lazinski, D. W., and 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, 5408-14.
- Xia, Y. P., and 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, 6641-8.
- Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97, 41-51.
- Yankulov, K., Blau, J., Purton, T., Roberts, S., and Bentley, D. L. (1994). Transcriptional elongation by RNA polymerase II is stimulated by transactivators. *Cell* 77, 749-759.
- Yankulov, K., Yamashita, K., Roy, R., Egly, J.-M., and Bentley, D. L. (1995). The transcriptional elongation inhibitor 5,6-dichloro-1- β -D-ribofuranosylbenzimidazole inhibits transcription factor IIH-associated protein kinase. *J Biol Chem* 280, 23922-23925.
- Yoo, O., Yoon, H., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Agarwal, K. (1991). Cloning, expression and characterization of the human transcription factor, TFIIS. *Nuc Acids Res* 19, 1073-1079.